Chapter 6
Self-optimising Production Systems

Robert Schmitt, Christian Brecher, Burkhard Corves, Thomas Gries, Sabina Jeschke, Fritz Klocke, Peter Loosen, Walter Michaeli, Rainer Müller, Reinhard Poprawe, Uwe Reisgen, Christopher M. Schlick, Günther Schuh, Thomas Auerbach, Fabian Bauhoff, Marion Beckers, Daniel Behnen, Tobias Brosze, Guido Buchholz, Christian Büscher, Urs Eppelt, Martin Esser, Daniel Ewert, Kamil Fayzullin, Reinhard Freudenberg, Peter Fritz, Sascha Fuchs, Yves-Simon Gloy, Sebastian Haag, Eckart Hauck, Werner Herfs, Niklas Hering, Mathias Hüsing, Mario Isermann, Markus Janßen, Bernhard Kausch, Tobias Kempf, Stephan Kratz, Sinem Kuz, Matthis Laass, Juliane Lose, Adam Malik, Marcel Ph. Mayer, Thomas Molitor, Simon Müller, Barbara Odenthal, Alberto Pavim, Dirk Petring, Till Potente, Nicolas Pyschny, Axel Reßmann, Martin Riedel, Simone Runge, Heiko Schenuit, Daniel Schilberg, Wolfgang Schulz, Maik Schürmeyer, Jens Schüttler, Ulrich Thombansen, Dražen Veselovac, Matthias Vette, Carsten Wagels and Konrad Willms

Contents

R. Schmitt (✉)
Werkzeugmaschinenlabor WZL der RWTH Aachen, Steinbachstr. 19,
52074 Aachen, Germany
e-mail: r.schmitt@wzl.rwth-aachen.de

C. Brecher (ed.), *Integrative Production Technology for High-Wage Countries,*
DOI 10.1007/978-3-642-21067-9_6, © Springer-Verlag Berlin Heidelberg 2012

6.1 Research Program on Self-optimising Production Systems

Robert Schmitt, Mario Isermann and Carsten Wagels

One of the central success factors for production in high-wage countries is the solution of the conflict that can be described with the term “planning efficiency”. Planning efficiency describes the relationship between the expenditure of planning and the profit generated by these expenditures. From the viewpoint of a successful business management, the challenge is to dynamically find the optimum between detailed planning and the immediate arrangement of the value stream. Planning-oriented approaches try to model the production system with as many of its characteristics and parameters as possible in order to avoid uncertainties and to allow rational decisions based on these models. The success of a planning-oriented approach depends on the transparency of business and production processes and on the quality of the applied models. Even though planning-oriented approaches are supported by a multitude of systems in industrial practice, an effective realisation is very intricate, so these models with their inherent structures tend to be matched to a current stationary condition of an enterprise. Every change within this enterprise, whether inherently structural or driven by altered input parameters, thus requires continuous updating and adjustment. This process is very cost-intensive and time-consuming; a direct transfer onto other enterprises or even other processes within the same enterprise is often impossible. This is also a result of the fact that planning usually occurs a priori and not in real-time. Therefore it is hard for completely planning-oriented systems to react to spontaneous deviations because the knowledge about those naturally only comes a posteriori.

As a reaction to the observation that in a networked enterprise non-stationary, transient Ramp-Up and Ramp-Down procedures are the rule rather than the exception, with the change of the predominant paradigm of Taylorism successful approaches were developed that are oriented at the so-called value stream of an enterprise. In these approaches, central planning activities were reduced in favor of de-central activities and integrated into the value-adding process. The distinguishing mark of value-oriented production systems is their high degree of flexibility, as decisions are made depending on the situation. Still, the downside of purely value-oriented approaches is that much potential for optimization remains unused as it cannot be tapped without a holistic view and the according target orientation.

A successful enterprise that connects both planning-oriented and value-oriented approaches thus uses the advantages of both strategies. In order to achieve this, on the one hand it is necessary to exactly model the production process. On the other hand, by identifying the essential parameters to be influenced, a foundation has to be laid for the ability to autonomously and flexibly make decisions. However, the existing possibilities to optimize the behavior of one element within the whole system can be focused too tightly and bind resources even though in certain situations this could lead to an adverse behavior in other areas. This optimization task within a production system usually cannot be solved analytically.

The solution of this conflict becomes possible if a system is designed that can adjust its goals situatively. While in most cases the optimization of a system is controlled from the outside, e.g. by a person, in many cases the optimization by the technical system itself is a possible option. The developments in automatization technology show, however, that even in comparatively simple matters this has not yet been accomplished. Therefore in many cases people still play an important role. The implementation of self-optimizing abilities provides a substantial possibility to reduce the area of tension of planning efficiency.

The following research question thus is the focus of explorations within the research area Self-Optimizing Production Systems (Fig. 6.1):

> How can large and small production systems be harmonised through common models in terms of their control up to the point of self-optimisation?

Speaking of the term "self-optimizing systems" we understand systems that are able to effect independent ("endogenous") changes of their inner states or structure based on varying input conditions or interferences. In production processes, according target values can be e.g. capacities, number of pieces, quality, costs or processing times.

Self-optimizing systems are defined by the interaction of contained elements and the recurring execution of the actions (Adelt et al. 2009)

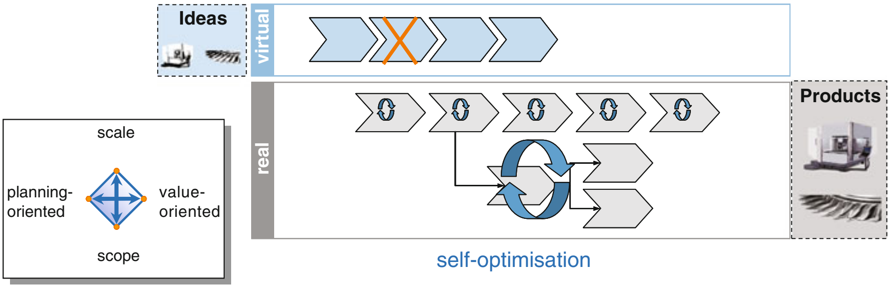


Fig. 6.1 Research question on self-optimizing production systems

- continuous analysis of the current situation,
- determination of targets, and
- adaptation of the system's behavior to achieve these targets.

From the perspective of a traditional control loop the behavior of the system is controlled by externally predetermined target parameters. If the control loop adapts control parameters to observed changes we speak of an adaptive system. The aspect of self-optimization puts the focus onto the dynamization of the target system. Compared to this traditional controlling a self-optimizing system is able to continuously determine the individual sub-targets based on internal decisions and to dynamically adapt the control path (Fig. 6.2).

In order to permit both the independent adaptation of targets and the change of the control path, e.g. of a production, assembly or even an organizational process (Fig. 6.3), self-optimizing production systems need an internal structure to register

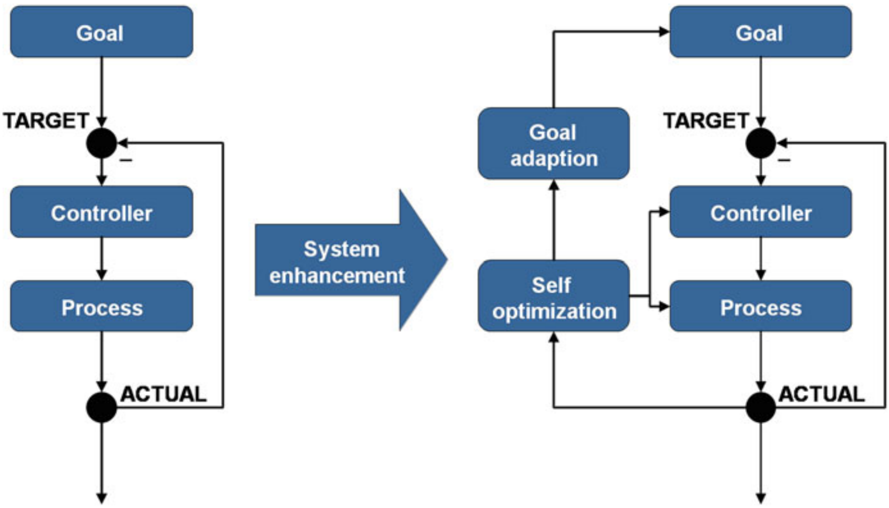


Fig. 6.2 From traditional controlling to self-optimization

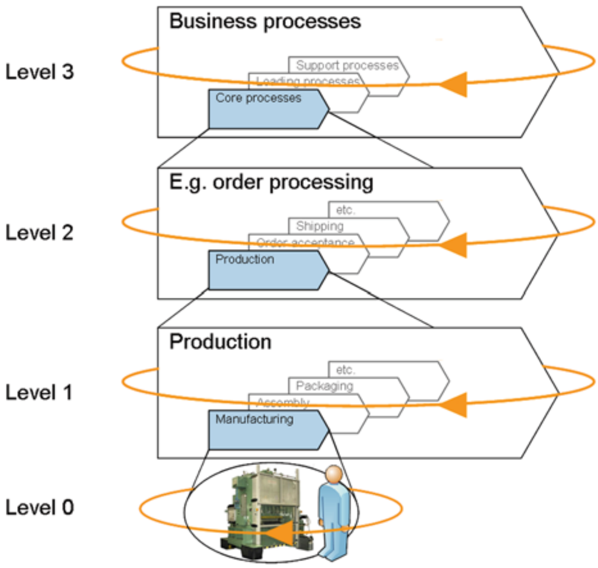


Fig. 6.3 Cross-level controlling in an enterprise

and analyze the current situation, to determine the system targets and to adapt system behavior. Self-optimizing systems are based on coupled sub-systems that can possess abilities of autonomy, versatility and cognition. Cognitive (sub-)systems are able to register information from their environment, to process it in a central processor and to convert it into behavior by being able to influence their environment.

Such a cognitive system can thus constitute the core of a self-optimizing production system. The transfer of cognitive mechanisms onto computer systems makes it possible to exactly process large quantities of data and thus to analyze even complex production processes with multi-level dependencies.

Therefore a self-optimizing system with its abilities to analyse data, to model and to make decisions offers an approach to master processes with no existing deterministic control function. It can basically be employed in various levels of an enterprise, from the direct control of machine parameters to assembly tasks and planning levels. Especially in cases where e.g. requirements of one-piece-flow or other non-analytical-deterministical cause-effect relationships prevail there are challenges to augment efficiency and cost-effectiveness through dynamic target synchronization of global and local target systems.

In production, a cognitive, self-optimizing system can intervene into a process and can concertedly vary individual values and tolerances in order to react to deviations in previous steps of the process. By doing this, scattering of the results of connected processes can be minimized and the quality of the end product can be augmented. Similarly, a cognitive system can act in the area of assembly that contains a high amount of manual labor, where the quality is decisively influenced by the experience-based knowledge of the workers and where aspects of assembly-compatibility of the construction also play a decisive role. Through specific matching of individual parts and through the variation of the according actuating variables from production resulting deviations can be compensated.

6.2 Integrative Self-optimising Process Chains

Robert Schmitt, Mario Isermann, Carsten Wagels and Matthis Laass

6.2.1 Abstract

6.2.1.1 Challenge

Only when a company can offer goods that satisfy the customer's requirements and for which a price can be obtained that will make a profit, can it be successful in the long run. The task of the management consists in optimising the product quality and thus the match of the customer's demands and the properties of the products produced. In addition to a highest possible liquidity ratio the other conditions to be taken into account in this optimisation task from a management standpoint include for example the pricing, production costs, the process through-put times or the flexibility of a complete production process as the basis for the ability to adapt rapidly and economically to changing governing conditions.

As a result consideration of Quality according to Garvin opens up further dimensions, which are to be taken into account when determining the target system. As a consequence a multi-dimensional target system is to be addressed during the optimisation process. In this case individual process optimisation alone is not expedient; in particular since the common optimum of the loss functions in the individual processes cannot be targeted in this way. Feedback of information from the individual processes to coordinate the whole process chain via overall control mechanisms offers a potentially successful approach. The challenge here consists in the developing the methodology and also specific applications for control and regulation.

6.2.1.2 Goal-Setting

The approach discussed here focuses upon the ability of production systems to adapt them-selves in self-optimising mode to changing conditions by using cognitive technologies. The main goal here is to ensure the product quality, wherein the process chains under consideration are aimed at fulfilling the required product functions by the design of pan-process closed-loop control systems. As a result the task of purely complying with tolerances of individual product components becomes a subordinate objective in the target system. This enables it (the target system) to permit deviations in the production system, dynamically during the production process, purposefully in favour of other targets (e.g. costs, quality, time) or to compensate these by targeted reactions, and increase its flexibility and hence also its competitiveness.

As a result it makes a contribution to the Cluster of Excellence "Integrated Production Technology for High Wage Countries to resolve the Polylemma of Production", while the dilemma between plan and value orientation by utilising self-optimising and cognitive technologies in the pan-process parameterisation of production systems is reduced.

The dichotomy Scale vs. Scope is being taken up by accelerating the identification and application of cause and effect relationships and the resulting associated faster deployment of these findings even with small quantities.

The influence of production parameters upon tolerances and consequently upon product quality will be considered along with the potential for making the production system more flexible by balancing the production system employing self-optimisation so that it will respond more rapidly to deviations in processes.

The aim is to fulfil the customers' requirements at lowest cost by taking into account possible constraints and at the same time safeguard the product quality economically in dynamic production environments.

Products and components are traditionally fabricated, geared towards dimensional, geometrical and positional tolerances and assessed in production, whereas customers primarily consider the production function.

The product function is determined essentially by geometric tolerances, but production is not considered explicitly. Function orientation is an approach that directs product features and hence the individual steps in the process chain towards the product function as the ultimate aim.

The product functions are awarded a higher priority than purely satisfying defined interim results of production steps, i.e. deviations and/or tolerances. Thus individual features may be varied as subordinate aims depending upon the given conditions. Deviations from previous processes can be compensated for in subsequent processes. This increases the flexibility of a production system and opens up additional potential for cost reduction, if costs for a tolerance can be set against the contribution to product function and consequently e.g. 'expensive' tolerances widened.

The objective is the ability of production to manufacture products based upon function and as a result to shift from production based upon tolerances to one based upon function.

If the target system is adapted and the derivation of control commands is carried out autonomously by continuous comparison with target conditions and the targets in the super-ordinate target system, this is described as self-optimisation.

With the combination of function orientation and self-optimisation a new approach for optimising production presents itself. It enables (the system) to respond autonomously to variations in processes and in the case of deviations in sub-targets of the target system to influence the remaining targets correctively so that attaining the primary target is still guaranteed.

This is however only possible if the processes are fully understood, how the processes integrate and which factors and production parameters influence the product parameters. The relevant information must be fed back to the correct points in the process as a basis for making closed-loop control decisions.

Accordingly the requirements are to identify factors, to make available all relevant production data, to analyse this data, to extract information from this data and finally to generate knowledge concerning the current cause and effect relationships from the information.

6.2.1.3 Results

With the aid of self-optimising mechanisms a production system may be optimised by considering the whole process chain. Self-optimisation offers the possibility of implementing improvements autonomously.

To determine production parameters in process chains, tools that can simulate technically the applicable procedure which mirror the human information process are suitable. To employ these tools to the greatest effect their specific strengths must be combined in operation with various types of tasks.

The tools can be combined by means of modular designed software architecture. The actual software implementation enables closed-loop control of production in production environments that are changing dynamically and controls dependencies between product and production parameters which cannot be described analytically.

It calculates solutions by means of efficient learning mechanisms and enables the closed-loop control of changing targets and disturbances to be adjusted rapidly in the cycle time. It implements the flexibility necessary for an individualised production scenario and at the same time helps to resolve the polylemma. Methodology and software are elucidated by means of application examples.

The transmission of torque by a shaft-hub connection is employed as a first application example, which permits a simple comprehensive explanation of the procedure. The second application example is the production of components for a car rear axle drive.

The focus here is on the acoustic emission of this drive during operation—a quality feature relevant to the customer that depends upon many components, component properties and in particular their tolerances. When the acoustic emission increases customers sense the generation of noise in the vehicle as annoying. If it lies below a vehicle-specific limiting value, then the drive noise will rarely be noticed.

The aim of optimisation is to meet the quality demands of the customer without additional cost and increased production costs. Increasing production flexibility to afford faster and efficient response to deviations will ensure the product quality and a relaxation of unnecessarily tight component tolerances favouring a homogeneous functional performance can reduce both fabrication costs and time.

This secures the competitiveness of the production. To achieve this, the interdependences between tolerances and product function fulfilment (here: low-noise power train) are analysed and fed back into the appropriate processes.

The application of cognitive methods in acquiring and evaluating production data and also generating decisions to optimise the production scenario is paramount in the application examples. Production data is evaluated with the aid of graphical data processing methods, modelling of the production processes is carried out by Artificial Neural Networks and the determination of process parameters is implemented by the cognitive architecture Soar as the core of optimisation software that can take decisions autonomously on the basis of control systems knowledge and practical knowledge.

In order to attain maximum production flexibility the capability of learning rapidly from available data is implemented in addition. Existing knowledge is updated via learning algorithms and the learning process when creating models is systematically supported and accelerated with quality management methods.

First the tasks and challenges of the developed cognitive software are derived. In Sect. 6.2.2 the motivation for the research projects and also the specific research question are derived from these, via which a contribution towards resolving the polylemma is evolved. Section 6.2.3 includes state-of-the-art techniques for self-optimising approaches to closed-loop control of process chains against the background of the polylemma to be resolved. The detailed work in the research projects is introduced in Sect. 6.2.4. This is divided according to the steps in human cognitive information processing. These comprise the cognitive extraction of information from production data, the analysis and interpretation of influences of specific feature characteristics upon the production results along with generating decisions to determine production parameters on the basis of knowledge and information, and also processes in order to learn from the combination of selected parameters and targeted results for future situations. Data mining processes are employed for analysing the data. The cognitive architecture by Soar is employed for decision making. Artificial Neural Networks are employed for illustrating cause-effect relationships and reinforcement learning is evaluated as learning process. In conclusion the industrial relevance of the solution concepts is evaluated in Sect. 6.2.5, before providing an overview upon the future research topics in Sect. 6.2.6.

6.2.2 Motivation and Research Question

The entrepreneurial task in design and production consists in optimising the liquidity ratio between the market demands and the properties of the products taking into account economic aspects. This means that a highest possible liquidity ratio is not to be striven for in principle, but that the constraints of the optimisation task, such as product pricing, throughput time in the process or the flexibility of a complete production on a company-specific basis are to be taken into account in all areas universally (Schmitt et al. 2007).

In the result a best possible quality of a product at best possible cost is to be delivered. Quality can be accessed from different points of view. Garvin differentiated the transcendental (subjective), the product-related, the customer-related and the value-based understanding of quality (Garvin 1984). The different views lead to various assessment results, since different criteria interact with each other, such as e.g. costs, functionality and service life. Objectives result from the various criteria for consideration and constraints which in combination form a multi-dimensional optimisation problem.

6.2.2.1 Tolerance Allocation in the Actual Process Chain

Tolerances in particular have a significant influence upon the perception of product quality, and upon the requirements and costs of production stemming from this. Quality demands determine and specify the subjective and objective functionalities of an overall product (Schmitt 2010).

Functionalities of a product arise due to the combination and/or interaction of its components, which in turn may be specified by their properties. A typical component property is its geometry that is defined by geometric dimensions and tolerances. Figure 6.4 shows the essential aspects of tolerances in the product life cycle: accordingly the product functions to be implemented are translated into geometries, nominal dimensions and tolerances in the design and construction phases of a product, which are the significant parameters in the assessment of the product components throughout the product development. These transformations are necessary since the fabrication according to the current state of art proceeds geometrically-based. The design of products or rather components is continued via the assembly of parts or components to form a sub-assembly through to the overall product (Jorden 2007). However the geometry is rarely the entire quality feature in the eyes of the customer, but the function that results from this. Geometries are however always prone to deviations in manufacture, that confronts designers with the task of specifying suitable dimensions and tolerances for components, so that the function of the sub-assembly or the overall product can be ensured and all customer requirements are at the same time taken into account optimally.

The area of tension in which the designer operates is characterised by two essential poles: firstly the functional capability must be ensured. Secondly the production costs rise with the requirements upon individual product properties (e.g. Accuracy of physical dimensions (Merget 2004; Jorden 2007)). A conflict of targets between production costs and a risk of dimensional deviations arises. Depending upon the

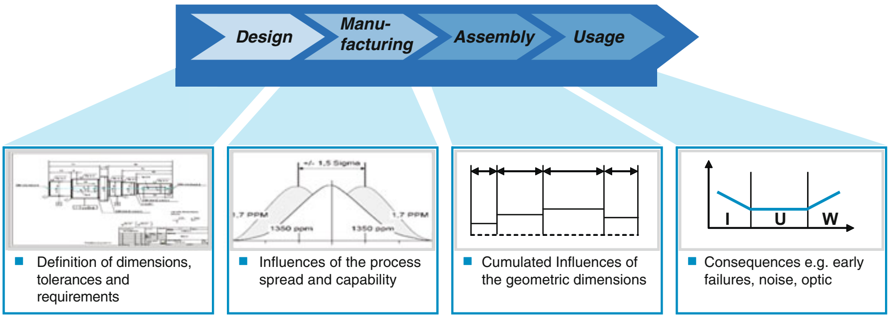


Fig. 6.4 Tolerances throughout the product life cycle

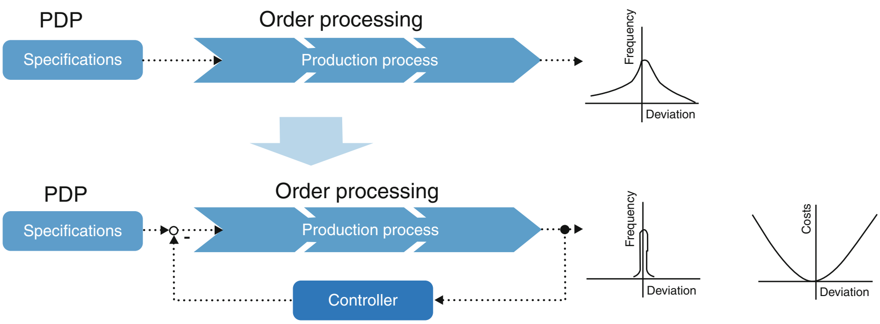


Fig. 6.5 Production control of a component

outcome of this conflict the compromise reached will lead either to quality problems on the product side or to cost problems on the manufacturing side. Geometrical product tolerances of product components are therefore one of the most frequent causes of quality problems in industrial manufacture (Scherms 1998).

Tolerance deviations occur with a distribution depending on the quality of the production process. Figure 6.5 shows the distribution of the tolerance deviations as an example of a production process for an individual component. With the aid of an effective production control (closed-loop) system the process quality was able to be improved and the probability of a tolerance deviation reduced. The precondition for this is the knowledge concerning the causes of deviations. The influence of tolerance deviations upon a product may be described by a loss function (Taguchi 1990). The loss function specifies the costs arising from a deviation from the optimum dimension. The loss function may be employed to decide for instance, at what dimensional deviation is re-work economic or the component must be treated as a reject. In most cases the production processes for individual parts of a product are controlled and optimised to the closest possible dimensional accuracy independently from one another.

A loss function example is shown in Fig. 6.5. The costs rise disproportionately to the deviation magnitude.

If the product or sub-assembly comprises several components, it is often possible for the deviations in down-stream process stages to cancel out, for example during assembly (Fig. 6.6).

For this reason, components typically from well controlled processes having a very small deviation from the target dimension are employed. Due to assembling these components based upon dimensions, dimensional deviations can be cancelled out and therefore rework prevented.

Due to the possibility of cancelling out deviations in individual components, the potential for optimisation arises. In order to exploit this however the loss functions can no longer be considered individually, but must be combined as multi-dimensional functions. This is for example practised in the assembly of roller bearings.

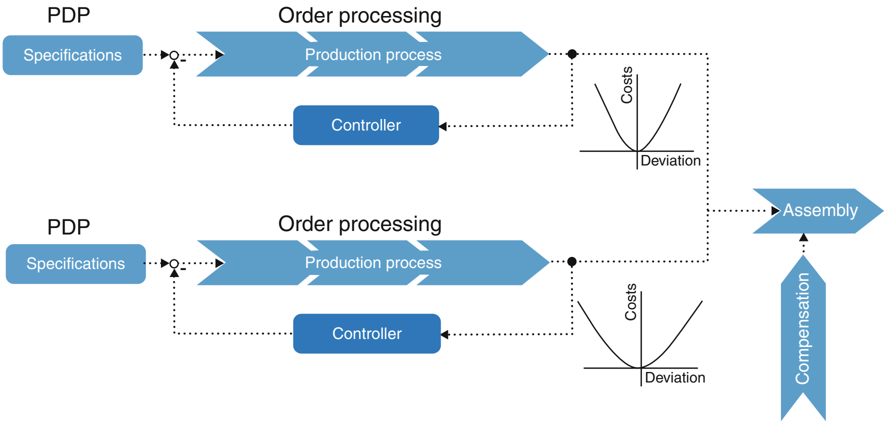


Fig. 6.6 Production control of several components

To achieve an optimum accuracy of fit between inner ring, rolling elements and outer ring a pairing by means of the dimensions achieved in the fabrication is carried out. Pairing is possible, since to fulfil the roller bearing function optimally the tolerancing of the individual components is not important in absolute terms, but is important relative to one another.

Figure 6.7 shows an example of the pairing between outer ring, rolling elements and inner ring carried out when manufacturing a motor car clutch release stop. The rolling elements must be manufactured with extreme dimensional accuracy and must enable the deviations arising from the pairing of outer ring and inner ring to be cancelled.

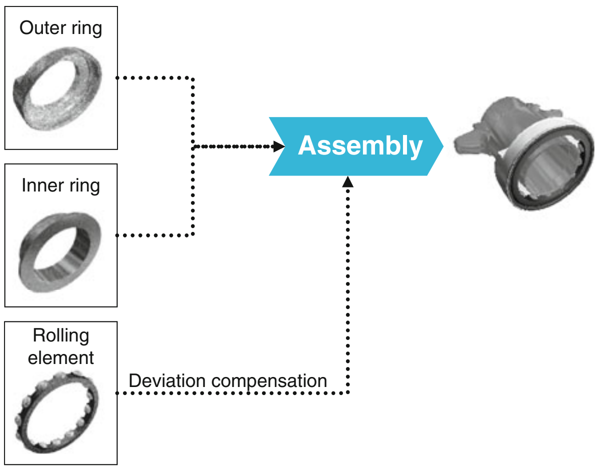


Fig. 6.7 Assembly of a roller bearing by pair-selection by means of rolling elements (e.g.: Clutch release stop). (Schmitt 2000)

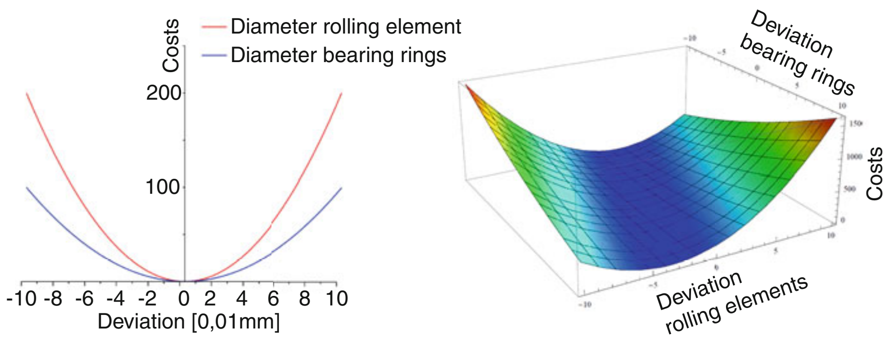


Fig. 6.8 Cost functions of bearing rings and rolling elements

This dimensional pairing of the individual parts permits cost-effective fabrication by simple component pairing, but nevertheless enables high precision bearings to be manufactured (Niemann et al. 2005).

Figure 6.8 shows the individual cost functions for the dimensional deviations in rolling elements and bearing rings on the left. The three-dimensional plane function shows the interactions of both cost functions. It clearly points out that there is a broad area for minimising the cost function, which may be attained by pairing individual components.

To establish the cost function it is necessary to be able to specify the relationships between individual deviations and the costs arising mathematically. Many production processes for demanding products however suffer from the lack of adequate knowledge of the relationships between production parameters and the functional performance of the component. Multi-faceted interdependences and a linking of several parameters are often present. An example of this is gearbox production, wherein the functional performance is characterised by dimensional stability, efficiency and acoustics. These dimensions may be influenced by numerous parameters in material choice, fabrication and assembly. In addition we are confronted by the challenge, that the functional parameters cannot be controlled individually, but in most cases stability, efficiency and acoustics are influenced interactively (Brecher et al. 2008d). It is not possible to establish an analytically minimising cost function.

Complex processes such as the gearbox production selected here as an application example consequently require methods that extend beyond simple pairing. A closed-loop control system is needed that can coordinate the various individual processes in parallel. Such a closed-loop controller is illustrated in Fig. 6.9 as an example for a bevelled gear wheel and ring gear pair. When the cause-effect relationships are unknown the closed-loop control system cannot however be reproduced analytically. In this case it is necessary to be able to simulate the cause-effect relationships in the production process via a model. A classical closed-loop controller is not possible. In this case it is necessary to be able to simulate the cause-effect relationships in the

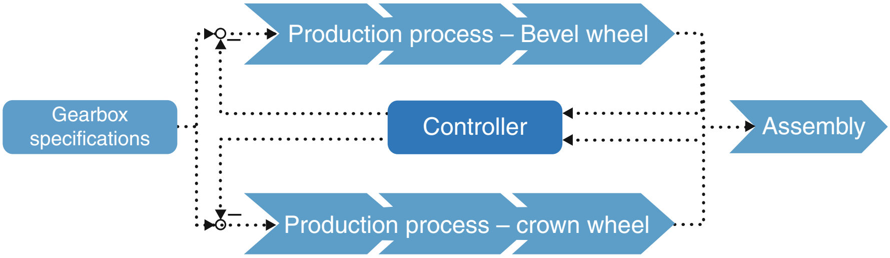


Fig. 6.9 Multiple target closed loop control in gearbox production

production process via a model. If an appropriate modelling mechanism is possible, an optimisation strategy is needed, which with the aid of the process model creates parameter suggestions for the production, with which the various defined goals can be fulfilled. As a result the targets that specify the product functional performance are of central significance. Subordinate aims such as for example tolerances of individual components, may be exceeded to a limited extent, if this can correct the result of another production stage in relation to the functional performance of the finished product.

Focusing upon the functional performance of the product beyond the features of individual components may be achieved by means of a self-optimising closed-loop controller. This is able to call into question also higher-level structures and interdependences. Classical approaches focus solely upon the optimisation of individual parts of the process (Pfeifer et al. 2003).

As part of this project a self-optimising system will be developed that enables processes to be controlled, whose analytical transfer functions are not known. The system is able to adapt itself flexibly to the process to be controlled and to adapt also to control targets, when this results in a higher level functional performance.

To realise this self-optimising system it is necessary to model the production system to be controlled and to take decisions for a closed-loop controller based upon the data generated by the model. Cognitive technologies are employed for modelling and decision-making. In particular this enables complex processes and ever-changing basic conditions to be optimised flexibly due to its inherent learning ability. The *research question* that arises from the starting position and the specified goal is formulated as follows:

> Can cognitive technologies control production processes having a number of inherent interdependences more flexibly, faster and using resources more efficiently, than planned production systems?

The work carried out is derived from this research question. This discusses, how and by which methods and techniques the possibilities of cognitive processes may be transferred to the challenges of optimising a complex production system.

6.2.3 *State of the Art*

Self-optimising control circuits can execute multi-target optimisation autonomously by open-loop and closed-loop applications that deploy cognitive technologies. Therefore an overview of existing approaches to self-optimisation in research is given below.

Existing technologies having cognitive character may be applied in data-recording and interpreting and also decision-making, while data-mining methods are being considered for recognition of cause-effect relationships. Artificial Neural Networks act as tools for modelling processes to evaluate Soar-generated decisions, prior to being utilised in the production system. The cognitive architecture by Soar is introduced at the application level for the optimisation. This is already employed in various application areas for generating decisions and implementing learning processes (Sutton and Barto 1999; Lehman et al. 2006).

6.2.3.1 Closed-Loop Control Systems in Production and Self-optimising Approaches

The application of closed-loop control systems is one of the value-based approaches and is utilised widely in industry at all levels of a company. Closed-loop control systems secure a corrective response when deviations from the target value occur in products and/or processes. Closed control-loops within the operative level aim at e.g. the condition of product features, by monitoring the maintenance of tolerances as closed-loop control systems near or inside machines (Conventional process control) The process concept of closed-loop control however must be broadened in the case of its application in quality management in contrast to the process concept of classical automatic control technology. All activities are described here by processes that influence the product development. Consequently a component of the control loop can also be people, machines and engineering processes, a method, products or materials and external influences in the work environment, such as the temperature or the humidity (Takeda 2009). The range and degree of difficulty of applications is increasing with an increasing complexity of modern products and heightened customer demand for a multiplicity of diverse variants. Companies are confronted with the challenge of reconfiguring their production systems into highly flexible, adaptable units, with which extremely individual products and the process variance resulting there from can be controlled.

Autonomous production systems pursue the goal of being able to adapt automatically to changing input conditions with the greatest possible precision (Among others: Pfeifer and Schmitt 2006; Chryssolouris and Mourtzis 2004). The monitoring scope of the closed-loop control circuit must be widened correspondingly in order to be able also to synchronise individual targets autonomously with one another across all processes. A ***Self-optimising production system*** pursues the goal of implementing the desired autonomy of the production systems, by carrying out the following steps in a recurring cycle, according to the definition by Gausemeier et al. (Gausemeier et al. 2009):

- Analyse the current situation
- Determine the (new) system targets
- Adjust the system behaviour

The Collaborative Research Centre 'Self-optimising mechanical engineering systems' (Adelt et al. 2009) as originator of this definition is admittedly concentrating upon mechatronic systems, thus primarily the product side, but offer among other things procedural models and tools along with practices, the use of which is also conceivable in production. In this way various methods of data analysis, a cognitive architecture having a 'Predictive algorithm', which is similar to an integrated model of the system to be controlled, possible future conditions are taken into account and various approaches for modelling and decision-making are employed. It is broken down into 'numeric processes based on a priori defined models of the physical system behaviour' as well as 'Learning methods and planning procedures based upon a black box specification of the system behaviour, which is arranged in part with the aid of accumulated experience'.

The Collaborative Research Centre 'Autonomous production cells' (Pfeifer and Schmitt 2006) along with further approaches by Mitsuishi et al. (2004) utilised an integration of physical process models in the machine control and installed model-based machine operation and monitoring as a constructional addition to this. Complex machining processes should be able to perform using optimal resources incurring minimum faults over longer time periods. Autonomous production cells do not have the possibilities of self-optimising production, since they do not have autonomous target adaption or decision-making.

The consideration of comprehensive influences during evaluation and decision-making requires a new approach, taking due account of this with the concept of a 'cognitive automation' (Onken and Schulte 2010). This concept is applied for example in the guidance of unmanned vehicles and also the control of virtual aircraft in flight simulation (Jones et al. 1999), which admittedly is comparable with an automated production in some aspects, but is only conditionally transferable.

With the so-called 'Cognitive factory' (Zaeh et al. 2009) there are approaches in research, to utilise cognitive mechanisms successfully in fabrication systems. All stages of the manufacturing processes are combined in the cognitive factory by means of EDP (e.g. Via Auto-ID-Systems such as RFID (Radio Frequency Identification)) and the information flow is implemented continuously from the customer requirements through to the delivered product, so that communication from component to production machine can be effected. Key areas of activity in the cognitive factory are:

Intelligent design: The CAD system provides knowledge regarding the capabilities of the production system employed. This avoids the need for long redesign-loops and for machines that are only used seldomly.

Artificial cognition in assembly scheduling: Production systems are equipped with their own planning capability. Thus new work sequences may also be generated and evaluated.
Cognitive assistance systems in assembly: Cognitive assistance systems permit the assembly robots to be adapted flexibly to the operators. By dint of this the operator can integrate flexibly with the robot. The robot detects the operator status and adapts its process steps correspondingly.

The individual modules in a cognitive factory are equipped with engineered systems. These comprise sensors and actuators for interacting with the environment, along with techniques for implementing perception, interpretation, learning and planning. The aim is to design engineered systems so that "they know what they are doing".

Technion (Israel Institute of Technology) is also grappling with the challenge of automated planning of assembly scheduling with the aid of cognitive technologies. Decisions for assembly scheduling are based upon a holistic model of the assembly process that takes into account both technological and economic aspects (Denkena and Shpitalni 2007).

The Fraunhofer-Institute for Optronics, Systems Technology and Image Assessment IOSB has developed a data mining tool for decision support in the production sector with ProDaMi (Assumed a company specialising in Data Mining). Data mining algorithms capable of learning enable intelligent evaluation of production and systems data. The goal is to generate knowledge capable of making decisions. The project makes it possible to recognise complex interdependences in production processes, but does not enable systems to be modelled and is unable to make independent decisions.

To implement a self-optimising system for closed-loop control of complex production processes technologies for the following functional areas must be employed:

- Data analysis
- Modelling
- Optimisation

The state-of-art of the technologies applicable here are as follows.

6.2.3.2 Data Mining for Cause-Effect Relationships Analysis

Prior to the actual closed-loop controller the data to be processed is analysed. The aim of the analysis is to recognise interdependences within the data and also to simplify the issues to be addressed by reduction of data to information sources relevant to the issues. Methods borrowed from the data mining area may be employed for this analysis.

Data mining is known as the recognition and extraction of non-trivial and previously unknown patterns from huge amounts of data. In addition various statistical and mathematical methods are utilised. Data mining itself is divided into five sub-categories (Hand et al. 2001):

1. Explorative Data analysis
2. Pattern recognition
3. Content recognition
4. Descriptive modelling
5. Predictive modelling

In the case of explorative data analysis large amounts of data are browsed, whereby a priori no knowledge exists upon the data searched. In the context of pattern recognition the challenge is to recognise typical or atypical patterns in the data investigated. With content recognition structures within the data similar to patterns which are known and already identified, are sought. Descriptive modelling aims to find descriptions for the existing data sets, with predictive modelling additional attributes to the value are to be deduced with the aid of the existing attributes. Various methods such as classification or regression are employed for predictive modelling (Hand et al. 2001). One of these classification methods involves decision trees, which are seen to be especially attractive to categorise within the data mining for the following considerations:

1. Decision trees can be generated without additional inputs from the user
2. Compared with alternative methods decision trees can be generated in a very short time (Shafer et al. 2005; Gehrke et al. 2000)
3. The model quality of decision trees is very good (Michie et al. 1994; Murthy 1995; Lim et al. 1997)
4. Decision trees offer a very intuitive presentation format (Larose 2005)

6.2.3.3 Artificial Neural Networks as a Modelling Tool

Modelling of (production) systems may be carried out by various modelling types and tools according to the intended application. The requirements upon a model in the environment under consideration however include in particular the demand for calculability executed by software. This is necessary in order to be able to evaluate computer-based optimisation proposals, prior to being employed in the actual production. This renders the application of a mathematical model desirable, which consists of the equations for specifying the system behaviour (Unbehauen 2008).

This is not possible in the case of severely non-linear production processes, which are characterised by complex interdependences of many parameters and in addition must take account of various stochastical influences. The creation of mathematical models having explicit information of the system-specifying equations is not possible in this case or not possible with justifiable computing time. They are therefore unable to be employed for production control in real time. This also applies to the application example of the manufacture of rear axle drives which is considered here. Blackbox models such as Artificial Neural Networks offer an alternative: although their behaviour follows mathematical laws, they do not need explicit information of the behaviour-specifying equations, since they can learn their behaviour numerically from sample data (Weinmann 1995). Artificial Neural Networks are therefore ideally qualified to be deployed for modelling purposes in the current application example.

Artificial Neural Networks are based upon the functionality of biological networks as information-processing systems. In 1943 McCulloch and Pitts described a neurological network for calculating logic and arithmetical functions that may be deployed also for pattern recognition. In 1949 Hebb created the Hebbsche learning rule for representing neural learning procedures. In 1960 a network was introduced with ADALINE (Adaptive linear Neuron) that uses delta rule learning for echo filtering in analogue telephones (Levine 2000; Golden 1996).

Following the introduction of Linear Associators by Kohonen in 1972 a non-linear neuron model was used in 1973 and Werbos defined the Back-Propagation method in 1974 for implementing a learning procedure. Mathematically based models of neural networks were developed by Grossenberg in 1976, prior to his developing an architecture for Artificial Neural Networks with Carpenter (Gurney 1997; Dören 2007; Müller and Reinhardt 1995; Zakharian et al. 1998)

The structure of an Artificial Neural Network in general comprises neurons arranged in layers, which exhibit links between layers. Artificial Neural Networks are characterised by these neurons, the architecture (structure of the layers, their nodes and the relevant links), the activation functions of the neurons and also the learning algorithms used.

Artificial Neural Networks have undergone specialised further development depending upon the challenges confronting various applications: (Dören 2007)

- Approximation of functions, e.g. Perceptron (Rosenblatt 1958)
- Image and speech recognition, e.g. The Kohonen Network (Kohonen 1984)
- Classification of patterns, e.g. Cooper's RCE (Zakharian et al. 1998)
- Prognosis, e.g. Perceptron (Rosenblatt 1958)
- Compressed storage, e.g. RAM
- Noise suppression, e.g. Adaline87 (Zakharian et al. 1998)

The Multi-Layer-Perceptron (MLP) was developed for approximating functions and for prognosis and is applied also in the software architecture introduced here. The sigmoid function is proven and widely used for activation (Kinnebrock 1994).

The network training function enables it firstly to anchor the interdependencies between input and output data to be mapped into the network. The Back propagation method (Error feedback method) is employed with MLP networks for this purpose. The Back propagation process is therefore defined as follows: "The term Back Propagation denotes the backwards propagation of an error signal through the network" (Nauck et al. 2002). The method attempts to reach a minimum deviation between calculated and actual trained results. Additionally the output pattern arising in the Artificial Neural Network is compared with a target and an error calculated from this. This error is then fed back into the network so that each internal unit is able to calculate its own error and thus to modify the individual weightings, so that the error is minimised (Nauck et al. 2002).

Neural networks can be found in many areas of production technology, in particular in process monitoring (Eversheim et al. 2006), such as e.g. in process monitoring and control of thermal spraying (Dören 2007) as well as monitoring tool wear (Fries 1999; Rehse 1999) and for decision support in grinding processes (Liao and Chen 1994).

6.2.3.4 Optimisation by Cognitive Technologies

Various algorithms are suitable for creating decision trees for deployment in production technology. An approach is chosen as optimisation technology in light of the present problem, which is designated as simulation-supported optimisation. In this case an optimiser and a process model are necessary for implementing the actual self-optimising closed-loop control. The optimiser generates parameter proposals, which are evaluated subsequently by the process model (among others in Fu 2002). In simulation supported optimisation simulation programs are employed for determining optimisation results (Sauer 2003). Here optimisation means seeking to achieve a desired process result by varying input parameters or variables. In contrast a conventional optimisation cycle achieves this by manual analysis of various process variants. For economic reasons however the process is not carried out physically at each optimisation stage, instead a numeric process analysis is brought into play, or the process is executed with the aid of simulation technology, or Artificial Neural Network, modelled in the computer. A number of strategies and algorithms are available for varying the parameters. With simple challenges a sensitivity analysis can be invoked (Assem 1986), and by these means the qualitative relationships between process parameter and result can be determined. Because of severe non-linearity and oscillations the target function is not practicable in the case of actual, complex production processes. Various investigations have revealed that it is sensible to separate optimising algorithm and process model completely. This separation also enables combinatorial optimisation procedures to be employed in which the relationships between process parameters and process result may be completely unknown (Becker 1991). Many modern optimisation systems are founded upon this concept, such as for example the Optimisation Shell by Grešovnik and Rodič. When aided by various combinatorial methods it offers the possibility of optimising diverse metal forming engineering problems automatically (Grešovnik and Rodič 1999; Posieek 2005).

Soar's cognitive architecture is suitable as an optimiser in the application examples described here. Soar is based on the KI systems GPS and OPS5, in which intelligence according to the principle of rationality is understood as the optimum achievement of targets. In Soar target-oriented problem-solving takes place as heuristic search in problem spaces. The search is affected through successive applications of operators until the target state is reached. In expanding classical planning systems the problem space search is built into a complex decision cycle. Soar is a condition-oriented program language, whose programs do not have a fixed sequential execution, but comprises rules that specify the space of possible conditions and the possible actions therein (Lehman et al. 2006; Laird and Bates Congdon 2008). The Soar program operates within the possibilities stipulated by the rules to reach a specified goal automatically. Initially the system acts like a random-based Monte-Carlo optimiser, but with advanced run-time the program learns from successes and failures and thus shortens the time required to solve the problem (Nason and Laird 2005).

6.2.4 Results

The challenge of solving the research question includes firstly the synchronisation of the technical dimensions in a company-wide closed-loop control system and secondly, an organisational dimension for coordinating all participants. This necessitates especially robust processes that are synchronised with one another. In particular the synchronising of processes with one another requires powerful approaches within a highly dynamic company environment. The implementation of closed-loop process control is effected by structured design of quality control systems in the company—and beyond the company's boundaries—by adopting analogies from closed-loop control techniques. Process-oriented closed-loop quality control supports the preparation and implementation of permanent competitive process landscapes, which adapt to the diverse and dynamic situations invoked by information and material flows (Schmitt et al. 2009a). This adaptation is effected by auxiliary tools from the cognitive sciences, which can recognise the information necessary for adaptation within all prevailing production data. As a result cognition provides the basis for the applicability of a closed-loop control system that will act upon this information. However while a classical closed-loop control system addresses the transfer function of the control loop, the closed-loop control process in the approach adopted here is carried out by a decision-making function by utilising self-optimising mechanisms, which constitute the various elements within the closed-loop quality control system.

6.2.4.1 Cascaded Closed-Loop Quality Control Systems

By designing closed-loop quality control systems to be deployed universally throughout the company relevant information can be distributed and used at the right time at the right place according to requirements. By intermeshing and linking closed-loop quality control systems at various levels in company unambiguous rules for decisions and escalation routines at engineering and organisational levels emerge. The continuous alignment between actual and target state enables continuous improvement to be institutionalised in the company.

By considering the processes at the various levels of the people and departments involved by means of cascaded closed-loop control systems, not only the improvement of individual partial aspects is ensured. On the contrary all the targets based upon the market demands from the company processes out to the finished products at the customer along with the extraction and recycling of user data will be achieved. Closed-loop quality control systems institutionalise this knowledge (Fig. 6.10).

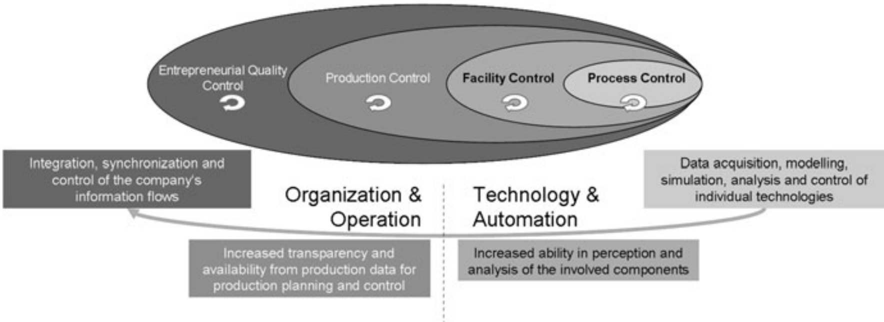


Fig. 6.10 Cascaded closed-loop quality control systems

The approach deployed makes use of the economic specification of tolerances against the background of the challenges discussed. Information from individual processes should be fed back to the correct points in the process chain with methodical support. Conclusions upon correlations in the manufacturing process are compiled; process and product parameters can then be adapted across all individual levels according to the situation. This adaptation allows individually specified parameters to be controlled and in this way the idea of self-optimisation with continuous alignment of target and actual states and the derivation of actions there from is realised.

6.2.4.2 Cognitive Information Processing

To execute closed-loop control across all levels, which adapts a target system automatically to the requirements of the various company levels, it needs intelligent coordination of the data flows, procedures for extracting exploitable information from the available data (knowledge extraction), and the preparation, making and validation of decisions. Conventional closed-loop control technology reaches its limits especially when making decisions autonomously. Making decisions to implement self-optimising production is comparable to human decision-making in its requirements. The engineering implementation of a self-optimising closed-loop production control in the style of human decision-making is the goal of the research. The engineering implementation of human decision-making is the cognitive science sphere of work and is explained in more detail as follows in the context of having transferred the procedures to the area of activity described.

The scientific definition denotes “Cognition” as a generic term for all processes and structures, which carry out the perception, classification and evaluation of facts and also decision-making (Strohner 1995). An example of this is human information processing: all actions that the human being executes in his environment are characterised by first perceiving the stimuli, then processing and subsequently acting upon these. This scheme is adapted and deployed here (Fig. 6.11).

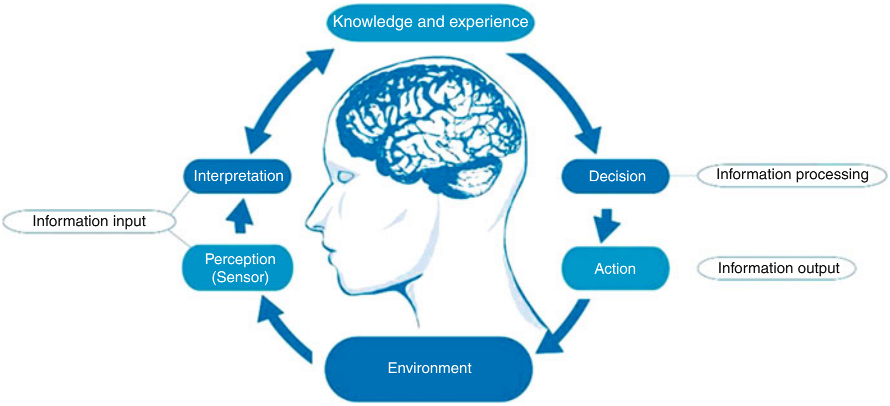


Fig. 6.11 Modelling cognitive processes

Immediately after acquiring the information, the perception, the information is processed. At this point the information acquired is first interpreted on the basis of knowledge and experience and a decision is then derived. Finally the information is output, i.e. decisions upon actions are implemented. The action executed represents a direct influence of the perceived environment, so that the cognitive closed circuit recycles.

Cognitive science pursues the aim of simulating this human information processing with the greatest possible precision by means of engineered systems. These engineered systems are called cognitive technologies. As part of the research work an architecture is being developed in which various cognitive technologies are employed. When developing the architecture the goal is to combine and shape the cognitive technologies, so enabling self-optimising closed-loop production control.

6.2.4.3 The Cognitive Tolerance Matching Architecture (CTM)

An essential requirement upon cognitive architecture for implementing a self-optimising production system is modelling the production processes to be optimised. The architecture must be able to generate production parameters, which can be evaluated by means of the process model. Also in order to adapt flexibly to varying basic conditions, the architecture must be equipped with an inherent learning capability.

The architecture developed enables self-optimised control of production processes by cognitive technologies. In addition it is being modularised according to the model of the processes and necessary functions of the human problem analysis and decision-making. It is composed of various problem-based modules, each representing parts of the overall process. The technologies of the individual modules may be selected individually with regard to the issues to be solved.

The architecture is designed as illustrated in Fig. 6.12. The perception and operation layer serve to exchange data and commands directly with the production

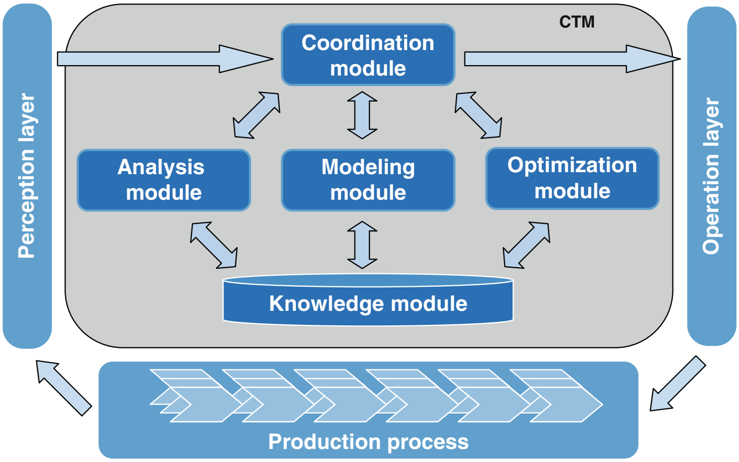


Fig. 6.12 The architecture of cognitive control software

system, while the central core of the architecture contains the elements necessary for the optimization. These provide the problem analysis and modelling, the decision-making (hence the actual optimization) and the mapping of knowledge with the aim of permanent storage and coordination of the individual elements. These tasks are represented respectively by their own modules or levels.

The *Perception layer* serves to acquire and forward sensor data and is employed as interface for the production system. Product and production data is acquired at this point. Implementation in engineering terms corresponds to a connection to the machines and measuring instruments. The perception layer makes the data available to the other modules with which these act, as a bundle and according to demand.

Coordination module communication between the various levels and modules. It provides the data employed and initiates the processes that are carried out by the other modules. In addition it forms the interface for direct interaction with the user.

The *Analysis module* provides the preliminary analysis of the data considered. It carries out e.g. correlation analyses and data reduction, in order to identify the essential data and keep the model complexity as low as possible, but nevertheless sufficiently accurate. The rules utilised for the optimisation are extracted on this basis.

The *Modelling module* is employed for mapping cause-effect relationships in production systems. It serves to predict results of individual production stages, which must be evaluated as a part of optimisation. In the modelling module both analytical and black-box models may be employed. Artificial Neural Networks are employed in the context of the previous implementations in light of a problem unable to be mapped analytically.

The *Optimisation module* creates the optimisation strategy autonomously and implements the flexible response to external influences. It deploys software by Soar in essence, to make optimisation decisions on the basis of adaptively expandable closed-loop control knowledge, and to adapt the production system behaviour autonomously by learning processes.

The *Knowledge module* is responsible for saving and making available the relevant data, which maps the knowledge necessary for the optimisation. This comprises both the rules that are applied in Soar and also the data (input and results), which has already been deployed in the actual system and thus may be considered as substantiated data points.

The Operation layer serves to implement the optimisation decisions made and acts as interface for the production system. Parameters are transferred to the production system via this interface, so that the resources concerned are regulated to the target communicated to the closed-loop control system. The aim in a production (system) being deployed is to be able to implement the control commands via automated interfaces directly by means of the resources.

The interaction between the individual modules begins with the data acquisition in the perception layer. In this case it concerns both the process data from production machines and the measurement data from various sections of the production process. This data is distributed by the coordination module to the individual modules in the architecture. In the next stage the optimisation module generates a parameter proposal for a subsequent stage based upon the current data. This proposal is transferred together with the current process data to the modelling module by the coordination module. The modelling module then generates a prediction for the process result targeted in this case. The coordination module sends this prediction result back to the optimisation module, where it is evaluated in relation to target achievement. Further parameter proposals are then generated, until a result deemed satisfactory in accordance with the requirements for the optimisation module is reached. The parameters thus found are then output to the operation layer by the coordination module and from there accordingly deployed in the production system.

Production Mapping in the Modelling Module

The software generated here enables various tools to be integrated for process modelling by dint of its modular design. In light of the characteristics of the problems specified Artificial Neural Networks were chosen to evaluate the optimisation proposals, since although their behaviour follows mathematical laws, they do not require explicit information for the equations specifying behaviour, since they are able to learn their behaviour numerically from sample data (Weinmann 1995). Artificial Neural Networks therefore offer ideal conditions for deployment for modelling purposes in the present application example.

Artificial Neural Networks

Artificial Neural Networks have been developed to simulate the behaviour of biological nerve cells by engineering means. They are able to learn a behavioural pattern by established training data. As a result they are able to map technical facts such as for example the behaviour of a production process. They have been developed for various applications customised depending upon the problem (See Sect. 6.2.3), wherein the multi-layer perceptron (MLP) was developed for approximating functions and prognosis and also applied in the software architecture application presented. The multi-layer perceptron comprises an input and output layer with respectively n and m neurons. In addition it possesses intermediate layers. Each neuron is connected with all neurons in the surrounding layers. The sigmoid function is of particular significance as the activation function, since it is continuous and enables non-linear functions to be mapped in Artificial Neural Network applications (Dören 2007). Artificial Neural Networks are characterised by these neurons, the architecture (Composition of the layers by their nodes and associated links), and the activation functions of the neurons along with the learning procedures deployed.

The back propagation method (see e.g. Gurney 1997) is deployed for training the MLP network in the present context, which is ranked among supervised learning procedures. By feeding back errors of the training phase, it moves each inner entity to the layer to calculate its own error and to modify its own weighting so that the error is minimised (Nauck et al. 2002). This capability is used here to train networks. Training Artificial Neural Networks with the back propagation method in the application functions in such a way that they optimise their internal transfer functions against the target by adjusting weightings and activation values, based on data from input parameters and results there from provided for learning purposes, to produce output data having the least possible differences in relation to the given input values of the data. Consequently the prediction quality of the network is improved with the number of available substantiated data points, therefore operating points in the present case. As a result Artificial Neural Networks are also able to model non-linear interdependences.

Process Modelling with Artificial Neural Networks

Artificial Neural Networks are currently deployed successfully in many application areas, which may be both discrete event-based and continuous in nature. In the present application they are deployed to optimise process and product parameters in multi-stage process chains. This concerns producing components for rear axle drives, for which industry experts from the individual parts of the processes employed are working on a comprehensive deterministic model as the basis for the optimisation. The number of parameters considered and also the number of unknown cause-effect relationships is cause for restricting the optimising processes to individual aspects of production. Artificial Neural Networks have been identified as an approach, which dispenses with the expensive creation of process models and nevertheless is suitable for assessing results of production under known conditions, provided the possibility is provided to train the network by means of adequate data (Dittmar and Pfeiffer 2004). Comparison with regression analytical calculations also reveals the potential of Artificial Neural Networks to be able to map non-linear interdependences very sensitively also. An additional advantage of Artificial Neural Networks is the possibility of being able to map current changes by regular training.

Architectural Aspects When Modelling

Artificial Neural Networks can be deployed as black-box models, in which the interdependences and mapping rules do not have to exist provided in explicit mathematical terms, but are learned numerically. This reduces the model creation to taking into account the essential characteristics of the present problem when creating basic structures of the neural networks. A mathematical specification of interdependences of the variables considered is not necessary, since the cause-effect relationships between the variables considered are learned automatically and deployed by the neural network. A "Supervised learning" procedure is deployed with back propagation in this case, with which an Artificial Neural Network can adapt its behaviour concerning the adjustment of the transfer functions. Consequently only its composition is relevant to the capabilities of Artificial Neural Networks, comprising the architecture of neurons and layers along with the links, the activation functions and the learning methods deployed. If multi-layer perceptron networks with sigmoid function and also back propagation as learning function are employed, only the neurons of the individual layers along with the number of layers remain to be determined Therefore the input layer consists of n neurons, whereby n is the number of input parameters, and the output layer of m neurons, where m represents the number of output parameters. Depending upon the problem complexity the formation of the middle layer(s) (called hidden layers) is still to be determined. This is affected in particular cases specific to the problem and is substantiated by appropriate studies, in which learning rates and deviations in results are compared (Pfeifer and Schmitt 2006).

Accelerating the Learning Process of Artificial Neural Networks by Means of Statistical Methods

To deploy Artificial Neural Networks as auxiliary resources when optimising production systems in real time, the speed of adapting to the given conditions is an essential aspect to consider. In addition to the required quality of results the demands upon deployment include the processing speed (Response time) as well as the time required to train a network with the actual conditions of the production to be optimised or to encounter deviations between calculated and actual result, as the internally mapped interactions of the model are updated.

In this connection the critical factor is "time", but not only dependent upon the processing speed of the network, but also of the production system to be mapped, the real time data for operating points to be learned must be made available. If this data must be generated only by trials in the production system, additional costs arise, which are to be minimised through reduction of the experimental expenditure by systematic design of experiments.

To accelerate the learning process of Artificial Neural Networks, aiming to set this up as rapidly and efficiently as possible, an efficient tool for reducing the experimental costs can be adopted from the area of six-sigma philosophy with the Design of Experiments (DoE). DoE includes various methods for systematic planning, execution and evaluation of trials for determining and mapping statistical parameters and their effects and interactions (Schmitt 2010). It helps to determine systematically the optimal parameters for processes subject to variation of the influences relevant to the process result. The aim in deploying this method is to obtain meaningful information and optimal process parameters with the lowest number of trials possible.

For application in the actual implementation experiments were planned and carried out systematically with the aim of comparing the learning process of neural networks (a) by applying DoE and (b) in conventional training. Artificial Neural Networks were deployed to determine parameters when assembling a shaft-hub joint in experiments conducted. A trial for calculating a transmittable torque has shown that the experimental expense was able to be reduced to 18 data sets taking as a starting point a conventional training run with 56 data sets from components actually constructed, whereas the prediction accuracy of the experimental results of individual parameter sets reduced by only 2% in the actual example (From 2.5–4.5% in the selected results range between 80 and 120 nm: 8 data sets determined systematically with DoE, 10 equidistant selected data points; the requested data points were not part of the training data).

In the experiment cited an Artificial Neural Network was first of all trained randomly with production data. The aim was to assess the ability to transmit torque on the basis of materials and tightening torques when screwing together a shaft-hub joint (an actual example is described in Sect. 6.2.4.3.2). A further Artificial Neural Network was trained by deploying DoE, wherein first the data on attribute and value ranges was investigated and then an experiment was designed. This experiment enabled known data points to be evaluated systematically as relevant. The data stemming from this then served as training data for the second Artificial Neural Network.

In a second experiment the whole spectrum of the results range of available data sets was utilised. The prediction accuracy of the Artificial Neural Network was compared (a) with training and validation with all 56 data sets available, (b) by deploying DoE with 8 data sets and subsequent post training with an availability of 16, 24 and 56 data sets. An Artificial Neural Network type MLP was deployed with one hidden layer, 3 input neurones and one output neurone, each fully connected with all neurons of each neighbouring layer, and by employing back propagation training each with 10,000 runs.

By means of the 56 data sets a prediction accuracy having a mean deviation of 4.2% was able to be achieved and after deploying DoE with only 8 data sets a mean deviation of 8.4%, (b) whereas the prediction accuracy with increasing number of training data sets (c) rose further likewise to 4.2% and thus to the level of experiment (a).

The learning curve for the prediction accuracy thus climbs at an early point substantially faster than when carrying out the conventional experiment.

If this system is deployed purposely in the training phase of Artificial Neural Networks, previous data is available that can be deployed for a training run (Fig. 6.13). This acceleration in the start-up phase is only detrimental to the prediction accuracy to a minor extent and incomparably better results can be achieved than with the random choice of training data far earlier than with conventional methods. Therefore the learning curve of the prediction accuracy climbs at an early point substantially faster than when carrying out the conventional experiment and nevertheless can be increased in further trials within the range possibilities, which conventional experiment additionally offers. Translating this to an application example of an actual production system time (reasonably good predictions available earlier) and money (experimental expense) are saved.

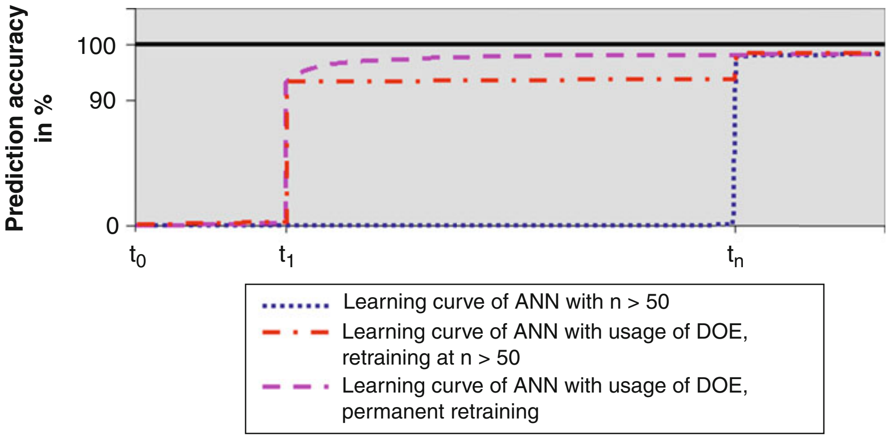


Fig. 6.13 Learning curves for artificial neural networks

Limits when deploying DoE however lie in the properties of the functions for mapping the interdependences. With a larger number of parameters and/or a non-linear function or a function able to approximate linearity only with large error, deployment must therefore be substantiated by studies. The DoE leads to a greater increase of the learning curve, but the prediction accuracy remains dependent upon the data employed.

Decision-making in the Optimisation Module

The scientific findings were established using, amongst other things, a specific shaft-hub model. This was based on a simplified problem suitable for developing and evaluating the cognitive architecture. In this respect, the ability to transfer a predetermined torque from the shaft to the hub represents the required functionality (product function). Here, the use case exhibits the property of controlling the transmission of torque through various influencing variables which influence each other through interactions. The software controls the assembly parameters for the shaft-hub connection so that the connection is capable of transmitting a required torque, but slips at a defined maximum load. Hence, the connection fulfils the function of a slip clutch or predetermined breaking point, which helps protect other components from excessive stress and the damage this can cause.

During the assembly process for the shaft-hub connection, the individual assembly parameters are adapted dynamically by the CTM software so that the transmissible torque or the torque required for the connection to slip is achieved regardless of the dimensional stability of the shaft and hub used. In detail, the connection is comprised of two hubs fixed onto one shaft:

Figure 6.14 is a schematic of the connection setup. The lower hub (hub 1) is connected with the shaft via a tapered interference fit. In this respect, the strength of the connection is determined by the tightening torque of the screw used to fix the hub to the shaft as well as by the material and design parameters. The upper hub (hub 2) is connected with the shaft via a three-jaw chuck. Here, the strength is determined by the

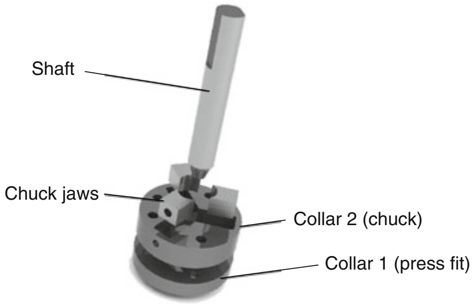


Fig. 6.14 Shaft-hub demonstrator

jaw material and the tightening torque of each screw. To simulate the predetermined breaking point, both hubs are fixed to one another and rotated against the shaft. The torque transmitted, i.e. the functional feature of the connection, can be measured.

The individual parameters for influencing the transmissible torque are dynamically calculated by the software. For this purpose, the assembly process for the shaft-hub connection is split into different subprocesses which are controlled by the software on a step-by-step basis. The individual steps are

- Hub 1 is predefined in each case,
- Selection of shaft,
- Assembly of tapered interference fit,
- Optional: measurement of the torque already achieved through hub 1,
- Selection of jaws and
- Assembly of the three-jaw chuck.

Thus, with respect to the assembly process, there is a wide variety of parameter variation possibilities that are calculated by the software and need to be checked. These variation possibilities allow a response to variations in the process, e.g. variations in the dimensional stability of the individual components. The Soar optimiser generates possible combinations of the individual assembly parameters, from which the modelling module extrapolates a resultant maximum torque. Next, the optimiser evaluates the torques provided by the modelling module with regard to the objectives to be fulfilled and varies the assembly parameters until the modelling module returns a result that meets the desired component function.

With reinforcement learning, the CTM system has an ability to learn that reduces the required queries of the modelling module and thus the computation time required for already known or similar starting situations (shaft-hub combinations).

Implementation of Reinforcement Learning in Assembly of the Shaft-Hub Connection

To be able to make an "intelligent" decision as to which of the actions or operators integrated in the Soar code should be selected, reinforcement learning is used. Reinforcement learning denotes a learning process based on reward and punishment. In Soar, positive or negative rewards are issued depending on whether an agent achieves its objectives and these rewards are allocated along the decision-making path that led to the corresponding result. The agent is thereby able to quantify the benefits of a decision and thus, step by step, learn the optimal way to achieve its objectives (Nason and Laird 2005; Sutton and Barto 1999).

At first, the agent makes random decisions about the subsequent operators. Once the agent has completed the search, the result is evaluated and reward points are issued for the decisions made depending on whether the objectives are met. The number of points awarded for fulfilling a particular objective can be fixed or freely selected by the programming. In the example of the shaft-hub production, dynamic point allocation was chosen. The nearer the operators bring the agent to the desired target torque, the higher the reward for the selected operators. Formula 1 shows the calculation for the rewards (r) as a function of the difference from the target torque (diff).

$$r = \frac{\max(r)}{(diff + 1) * (diff + 1)} \tag{6.1}$$

In further iterations of the agent, operators with high preferences are chosen more often. To avoid a situation where supposedly poor operators, or operators that have not yet been tried, are never selected, selection of a certain percentage of supposedly poor operators is always rewarded with a lower number of points. The frequency of selection of such operators can be freely determined. Experiments with 10 and 30% are presented below. This maintains the balance between exploration (undiscovered possibilities) and exploitation (of current knowledge).

Evaluation of Reinforcement Learning Using the Shaft-Hub Model

For evaluating reinforcement learning in respect of the shaft-hub production, a multi-step approach is selected. In the first step, there is no variation of hub 1 (tapered interference fit). On this basis, both the learning algorithms integrated in Soar are evaluated using Sarsa and Q-Learning. In the second series of experiments, hub 1 is varied too. The Sarsa and Q-Learning algorithms are analysed with a variable hub 1 as well.

Rewards points are distributed both after hub 1 assembly and after hub 2 assembly. For the transmission behaviour of the first hub, with 39–63 Nm transmissible torque, a comparatively greater range is considered able to fulfil the function. Ideally, the total torque should be 105 Nm; deviations of +/− 7 Nm are permissible. The agent receives positive rewards within these limits and negative rewards outside them. The points are issued according to the degree to which the objectives are met, as per formula 6.1.

In generating parameter proposals, it is important that new values are also generated in addition to the combinations already tried and tested. This is the only way to ensure that the optimiser does not just stick with in one local optimum, thereby leaving a potentially more favourable parameter combination undiscovered. Furthermore, a balance between tried and tested and new proposals facilitates a flexible reaction to variations in the production system. This deliberate selection of new parameter combinations is known as an indifferent selection strategy. In the experiments described here, the most widely-used procedure (Nason and Laird 2005) was selected using the epsilon-greedy method. This method provides the facility to specify an ε value of between 0 and 1 as the probability that new parameter combinations that deviate from the prioritisation will be selected. An ε of 0.3, therefore, means that in

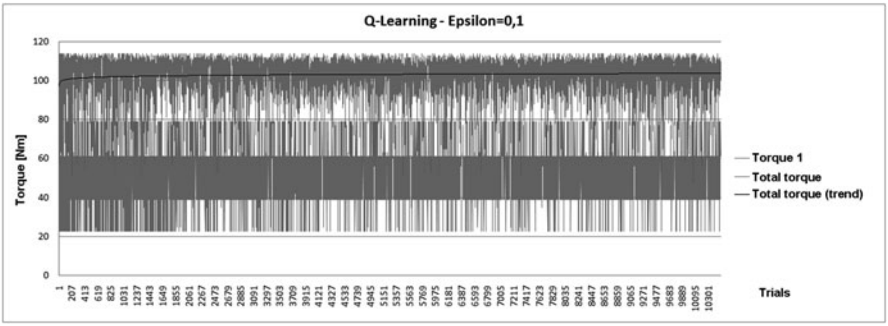


Fig. 6.15 Learning success without variation of hub 1

30% of cases, a new proposal will be selected (Nason and Laird 2005; Rummery and Niranjan 1994). Accordingly, when setting the value for ε, it is necessary to weigh up in each case whether the increasing flexibility produced by a higher ε is necessary, or whether the efficiency of the optimiser is reduced too much through more frequent testing of alternative parameter proposals. Frequently varying processes tend to require a higher value for ε whilst with very stable processes, the value for ε can be reduced.

Experiment 1: No Variation of Hub 1
With no variation of hub 1, the success of reinforcement learning can be measured at a very early stage. In carrying out the experiments, two learning algorithms were evaluated using Sarsa and Q-Learning. In addition, the learning rate ε was varied between 0.1 und 0.3. Figure 6.15 shows an example of this for Sarsa and an ε of 0.1. The diagram shows the transmissible torque resulting from the proposals of the optimisation module for hub 1 (torque 1) and the torque transmissible by both hubs together (total torque) across the total number of proposals executed. In all experiments, a significant improvement in the result quality is revealed after only approximately 100 proposals. With this learning success, the modelling module needs to be requested to find adequate assembly parameters less often, hence a result can be achieved in less time.

Figure 6.15 shows that over the entire length of the experiment, there are strong fluctuations in the predicted torque. These fluctuations are inevitable due to the use of the indifferent selection strategy and they enable a flexible response to changing boundary conditions. To compare the learning success of the various algorithms and the learning rates, the mean values and standard deviations of the calculated total torques are compared once at the beginning of the experiment and once at the end of the experiment. The study shows that after numerous experiments, a high epsilon also results in strong fluctuations of the predicted torque. This is reflected in an increased standard deviation of the values proposed by Soar (see Table 6.1). However, no significant differences due to the learning algorithm used are observed.

Experiment 2: Variation of Hub 1
If, in addition, the parameters of hub 1 are varied, this results in a significant increase in the complexity of the problem. The Soar program for manufacturing and

Table 6.1 Mean value and standard deviation without variation of hub 1

Learning algorithm	Epsilon	Tests	Mean value (Nm)	Standard deviation (Nm)
Q-Learning	0.1	1–30	89.67	17.23
Q-Learning	0.1	1,000–1,030	104.35	6.94
Q-Learning	0.3	1–30	97.58	16.73
Q-Learning	0.3	1,000–1,030	101.94	9.76
Sarsa	0.1	1–30	91.68	19.46
Sarsa	0.1	1,000–1,030	103.45	6.27
Sarsa	0.3	1–30	100.84	12.13
Sarsa	0.3	1,000–1,030	106.32	8.31

assembly of the shaft-hub connection contains approximately 100,000 rules. Due to the substantial increase in the number of parameter combinations, it takes longer to reduce the deviations of the total torque resulting from the proposals. A significant improvement in the result quality can be measured after approximately 2,000 proposals. Figure 6.16 shows an example of a learning curve for Q-Learning and an ε of 0.1.

Similar to the experiment without variation of hub 1, the graph shows strong fluctuations in the torque reached across the entire process. The fluctuations are much more pronounced over the first 1,000 tests. For a quantitative comparison of the learning success, as with the above tests, mean values and standard deviations are determined under variation of the algorithm and the ε. Compared with the permissible tolerance of $\pm$ 7 Nm, an ε of 0.1 generates very good values. With an ε of 0.3, the standard deviation of the total torque achieved increases significantly. In this experiment, which is characterised by a much higher problem complexity compared with the non-varying hub, Q-Learning exhibits significantly better results than Sarsa (see Table 6.2).

Rapidly identifiable learning success in relation to the possible parameter combinations is essentially connected with the distribution of reward points for interim results. In this case, the procedure makes it possible to evaluate unfavourable interim results, i.e. in this instance, the torque transmissible through hub 1, in advance. Thus,

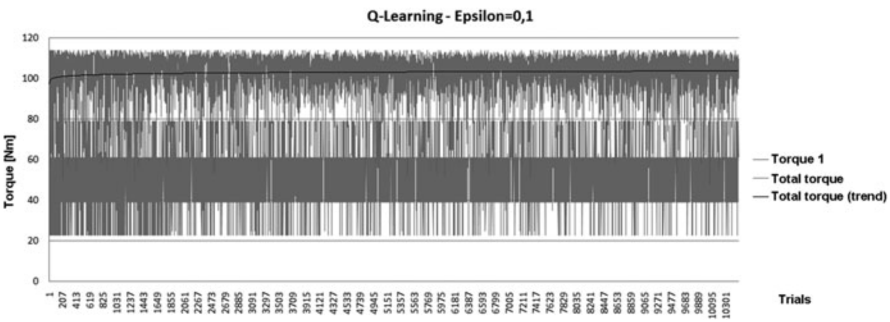


Fig. 6.16 Learning success with variation of hub 1

Table 6.2 Mean value and standard deviation with variation of hub 1

Learning algorithm	Epsilon	Tests	Mean value (Nm)	Standard deviation (Nm)
Q-Learning	0.1	1–30	92.85	18.41
Q-Learning	0.1	10,000–10,030	103.46	6.25
Q-Learning	0.3	1–30	97.42	15.23
Q-Learning	0.3	10,000–10,030	104.89	7.01
Sarsa	0.1	1–30	100.43	14.09
Sarsa	0.1	10,000–10,030	104.26	8.99
Sarsa	0.3	1–30	100.73	15.52
Sarsa	0.3	10,000–10,030	103.00	11.39

the potential, but not necessarily instrumental, degrees of freedom are effectively reduced without categorically excluding them through fixed rules or restricting the flexibility of the optimiser.

Generally, however, when using interim results, it is important to carefully check that, disconnected from the previous settings, they do not lead to interferences between the new and the previous settings.

Conclusion

Evaluation of reinforcement learning in the shaft-hub connection model is important for application of the CTM system in real industrial applications. These are usually characterised by greater complexity, i.e. a higher number of varying production parameters. Without an effective learning process, an optimiser with no physical model would not be able, within a practical period of time, to determine the production parameters that ensure the functional performance of the product to be manufactured.

However, the use of reinforcement learning in the shaft-hub model also demands a considerable amount of memory. At the beginning of optimisation, Soar initialises the learned value of each possible state with zero, i.e. the higher the number of possible states (in other words the number of possible parameter combinations), the greater the memory requirement of the optimisation program. The manufacturing software for the shaft-hub demonstrator requires up to one gigabyte of working memory depending on the degree of discretion of each parameter. The memory requirement increases exponentially with the concatenation of subprocesses. For this reason, preliminary analysis using data mining (see 6.2.4.3.3) is essential.

Another way to reduce the memory requirement, and thus also the computational complexity, is to split the Soar optimiser into several independent agents that process different sections of the production process. To create a working interface between the agents, it is important that only the information relevant for the next step is passed on. Since this can only happen reliably at certain points in the process, the use of multiple Soar agents must be tested in each case.

Data Analysis in the Analysis Module

The data recorded in the perception layer forms the basis for decision-making in the optimisation layer. Decision-making is implemented by the Soar cognitive architecture in cognitive tolerance matching. Both the effectiveness and the efficiency of decision-making decreases exponentially with an increasing number of parameters to be optimised ("curse of dimensionality", (Priddy and Keller 2005)). Hence for parameter optimisation that is useful in industrial practice, it is necessary to reduce the parameters to be optimised as much as possible.

In terms of parameter reduction, the challenge lies in identifying the parameters that have a significant influence on the product feature to be controlled. Often a large number of parameters, such as tolerances, dimensional measurements or machine settings, are available and their precise influence on the product feature to be controlled is not fully known. A prerequisite for controlling process chains with cognitive tolerance matching is to consider all relevant parameters that have an influence on the feature to be optimised. If a relevant parameter is not taken into account, the existing dependencies cannot be uncovered. With respect to data analysis, the challenge lies in reducing the scope of consideration by non-relevant parameters.

In the context of data mining, there are various statistical methods for interdependence analysis of large amounts of data. These make it possible to identify data or parameters that have a significant influence on the product feature to be controlled. As part of the research work, various data mining methods, including correlation analyses, principal component analyses and various classification methods, were examined for applicability in the cognitive tolerance matching system. A methodology was developed that can be established upstream of an optimiser in such a way that the optimiser only has to consider a greatly reduced number of parameters and hence the memory requirement and performance of the optimiser remain within practical limits, but the relevant area of the parameter space is covered.

Decision Trees as a Method of Data Mining

Based on the advantages described in Sect. 6.2.3, decision trees were identified in a preliminary study as the data mining method most suitable for CTM and thus were evaluated in detail.

Figure 6.17 shows an example of a decision tree for the quality of a shaft-hub connection. A decision tree is a directed cycle-free graph with a tree-like structure

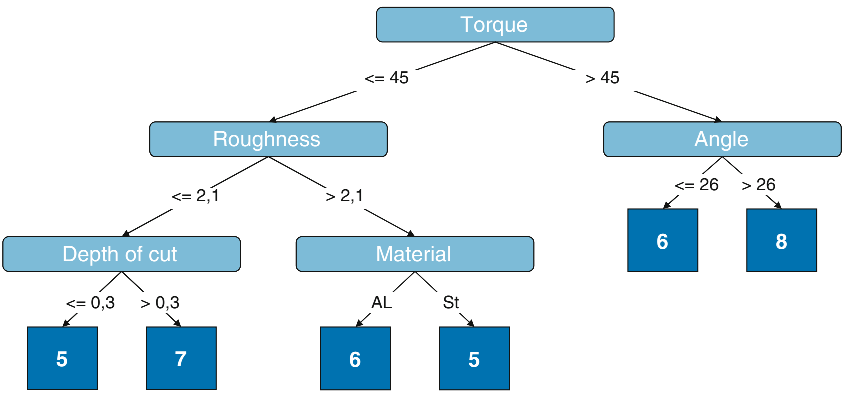


Fig. 6.17 Illustration of a decision tree

comprised of a root, edges, nodes and leaves. In the tree, each node denotes the test for an attribute, each edge indicates the result of such a test and each leaf represents a class. A decision tree makes it possible to intuitively visualise dependencies in relation to a product feature to be examined. At each node of a decision tree, all the available attributes are ranked. The attribute that is most useful for partitioning the tuple at this node into classes is selected as the test criterion. In the relevant literature, this test criterion is also known as the splitting criterion (Han and Kamber 2006; Agrawal et al. 2004).

In the leaves, the quality of the connection is rated using values between 5 and 8. The nodes and edges quantify the influencing factors that lead to achievement of a particular quality class. Hence a connection with a torque of 50 Nm, for example, and an angle of 31° is rated with an 8.

Evaluation of the Method

With respect to the applicability of data mining algorithms in general and for decision trees in particular, the effectiveness of the various algorithms was investigated using actual industrial data. The data used for this was from the industrial application of BMW rear axle drive production. In gathering the dataset, the aim was to use data mining to identify the parameters that have a significant influence on the acoustic behaviour of a rear axle drive. Therefore, ideal acoustics is the product feature to be optimised by the CTM system. In a quality control loop, the acoustics is the control variable.

To decide which parameters to include in the study, expert workshops were held at BMW in order to select all variables that, in the experts' opinion, could influence the acoustics. As a precautionary measure, all parameters whose influence could not be ruled out were also selected. In total, 193 different parameters were identified in the expert workshops. To gather the full database with respect to these parameters, 80 rear axle drives were produced and measured completely. The resulting data thus consists of 193 different measured and preference values which represent the process chain of production of a rear axle drive for the optimisation variable "acoustic behaviour". Figure 6.18 shows the distribution of the identified parameters along the process chain of rear axle drive production.

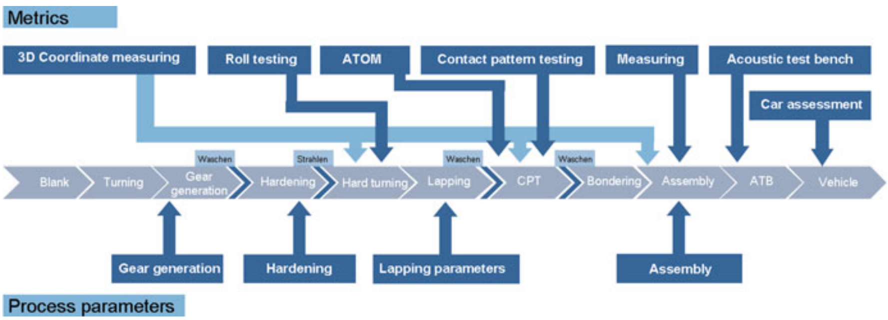


Fig. 6.18 Influencing parameters

Table 6.3 Assignment of acoustic test bed and evaluation indicator

	KPI <= 5			
KPI =>7	3,75	4	4,25	4,5
1,5	x	x	x	x
1,75	x	x	x	x
2	x	x	x	x
2,25	x	x	x	x
2,5	x	x	x	x

The acoustic behaviour of a drive is evaluated using a subjective key performance indicator (KPI) between 1 and 10. The aim is to achieve the highest possible value for the rear axle drive acoustics. A value greater than or equal to 7 is deemed exceeding requirements and a value smaller than or equal to 5 is generally perceived by customers as annoying and they would complain.

The KPI is heavily influenced by the subjectiveness of the test participant. For the sake of objectification, the measured values of the acoustics test bed were assigned to various KPIs. To this end, first the measured structure-borne sound signals of the acoustics test bed were standardised and then totalled for all measured axes. The measured values cannot be rigidly assigned to the KPIs because slight deviations may arise depending on the perception of the examiner making the evaluation. To compensate for subjective deviations, Table 6.3 was developed as an experimental design for variable assignment. The table shows 20 possible assignments for acoustically unobtrusive drives (KPI greater than or equal to 7) and acoustically obtrusive drives (KPI smaller than or equal to 5). The crosses each stand for one possible assignment of the limits for obtrusive and unobtrusive drives. The numbers represent the standardised and totalled measured values of the acoustic test bed. For example, the cross on the left on the first line means that a measured excitation smaller than 1.5 indicates an unobtrusive drive and a measured excitation greater than 3.75 indicates an obtrusive drive.

The evaluated measurement variable of the test bed is the integrated amplitude of individual frequency bands that can be perceived as annoying in the vehicle. To avoid subjective influence with respect to the evaluation indicator rating, decision trees were created for all 20 possibilities. For creating the decision trees, the three open-source methods “information gain”, “gain ratio” and “Gini index” (Han and Kamber 2006) were evaluated, along with “Accuracy”, a proprietary method of the “Rapid Miner” software used. Combining the four methods presented with the 20 possible evaluation indicator assignments produced, in 80 validation runs, decision trees which formed the basis for parameter classification independent of potential method influences.

The results from the analysis of the data mining process show that these vary considerably from method to method. Due to the significant differences in the data structures of various industrial applications, valid selection of a single method is not possible at this point in time. Automated use within a CTM architecture for prioritising the data gathered is therefore not feasible. The challenge in selecting the method suitable for the data to be examined lies in not knowing the relationships within the existing data. If it is not known whether, or in what way, dependencies exist between individual parameters, then precise selection of a suitable method of analysis is not possible. Hence, combining the individual methods was investigated as a possible solution. Adding up the results of the individual methods across all 20 evaluation indicator combinations clearly highlighted certain parameters so it can be assumed that these are of greater relevance. For example, the circumferential backlash and the block size of the gear set were identified as the most influential parameters with respect to the acoustics of the rear axle drive. A feasibility test by the experts at the BMW Group has confirmed the parameters obtained through data mining are appropriate. Industrial application of the methods therefore also seems thoroughly appropriate. However, this cannot be automated as yet; preparation with the help of the experts is needed.

Using data mining and decision trees in this way, complex process chains can be analysed with a variety of parameters that are considered relevant. Identifying subprocesses and individual parameters that have an increased influence on the product feature to be optimised only enables a cognitive optimiser to be used for those subprocesses.

Control Via the Coordination Module

To coordinate the individual tools, a comprehensive module has been installed that forms the interface to the user, controls execution of Soar programs as well as those of the modelling module and is responsible for the data exchange between individual modules. In real-time control of a production system, the coordination module also connects the production system, and thus the production machines, by transmitting the settings and control commands created for actions via the operation layer.

Soar Sequential Control

A Soar program is executed in “steps”. Each step involves the execution of one Soar rule and thus the making of a decision. Between the executions of individual steps, it is possible to make Soar data available, or to read data from Soar. The coordination module controls the sequence of the individual Soar decisions and is responsible for the data exchange to and from Soar. For controlling a Soar agent, there is a C++ library which makes it possible to load Soar codes and execute them step-by-step as well as read and describe the individual variables of the Soar working memory.

Modelling Module Control

Similar to optimisation module control, the coordination module controls the data exchange between the modelling module and the other components. In the specific implementation of the software, the modelling module is represented by the artificial neural network software SNNS (Zell 1994). SNNS stands for “Stuttgart Neural Network Simulator” and simulates the components and processes of neural networks using technical means so that they may be recreated using a computer. SNNS allows artificial neural networks to be exported as a software library that can be linked directly to the coordination module and enables batch-based execution and querying of network actions.

The coordination module uses these functionalities and can access the relevant SNNS data and functions directly through specially designed interfaces in order to execute high-performance network queries. In addition to purely querying a network, the coordination module supports post-training of an existing network during run time. Hence in the event of deviations between a value predicted by the artificial neural network and a real measurement result, the network can be post-trained using the relevant data points (i.e. measurement results). The functionality provided by SNNS was extended accordingly for this purpose.

Storage of Knowledge and Experience in the Knowledge Module

Knowledge is provided in cognitive tolerance matching within Soar and by artificial neural networks. Soar provides two concepts for representing knowledge in the form of a long-term and a short-term memory:

The long-term memory stores operative knowledge for forming the problem spaces and control knowledge for controlling the search processes consistently in the form of productions.

The Soar memory model (Fig. 6.19) provides a more precise description of the long-term memory: it is divided into the semantic, the episodic and the procedural memory. Declarative knowledge, i.e. semantic and episodic knowledge, is organised in the form of so-called semantic networks. Short-term knowledge is represented by attribute-value lists which are grouped together to form objects. The short-term memory forms the working memory where all information processing takes place. The short-term memory is characterised in particular by its limited storage capacity.

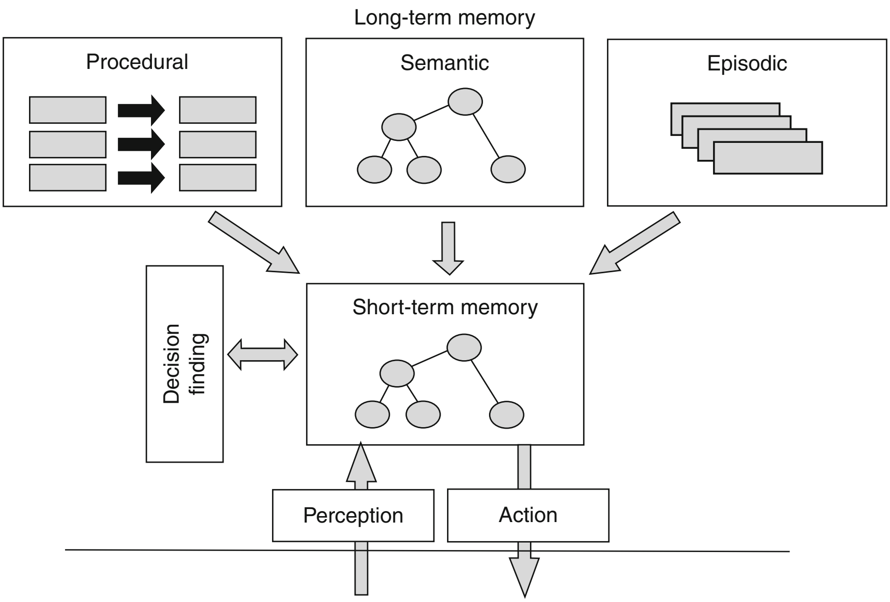


Fig. 6.19 Soar memory model

Long-term information must therefore be transferred into the long-term memory. Short-term memory is often referred to as working memory. The short-term memory processes the current to-do tasks.

The uniform representation and access mechanism as well as the ability to structure the working memory into areas means there is a strong similarity to blackboards. The open design of the working memory allows the addition of any modules that can use this memory, or an assigned segment thereof, for information exchange and coordination.

Information processing involves two phases. In the first phase of the knowledge search, applicable long-term memory productions which operate on the working memory fire. On the one hand, this process leads to the generation of new objects which can in turn allow other productions to apply. On the other hand, preferences emerge and vary, and these are then used in the second phase for control with respect to the further application of goals, problem spaces, states and operators.

In the second phase, based on the current knowledge in the short-term memory and with additional help from existing preferences, the decision procedure selects an operator and applies this to the associated problem space. Through successive application of operators, either the goal is met at some point or a dead end is encountered. In this case, a subordinate goal is generated and its task is to direct the search process out of the dead end. If a dead end cannot be resolved this way, then the problem space reverts to independent mechanisms such as backtracking (Laird and Bates Congdon 2008).

Through the use of reinforcement learning, the knowledge base is dynamically expanded during run time. The learned knowledge is taken into account when further decisions are made, thus reducing the likelihood of reaching a dead end (Nason and Laird 2005).

The knowledge about cause-effect relationships in the model is stored in artificial neural networks which serve as a black box model. To this end, using saved data, artificial neural networks are trained to calculate or estimate results for possible input parameters of the production system. If the input parameters are applied to a calculated result in real production, the real result obtained can be compared with the calculated result. If a significant difference is discovered, the model is updated with the corresponding newly acquired data through the learning process used, hence the newly acquired knowledge is added to the implicitly stored pool of experiences.

Data Acquisition by the Perception Layer

As a provider of information to the CTM architecture, the perception layer is dependent on the processes and machines to be controlled. The perception layer's task is to acquire all the characteristic variables of a production process that are considered relevant. Preliminary data analysis in the analysis module and optimisation by the control software requires the continuous recording and digital availability of all relevant production data. However, the challenges in this respect lie not only in the interfaces that must be created between the production means and the IT systems concerned, but also in interpreting the data in order to obtain useful information from it. Often, data is not available as concrete figures and thus not in a form that can be immediately analysed; instead, it often needs to be converted from a qualitative

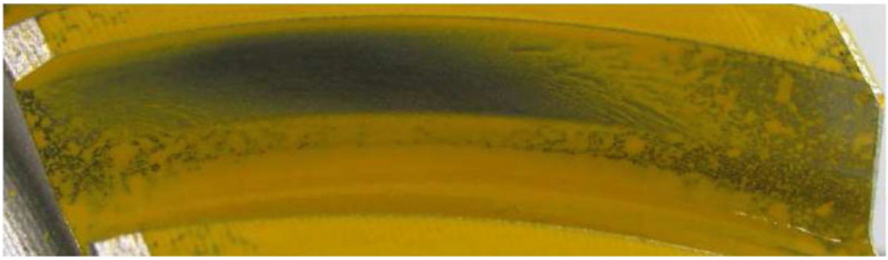


Fig. 6.20 Contact pattern of a ring gear

form into quantifiable values before the data, classified by the experts as potentially relevant, can be analysed in terms of its significance.

Hence when assessing a rear axle drive, for example, the contact pattern (Fig. 6.20) is an important evaluation indicator. The contact conditions of a gear set pair are visualised using contact pattern paste. Based on the quality of the tooth flank contact, conclusions may be drawn regarding, for example, the acoustics of a drive. However, to date, evaluation of the contact pattern has not been automated; instead, it has been conducted by trained staff and is therefore always subject to subjective influences and evaluation criteria. The evaluation results vary depending on the examiner and are therefore only suitable for further digital processing to a limited extent.

It is therefore necessary to evaluate this data deterministically and thus create reproducible evaluations. In the specific use case of the rear axle drive, an image processing system for automatic contact pattern recognition is developed. The measuring device is equipped with cameras to create colour images of the tooth flanks. Once the image has been captured, the tooth flanks are detected and extracted. Using a variety of graphic image processing algorithms, the contact patterns are identified and classified so they can then be converted into binary matrices. The relevant parameters of the contact pattern of rear axle drives are determined using these matrices. These parameters are optically recorded by the developed measuring device, they are captured by digital image processing and are thus used as input parameters for the optimisation software. In the next step, they are used by the analysis module which identifies relevant characteristic variables with the aid of data mining algorithms for the production process.

6.2.5 Industrial Relevance

The boundary conditions for the manufacturing industry in high-wage countries have changed immensely in recent years. Due to increased global competition, besides the demand for high-quality products, the demand for cost-effective production processes, in particular, has also grown. Many companies are no longer able to meet these challenges in Germany and are shifting their production to low-wage countries. There, lower wages and social security contributions as well as economies of scale through cost-effective mass production are leading to an apparently higher productivity (Pfeifer and Schmitt 2006; Beckmann 2009).

Another global trend is that customers are increasingly demanding individualised products available to them within a very short period of time. This is particularly evident in the automotive industry where customers are able to amend the configuration of their vehicle even shortly before its delivery (Lindemann 2006; Krause 2007).

So, successful companies must be able to adapt quickly to individual customer requirements which, in addition to short lead times, also requires an increased flexibility in production as well as stable processes (Pfeifer and Schmitt 2006).

Manufacturing processes for technically sophisticated products, in particular, increasingly have such complex dependencies between a high number of variable production parameters with an influence on the product to be produced that those relationships can no longer be fully comprehended. This leads to decreasing process stability. Production processes must be continuously optimised to achieve the desired product features and tolerances. However, these boundary conditions do not allow any last-minute product variations and, in addition, the processes are often not ideally designed because they can usually only be optimised in parts, especially in the case of multi-stage production processes. This results in the separate optimisation of individual elements of a system without the ability to fully assess the interaction of these changes with respect to the properties of the final product (Best 2010).

Function-oriented optimisation of the end product is more effective than optimising individual production steps. This is possible if individual targets, e.g. the dimensions to be obtained, can be adapted dynamically within the production process. That way, the superordinate goal, namely the functionality of the product, can be achieved.

For one thing, this kind of dynamisation of key process parameters will help reduce costs because individual tolerances can be selectively expanded without overlooking the required product features. For another thing, the flexibility of the production process will be greatly increased with respect to product modifications.

Self-optimising production systems are a crucial method of dynamisation. They allow value stream-oriented approaches to be adopted whilst at the same time enabling planning efficiency to be increased through the transfer of previously acquired knowledge into similar scenarios within production engineering. This facilitates brand new approaches for both production and assembly systems that constantly analyse and evaluate the current situation and dynamically adapt the system to changing goals.

Cognitive tolerance matching allows cross-level approaches to be developed in the field of coordination, planning, control and man-machine interaction. Hence, a more integrative framework of action is generated, enabling production systems to self-optimise with respect to different goals. This is achieved through the creation and implementation of cognitive mechanisms at planning and organisational levels across processes and process chains and through the creation of the ability to communicate in the areas concerned.

The business and technology case of rear axle drive production at the BMW Group in Dingolfing, Germany, is used to exemplify the industrial relevance. At this site, cognitive tolerance matching is used to optimise the production of rear axle drives with respect to the acoustics emitted during operation. The rear axle drive selected

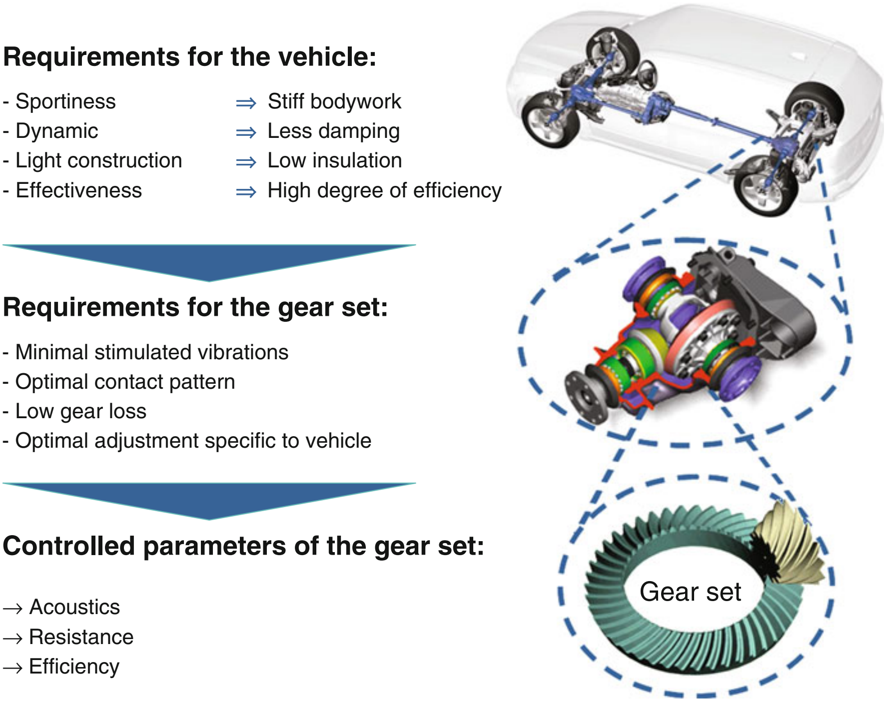


Fig. 6.21 List of requirements on the Hypoid gear sets for passenger vehicle axle drives of standard design

for the development of cognitive tolerance matching has a pronounced effect on the acoustics within the vehicle. It has very complex and multifaceted mechanisms that generate structure-borne sound which is then transferred to the interior of the vehicle via the car body. When the noise reaches a particular level inside the car, it may be perceived by the customers as annoying. The axle drive acoustics can sometimes be adjusted depending on the vehicle via the axle drive teeth (gear set). This presents a major challenge in connection with the occasionally contradictory demands for a high level of efficiency and strength. The vehicle acoustics, and therefore also the axle drive acoustics, are product features that are important to the customer (see Fig. 6.21).

Figure 6.22 illustrates the production process. It ranges from production of the gear set through to assembly and is characterised by many sensitive tolerances with complex dependencies.

The approach for optimising the acoustics is to optimise the specified tolerances throughout the entire production process. This means both expanding any unnecessarily tight tolerances to save costs and defining the critical tolerances more precisely to ensure the desired functionality of the end product.

The challenge lies in understanding the process, a pre-requisite for implementing the right optimisations at the crucial points.

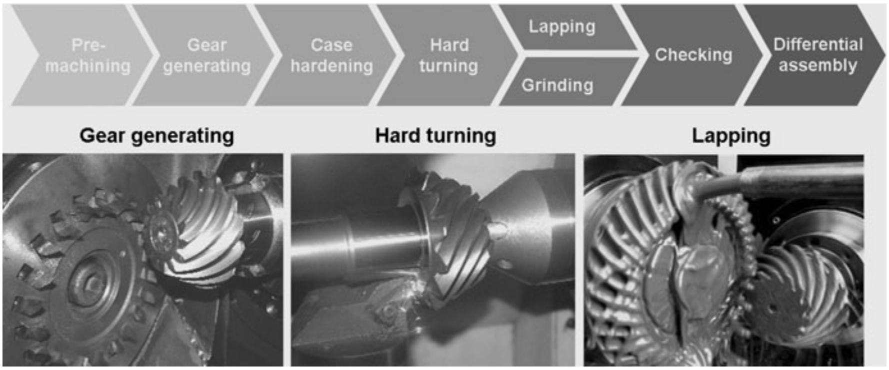


Fig. 6.22 Illustration of the process chain for manufacturing a hypoid gear set

Besides the objective of producing quieter vehicles, it is also important to identify the influencing parameters relevant for vehicle acoustics in design and production and to ensure that vehicles can be produced with constant acoustics without spread. The challenge here lies not just in merely identifying the acoustically relevant vehicle components and parameters, but, more specifically, in modifying them effectively. In this respect, the complex acoustic interactions between the individual vehicle components need to be taken into account.

Figure 6.23 illustrates the impact of cognitive tolerance matching throughout the entire production process. Firstly, the knowledge gained from production, assembly and use is fed back into development so that the production parameters can be adapted

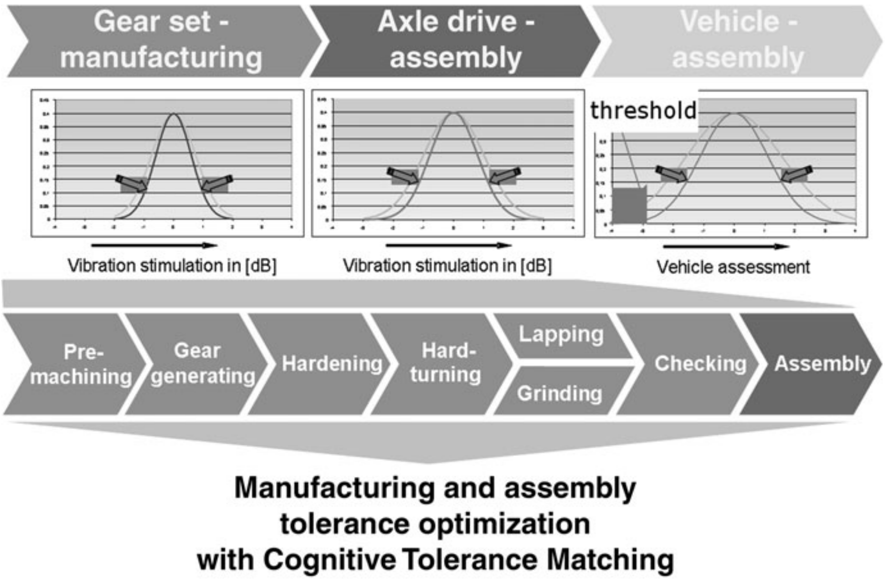


Fig. 6.23 Implementation of CTM in the production process

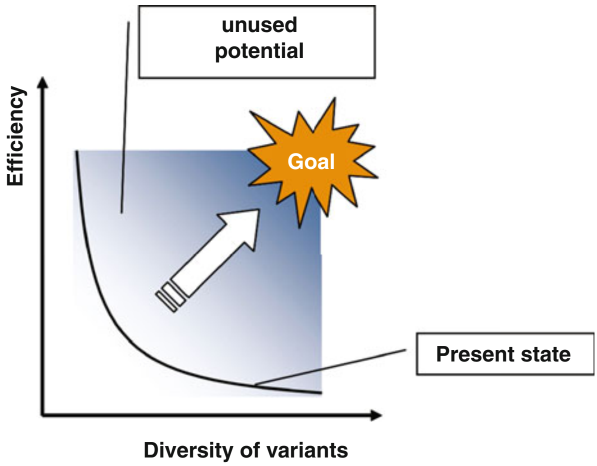


Fig. 6.24 Conflict between product variety and efficiency

accordingly. Secondly, through cognitive tolerance matching, individual process steps are interlinked, making it possible to respond immediately to deviations in the process. Hence, CTM contributes to the continuous improvement process in production. It also reacts specifically to deviations in upstream process steps, thereby ensuring the superordinate goal of an optimised drive in terms of acoustics, strength and level of efficiency can still be met. Thus the focus of deliberations is less on tolerance-oriented and much more on function-oriented production. Through these two mechanisms of action, the perceived quality can be both stabilised and increased.

In particular, the demand for increased product variety with increasingly complex products causes decreasing efficiency of the production process (see Fig. 6.24). For Germany to continue to be attractive as a production location, production efficiency must be increased as product variety increases.

The industry has recognised that in addition to focusing on how individual processes work, it is necessary to adopt a holistic view of the causal relationships within the value-adding chain.

6.2.6 Future Research Topics

Research and development in the field of self-optimisation in production and assembly already offers some promising approaches in the form of specialised solutions for individual problems. Technologies for implementing self-optimising applications in production engineering are being identified and successfully applied both in this and in other related research projects, e.g. in assembly scheduling (Sect. 6.6) or in production process parameterisation. Here, using self-optimisation approaches, significant improvements have been achieved in production and competitive advantages have been created. However, the research described herein presents just a small fraction of the possible applications that, in theory, can be improved through cognitive automated systems in production. Accordingly, there are other research tasks which may yield additional potential for these approaches. Deriving a systematic approach from the research presented makes it possible to transfer the methods to other areas of application. In this respect, their use is not limited to the operational level of a business, but can be applied at the planning and control levels too and, furthermore, these can be linked with the other levels. The following areas are therefore suggested as priorities for continuation of the research work:

- Development of a control model for the various production engineering applications within the levels of a company as well as on an inter-level basis: for planning the implementations of self-optimising control loops and the applications of cognitive components, it is necessary to model the use case as well as the solution components. For this purpose, analogous to classical control technology, mapping facilities need to be designed. In this respect, models need to be created in such a way as to be cascading and able to adapt to the granularity of the respective considerations.
- Development of modelling languages and standards: for modelling the applications of self-optimising control and regulation systems, it is necessary to develop modelling standards and languages that can serve as the basis for data exchange both between the models and between actual applications. The aim is to create standardised interfaces between the controlling and the controlled components. This includes creating semantic links between the individual cognitive architectures and technologies.
- Development of a software architecture as a modular framework: mapping the software technology in a modular framework allows standardised creation of control and regulation applications according to uniform guidelines. That way, different approaches can be concurrently evaluated quickly and easily so that the respective advantages and disadvantages can be compared instantly. The framework should be created using data structures based on the modelling languages to be developed.
- Further development of existing cognitive technologies: universal cognitive applications, such as SOAR, 3CAPS and ACT-R, offer significant potential for cognitive decision-making using technical means. To best exploit this potential, the capabilities need to be expanded with respect to the issues involved in production engineering.
- Evaluation of technologies for self-optimisation and cognition: the various technologies and cognitive tools used in the context of the Cluster of Excellence for implementing self-optimisation should be examined for their applicability in other cases. Capability profiles are useful in terms of selecting the best tool for the respective use case. Benchmarking in different applications provides support for such selection. A systematic selection methodology helps validate the decision for a specific technology.
- Evaluation and comparison of learning processes: learning processes provide the basis for generating knowledge from the available data and using it to positively

influence future situations. Different learning processes demonstrate different strengths and weaknesses depending on the application and these have to be balanced out against each other in accordance with the problem. The aim is to identify and implement stable and efficient learning processes depending on the application.

- Analysis of the effectiveness of cognitive and self-optimising technologies with respect to increasing the flexibility of production systems: individual studies within the problem area in question provide support for computer-based analysis of the effectiveness of various approaches as well as quantification of both the effectiveness and the efficiency prior to implementation.
- Cooperation of cognitive systems: cognitive systems must be capable of cooperation in order to be implemented on a cooperative or cascading basis at the same level. This includes strategy and sequence planning, coordination of a common component handling approach and the execution of transfers to the respective system boundaries. In addition to cooperation between machines, this also includes cooperation between man and machine and thus leads on to the task field of man-machine interaction.
- Man-machine interaction and cooperation: this includes the development of concepts and technologies for intuitive and safe interaction between man and machine in order to best combine the respective advantages of the possible scenarios. Amongst other things, this includes considering the challenges posed by the demographic change in production.

The afore-mentioned task fields address the question of how self-optimising control and regulation systems can be most effectively used throughout a production process in order to enhance the competitiveness of a production operation. This requires modelling the control loops to be depicted and modelling the specific use case in order to create a basis for identifying the right technologies for implementing the control loops. The optimal degree of automation with respect to throughput, flexibility and scalability needs to be identified through comparative assessment of possible approaches. In this connection, the optimal degree of integration of conventional systems and manual activities of human workers also needs to be determined. Changes in the boundary conditions (throughput, error rate) may also require adaptation of the production system configuration, even during the production, in order to enable a faster response to unforeseen interruptions. One example of this is an adaptive production line, e.g. in the automotive industry, which allows cognitive automation of customisable assemblies for multi-variant products with the involvement of humans.

6.3 Integrative, High-resolution Supply Chain Management

Günther Schuh, Tobias Brosze, Fabian Bauhoff, Niklas Hering, Simone Runge, Maik Schürmeyer, Sascha Fuchs and Till Potente

6.3.1 Abstract

6.3.1.1 Initial Situation and Problem

From the perspective of logistical planning and control, managing the increasing dynamics in customer-specific production and assembly is the most important challenge of the next few years (Zäh 2005). Causes of the increasing internal dynamics are shorter delivery times, greater variety of production and assembly processes (caused by the increasing product variety) and the use of more complex production systems (substitution of capital factor for work factor). The drastic reduction in delivery times has fundamentally changed the order situation and the capacity requirements of producing companies (Wiendahl 2006, p. 29 ff.). The necessary reductions in throughput time have only been made possible by the corresponding reduction of work in progress. This reduction in stock has necessarily led to more intensive coupling of the individual product resources. Capacity fluctuations and process instabilities in one resource thus affect the stability of the overall systems far more because stock can no longer be used as a buffer. At the same time, macroscopic industry-wide capacity fluctuations within the supply chain are increasing because there is no time buffer.

The increasing variation in process chains and times potentially intensify the effect of the capacity and throughput fluctuations described above. The “average value-based PPC” therefore no longer fits the requirements (Günther and Tempelmeier 2007). Conventional planning and control concepts which cannot respond to this variation necessarily impose further fluctuations on the system. The lack of a buffer and fluctuating requirements lead to loss of efficiency and increasing backlogs in production.

Absorption of the dynamics by means of stock levels and decoupling or maintaining reserve capacities for process synchronisation is no longer possible for reasons of cost pressure and customer-specification of products. In fact, new approaches for the planning and control of internal and inter-company production processes are required. These solutions need to cope with the dynamics of processes and capacity requirements and to incorporate these into the superordinate network and to engender further solutions to absorb the dynamics taking into account commercial and logistical goals.

6.3.1.2 Purpose of Research

High Resolution Supply Chain Management (HRSCM) describes the creation of a company-wide information transparency, which includes the interfaces with value chain partners, with the aim of ensuring product availability through decentralised, self-optimised control loops in industrial value chain networks. HRSCM pursues the aim of enabling organisational structures and processes to self-optimise, i.e. to adjust, according to consistent goals, to constantly changing framework requirements by means of decentralised production control mechanisms in the form of a cascaded control loop model. The designation "High Resolution Supply Chain" therefore relates to the almost unlimited transparency achieved at practically all levels of the industrial supply chains through the ubiquitous use of IT.

Information transparency plays a key role in the planning and control of a producing company. It is the basis for making decisions about the actual need for action and also provides support for actions to be undertaken. The central challenges are the complexity of information and their management along with the effective integration of human intuition and experience into the control loop of the supply chain management.

On the basis of adequate information transparency, this project pursues the scientific aim of radically improving the planning quality of planning and control processes while simultaneously reducing planning times and costs. This can only succeed when the currently prevailing average value-based planning logic of the MRP II concept is resolved by a decentralised actual data-based planning logic.

6.3.1.3 Results

To manage the initial situation and to implement the above mentioned aims, a comprehensive model of cybernetic production management was developed which represents the structural framework of the HRSCM approach. Furthermore, in order to enable the systems and subsystems to optimise dynamically and independently, logics were adopted from control theory and transferred to production planning and control and the optimisation of company processes.

To create a cybernetic production management model firstly a unified understanding of the production system processes to be controlled was required. For this purpose the processes and information flows involved in technical order processing and the tasks of the order-independent networking and cross-sectional functions were described The basis was the Aachen PPC model extended by Schmidt, in which the traditional processes of production planning and control were augmented by the processes of construction, assembly, dispatch and commissioning as well as project monitoring and control (Schmidt 2008; Schuh 2006). The result is a process-orientated illustration of the tasks and related information flows required to execute the operative order processing processes and thus a description of the production system to be controlled.

Thereafter the basic structures and mechanisms of a viable and changeable system were described from the perspective of management cybernetics. The principles, organisational structures, elements, information channels and mechanisms required by a cybernetic management system in order to manage the variety of complex system under dynamic influences were specified. The necessary expansion of the Aachen PPC model was then defined on the basis of these results with respect to the requirements of the cybernetic management theory. It was found that the model needed to be augmented by the principles of recursiveness, autonomy and synchronisation of target systems. On the other hand, it was also demonstrated that essential elements for the recording and use of high resolution information were as yet insufficiently defined.

The core of the cybernetic production management model is a reference model derived from the modules outlined above. Thereby the viable system model served as a structural template for the variety-orientated development and location of production management tasks, processes and information flows within an all-encompassing production management model. The architecture of the model is based on the order processing processes as the base units and is divided into four recursive structural levels. The autonomy principle was taken into account in the allocation of tasks in a way that in normal cases an order can proceed through the operative processes without intervention from the management system. To enable a cybernetic exertion of the management tasks, taking into account the manageability of the subtasks, the overall system was divided into five management units.

Based on the cybernetic production management model, control logics were applied to production management. Closed loop systems were modelled to permit automatic adjustment of the actual parameters. These dynamic adjustments of planning parameters require minimal planning time due to the high degree of automation but nevertheless permit high planning quality on the basis of data currency and granularity.

6.3.2 State of the Art

In the following, a summary of the state of the art in production planning and control is given. Furthermore approaches to the changeable design and control of production systems (Sect. 6.3.2.2) are presented.

6.3.2.1 State of the Art in Production Planning and Control

The operative planning and control of production is the core of production management. Current PPC systems are the result of various stages of development since the 1960s. The Order-Point System and Material-Requirements Planning (MRP) were the first of these development stages. Unlike the Order-Point concept, MRP is based on deterministic requirements in which part lists and associated gross secondary requirements are incorporated into the demand planning as influential values (Dangelmaier and Warnecke 1997, p. 258). The net demands which result in the order or the initiation of a production order are the result of balancing the gross secondary requirements with stock levels (Orlicky and Plossl 1994, p. 24). This material-requirement determination logic is still used today due to its simplicity and simple calculation algorithms, primarily for push-control of expensive and sporadically required parts (Ahsan et al. 1996; Askin and Goldberg 2002; Kiener 2006). Closed-Loop MRP was developed in 1969 as a refinement of the traditional MRP concept. At the heart of this refined concept is the assurance that all capacities required to implement the production plan are available (Kurbel and Endres 2005, p. 135 f.).

The term Manufacturing Resource Planning (MRP II) was coined in 1984 by Oliver Wight (Wight 1984, p. 51 ff.). Its most important addition is planning according to limited capacities. The integration of commercial and sales planning also enables analysis of implicit and financial aspects in PPC (Ahsan et al. 1996, p. 19).

ERP systems were first developed in the early 1990s, adding accounting-orientated modules, complex modules for operative production planning and control, maintenance, order management and human resources to the MRP-II approach (Kiener 2006, p. 267; Busch and Dangelmaier 2004b, p. 424 f.). The introduction of ERP systems and the use of Electronic Data Interchange (EDI) greatly improved communication and therefore also coordination within the supply chain. However, the MRP-II concept has some fundamental weaknesses. The centralised and push-orientated MRP-II planning logic, according to Pfohl, cannot adequately detect or plan the dynamic processes which frequently occur in a production environment as a result of problems (Pfohl 2004, p. 160). Weaknesses of MRP-II-based systems include lack of support for order release, a planning principle based on average values, the successive planning method and the use of limited partial models. The successive planning method breaks down PPC tasks into smaller task packages which prevents holistic analysis and achievement of optimum solutions (Hellmich 2003, p. 208). It is not possible to plan according to a higher-level commercial target system using the partial planning approach because of its inherently isolated partial analyses and reduced computation complexity (Finkler 2006, p. 23 ff.; Hufgard 2005, p. 152 f.; Kiener 2006, p. 269). Further weaknesses of MRP-II-based PPC systems according to Kletti are the inadequate analysis of the prevailing load horizon and the actual capacity exploitation, absence or delay of feedback about order progress, problems and poor information availability and transparency (Kletti 2006).

Since the 1960s there have been many attempts to expand the MRP-II approach with the aim of eliminating these weaknesses. At the end of the 1990s, the MRP-II concept was supplemented by so-called APS-Systems (Advanced Planning and Scheduling systems) to make them more useful by systematic improvements in planning (Gesatzki 2002, p. 1). APS systems are modular software systems which enable the integrated planning and control of companywide business processes (Albert and Fuchs 2007, p. 6). Exact mathematical optimisation processes are used for planning as well as heuristics for bottleneck-orientated planning (Schwindt and Trautmann 2004, p. 5). The use of APS systems makes it theoretically possible to incorporate all restrictions and available capacities simultaneously into the planning process and to continually update the plans (Albert and Fuchs 2007, p. 6 f.). Unlike ERP systems, APS technologies can be used to analyse alternative planning scenarios. In this case existing problems are not only resolved by increasing the efficiency of processes towards the customer but also optimizations at the interface to suppliers are taken into account (Gesatzki 2002, p. 4). APS systems are an attempt to obtain a more accurate prediction of the future using mathematical models with refined data. The

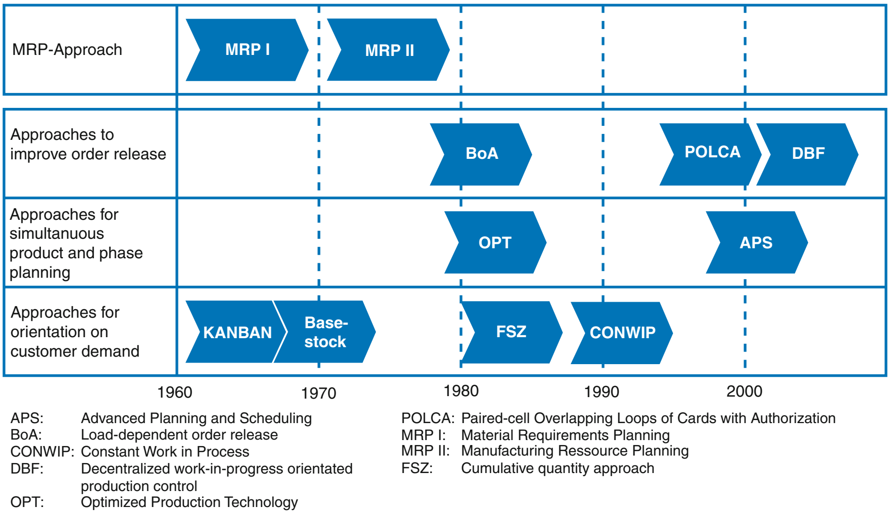


Fig. 6.25 The development of PPC concepts. (Meyer 2006)

computational results are nevertheless based on assumptions which do not fit the dynamic company environment, for example "an average production time per part".

Figure 6.25 provides an overview of the development of PPC systems from 1960 to the present day.

The following approaches represent the next development stage of the PPC whose basic characteristics are increasing decentralisation and self-organisation within PPC.

Scholz-Reiter et al. choose as the key issue the correct degree of use of self-control concepts for logistical processes (Scholz-Reiter et al. 2007, p. 1). The relationship between the degree of self-control of logistical systems and the stated degree of complexity to achieve an optimum target was analysed in an evaluation (Scholz-Reiter et al. 2007, p. 2 f.). The optimum degree of self-control that can contribute to the achievement of the desired targets can be identified with the help of the concept of positive emergence. Emergent characteristics of a system develop through the interaction of system components and not through the characteristics of individual system elements (Scholz-Reiter et al. 2007, p. 4.). These emergent characteristics may be quantifiable characteristic values (average throughput times, stock levels, use of resources, schedule-adherence etc.) or they may represent values which are not immediately quantifiable (flexibility, adaptability, robustness of the system etc). The supplementation "positive" in this case relates to the required characteristics and therefore to the achievement of the declared targets (Scholz-Reiter et al. 2007, p. 5 f.). The investigation of Scholz-Reiter et al. shows that a moderate use of self-control concepts has a demonstrably positive effect on production control.

SFB 614 at the University of Paderborn is involved with the use of self-optimising systems in mechanical engineering. The focus of the research work is on autonomously-acting technical systems which react flexibly to changing environmental conditions. Besides the establishment of self-optimisation within the scientific world, it is necessary to develop a method for the design of self-optimising systems (Gausemeier et al. 2006, p. 50 ff.).

The so called Hanover School has been looking at production control for some years. In this case control of the production system rather than the progress of individual orders is the focal point. Wiendahl is using basic principles of control theory for production planning and control to better engage with the deviation of production logistical processes from production schedules. The use of feedback mechanisms is designed to increase the speed of responses in the event of problems and therefore to reduce the gap between the actual and reference values (Wiendahl 2005, p. 348 ff.). The basic idea of the approach is to balance the planned values generated by the PPC for in-house and external production with actual production values fed back in the form of characteristic values. The analysis of the resulting discrepancies can be used to introduce suitable measures for their reduction with minimum delay. This type of production control enables target specifications to be achieved more quickly and more accurately (Wiendahl 2005, p. 348 ff.).

Nyhuis offers an approach wherein the expansion of production characteristics enables an increase in the degree of synchronisation of production goals. The characteristic curve theory is based on the progress elements which include the working processes and the throughput time of a production order. These throughput elements are the basis for different descriptive models, e.g. the hopper model developed by Wiendahl and the throughput diagram which is based thereon (Nyhuis 2008, p. 191 ff.).

In a general form, the production characteristics are valid for any production system. Specific characteristic curves for an analysed working system are possible however, taking into account additional framework conditions such as capacity, ongoing orders and the integration of the system into the material flow. Therefore a mathematical approach for the straightforward production of logistical characteristic curves for production processes is being developed at the Institut für Fabrikanlagen und Logistik (Nyhuis and Wiendahl 2003; Nyhuis 2008, p. 198). Determining production characteristics provides comprehensive assistance in the positioning of target systems in which relevant relationships between the stock and the throughput time are placed within a graphical context along with the use of resources in a normalised case.

6.3.2.2 Viable and Changeable Design of Production Systems

We now present actual approaches which particularly address the question of how cybernetic principles can be integrated into production systems to increase their changeability. Unlike conventional methods and general procedures in engineering or management sciences, system theory and cybernetics do not exclude complexity by restrictions and simplifying assumptions. It is in fact at the centre of all analyses.

Westkämper and Zahn present the Stuttgart company model, based on a holistic production system, which incorporates methods, tools and approaches for improving changeability in production companies with highly varied series production

(Westkämper and Zahn 2009, p. 25 ff.). The Stuttgart company model is a holistic model, which describes companies as complex systems which in turn comprise semi-autonomous performance units. In order to cope with highly dynamic production, decentralised, semi-autonomous organisms (performance units) form the basis of the Stuttgart company model. The performance unit comprises an organisational unit of a company of one or more employees, who pursue specific goals using resources (Westkämper and Zahn 2009, p. 49 ff.).

The investigations of Wiendahl, which were used within the context of SFB 467, are aimed at developing a situative configuration of order management in a turbulent environment. The term order management is used in order to differentiate it from PPC in terms of planning and control of orders (Wiendahl 2002, p. 18). Turbulence emphasises the fact that the environment is very unpredictable (Wiendahl 2002, p. 14), which leads to requirement fluctuations and unforeseeable events (Wiendahl 2002, p. 17).

Within SFB 467 Balve is developing a framework concept for the design of changeable order management systems (Balve 2002) on the basis of the Viable System Model (VSM). In this case an order management system is defined as an ideal and/or a real tool, supporting the economic and flexible processing of customer orders. The model described is incorporated into an overall procedural concept for changeable order management systems. This comprises the phases of team-building, situation analysis, target formulation, solution synthesis and subsequent evaluation (Balve 2002, p. 95 ff.)

Thiem provides another model for the socio-technical structuring of production systems based on the VSM. In his dissertation the author transfers the concepts of production control to the communication channels and systems of the VSM. Task steps and supporting tools (autonomy profile, information flow diagram) are assigned to stages of the procedural model (analysis of the actual situation, weakness analysis and development of reference structure). The design of the procedural model is followed by its application within a medium sized gearbox producer.

Haats proposed a solution for creating a lean computer-supported system for production planning and control based on VSM. Sub-concepts are developed from the overall concept of a lean PPC model. These sub-concepts are appropriated and concretised to form a catalogue of minimum requirements for informational processes. Haats demonstrates that with increasing decentralisation, while the requirements of communication support and integrated data access increase, the functional complexity of the communication systems overall can be reduced. The implementation and effectiveness of the concept is finally demonstrated in a case study (Haats 2000).

Espejo presents further application reports from various authors in his publication (Espejo and Harnden 1989, p. 103 ff.). The publisher opens the collection himself with his report on the application of the VSM as a diagnostic instrument within P.M. Manufacturers, a medium sized electrical engineering production company. The result of the project is a detailed performance specification. The application illustrates basic techniques for defining recursion levels and implementing the tasks of the VSM mechanisms within the organisational context.

6.3.2.3 Critical Evaluation of the State of the Art and Derivation of the Research Requirement

With regard to the PPC process and its development over time it is clear that many terms and organisational concepts for management cybernetics have also been established within operative processes. Therefore the development of the early deterministic successive planning models can be tracked to strategies and processes of self-control and optimisation in production and logistics. Existing approaches to production planning and control are inadequate for the dynamic processes often prevailing in today's production environments. The planning principle of today's ERP systems which is based on average values and the underlying successive planning prevents an integrated analysis of the PPC tasks and in most cases leads to less-than-optimum planning results. Planning, which is orientated towards an overall commercial target system, is not possible with today's PPC solutions because of their inherently isolated partial analyses and reduced computation complexity.

The approaches for self-control developed within various scientific disciplines show that in the analogy of naturally autonomous behaviour, forms of self-organisation can be transferred to social or logistics systems. The research finds that self-control strategies are suitable under certain circumstances for the management of dynamics and complexity in logistical production systems and exhibit a stable system behaviour. The analysis of the use of self-control strategies in the research we examined is often based on limiting assumptions with respect to the material flow structures and relates to simplified abstract models.

In conclusion, there are so far no models for defining in detail a structure for the decentralised, self-optimised planning and control of production systems and its incorporation into strategic and normative production management systems.

6.3.3 Motivation and Research Question

According to the state of the art, producing companies these days face an increasingly turbulent environment along with the associated increasing process complexity with rigid order processing coordination, which limits their response capability. The lack of reconciliation of elements within order processing due to the strict separation of planning and control means that a rigid and inefficient system behaviour has developed.

The planning and control of customer orders takes place in most companies within the context of operative production planning and control. Today, these tasks are supported by IT systems, so-called Enterprise Resource Planning (ERP) or production planning and control systems (PPC) in most cases.

Particularly planning based on static data (standard restocking times, standard throughput times etc.) is no longer appropriate in commercial practice as it is based on average values. The highly dynamic situation can therefore not be managed with the present planning systematic of many ERP/PPC systems.

High Resolution Supply Chain Management (HRSCM) however describes the creation of a company-wide information transparency, incorporating the interfaces with value chain partners, with the aim of ensuring product availability through decentralised, self optimised control loops in the PPC area. High Resolution Supply Chain Management pursues the aim of enabling organisational structures and processes to self-optimise, i.e. to adjust, according to consistent goals, to constantly changing framework requirements by means of decentralised production control mechanisms in the form of a cascaded control loop model. The potential of the High Resolution Supply Chain Management lies in the possibility of achieving a new level of capacity synchronisation through improved planning and control quality within the context of the PPC.

The designation "High Resolution Supply Chain" therefore relates to the almost unlimited transparency achieved at practically all levels of the industrial supply chains thanks to ubiquitous IT use. Information transparency plays a key role in the planning and control of a producing company. It is the basis for making decisions about the actual need for action and also provides support for actions to be undertaken. Companywide information availability of decision-relevant parameters accordingly serves as a fundamental basis for the operative planning and control of production value chains. The central challenges are the complexity of information and their management along with the effective integration of human intuition and experience into the control system of the High Resolution Supply Chain Management.

On the basis of adequate information transparency, the project pursues the scientific aim of radically improving the planning quality of planning and control processes while simultaneously reducing planning times and costs. This can only succeed when the current prevailing planning logic of the MRP II concept (Hellmich 2002, p. 27 ff.; Spath et al. 2002, p. 130 ff.; Wiendahl 2002) is resolved by a decentralised ACTUAL data-based planning logic.

The aim of the sub project "High Resolution Supply Chain Management" is accordingly limited to the significant increase of the planning quality with simultaneous reduction of the planning times and costs based on adequate information transparency in terms of a "High Resolution Supply Chain". This aim brings us to the following research question:

> How can we structure production planning and control considering the background of dynamic framework conditions to achieve a significant increase in planning quality while reducing planning times and costs?

In the future it will be possible to achieve the subgoal of reducing times and costs by increasing the rate of automation in data collection with the aid of auto ID technologies (e.g. with RFID), by situative aggregation and or disaggregation of information and by the direct availability of planning-relevant data.

The planning quality however is improved by the optimal use of ACTUAL values (information availability) and by the selection of an adequate level of detail (aggregation level) within the context of appropriate planning and control logics. For example, improved forecasting with the associated improved planning reliability can make a substantial contribution to innovative planning and control of a value chain.

6.3.4 Results

6.3.4.1 Principles

Production management is defined according to Kämpf as follows: *"Tasks, people, machinery and materials should be employed, controlled and co-ordinated such that products and services as a result of this activity are produced in the required quantity and quality at the specified time with the minimum outlay of costs and capital."* (Kämpf 2007, p. 5 ff.). In the following, production management should be understood as the planned and controlled employment of production factors in decentralised production systems (production network), in order to provide material products or services according to customer requirements with respect to quantity and quality at the required time with minimum costs and time.

At first, conventional production management is described in this section. It is followed by an explanation of the principles of cybernetic production management from which the preconditions for High Resolution Supply Chain Management (HRSCM) are derived.

Conventional Strategic and Normative Production Management The St. Galler Management Concept based on Ulrich's system approach differentiates between three levels of management: normative, strategic and operative. While normative and strategic management basically relate to design and development functions, operative management focuses on the steering function (Bleicher 2004, p. 80 ff.).

Normative management of a production company does not differ substantially from the normative management of any other company. The general company aims, principles, standards and company culture are defined and are intended to ensure the viability and development potential of the company. The overall company aim of a production company is typically to succeed in securing its existence.

The general aims of normative production management can be:

- the balanced fulfilment of the requirements of the various interested parties (partners e.g. neighbours, customers, shareholders and employees)
- obtaining or consolidating a significant position within
  - an industry,
  - a technology (e.g. laser technology) or
  - with respect to a material (e.g. special glass)
- concentration on the processes with the greatest added value irrespective of the traditional strengths of the company or of the industry in order to maintain the company sites and the company's size

Building on the aims of normative production management, we now address strategic production management.

Strategic production management in this case has two outstanding aims (Eversheim and Schuh 1996):

- Dynamic, market orientated development of the goods and services of the company
- Creation of long-term competitive advantages by establishing core competencies

To achieve long-term competitive advantages, core competencies, from which surprising products result, must be developed faster and more cost effective than those of competitors. The real sources of a strategic benefit therefore lie in the ability of the management to combine technologies and production capabilities (companywide) to form competences. Only then the business units will be strong enough to be able to respond rapidly to emerging opportunities in the future (Westkämper and Zahn 2009).

Conventional Operative and Tactical Production Management A comprehensive analysis of the task areas of operative and tactical production management was undertaken within the context of the Aachen PPC model (see Schuh 2006). The Aachen PPC model extended by Schuh serves firstly to describe various aspects of production planning and control (PPC) and secondly offers support in the definition of PPC aims and the application of various design and optimising methodologies. The main task is to divide the PPC into several smaller models, each taking a different perspective. This breakdown is required to identify the different aspect groups (personnel aspects, information technology aspects and commercial aspects) which have different influences on the aims and model requirements of the PPC. The four perspectives used in the Aachen PPC model are:

- the task perspective
- the process architecture perspective
- the process perspective and
- the function perspective.

The task perspective describes the tasks of the PPC on a generally applicable hierarchical level of abstraction. The process architecture perspective forms the interface between the task and process perspectives by creating connections between tasks at the network level and the company level. The process perspective identifies processes of the tasks in the task perspective in a chronological sequence in order to describe the order processing more precisely. Finally the function perspective describes the requirements for IT systems, which support the internal PPC (ERP-/PPC systems). The primary tasks of the task model can be broken down on the basis of the following structure (Fig. 6.26).

This structure divides tasks into industry-wide network tasks and business-wide core and cross-section tasks. The core tasks of production management in this case comprise production program planning, production requirements planning, procurement planning and scheduling and in-plant production planning and scheduling. Network tasks add industry-wide aspects at a strategic level to the task interpretation of the original PPC model (according to Luczak and Eversheim 1999) and include network configuration, network sales planning and network requirements planning. The cross-section tasks integrate the network tasks and the core tasks. These include order management, inventory management and financial controlling. All task areas

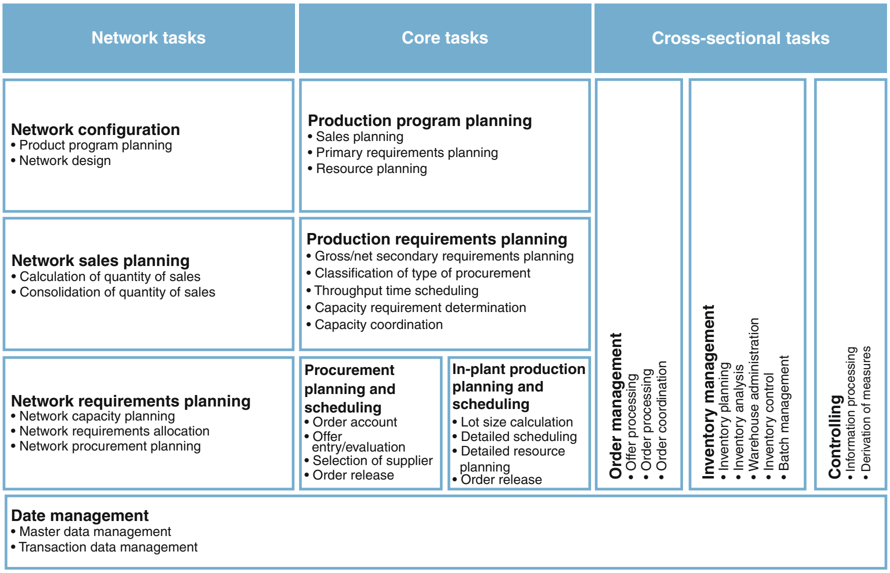


Fig. 6.26 Production management tasks. (From Schuh and Roesgen 2006, p. 28)

require access to the data management, which is therefore attributed to every task (Schuh 2006, p. 20 ff.)

Structure and Tasks of a Cybernetic Management System The basic ideas of cybernetics originated in the 1940s and 1950s. New methods were sought to solve complex problems. Central to cybernetics are the regarding of information as a central value and the use of closed systems and control loops. Already in the early phases of cybernetics, an analogy with living systems and natural organisms was sought. Stafford Beer adopted the principles of cybernetics into his Viable System Model (VSM), which serves as the reference model for describing, diagnosing and designing the management of organisations. The basis of the Viable System Model is the analogy with the human nervous system. Beer divides the organisational model into five subsystems, which can be assigned to three structural levels. The operative level of the whole system contains systems 1–3 and relates to the present and to the inner world of the system. At this system level, actions are initiated autonomously through routine behaviour and reflex-type adjustments to changes in the environment. System 4 aims at stabilising the system within the external world. System 5 represents the third structural level and is the normative instance. The following illustration (Fig. 6.27) is based on the cybernetic derivation of the model according to Beer, Malik, Gomez and Espejo (Beer 1979, p. 319; Malik 2006, p. 84; Gomez 1978, p. 24; Espejo and Harnden 1989, p. 99).

System 1 contains the management units (squares) and basic units (circles) in analogy with organs and muscles which implement the company processes. The steering units optimise the daily activities of the largely autonomous process units.

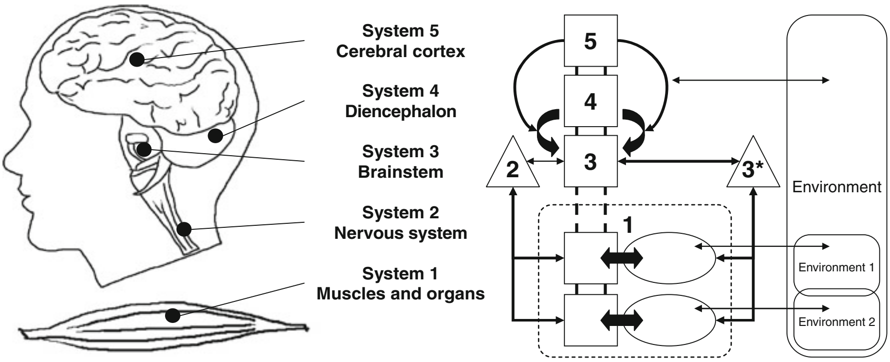


Fig. 6.27 The structure of the VSM. (From Beer 1979, p. 319; Malik 2006, p. 84; Gomez 1978, p. 24; Espejo and Harnden 1989, p. 99)

The operative units act autonomously within fixed limits at the horizontal level. The link between the individual systems 1 and the metasystem should be understood as a list and not as an hierarchical order (Beer 1979, p. 121).

System 2 coordinates and regulates the semi-autonomous systems in analogy with the nervous system. The core task of the system is to reinforce the self-regulating capacities through the supply of information and damping of oscillations between the operative units of system 1.

Like a brain stem, system 3 has an overall model which is superior to all systems 1 and their interactions. The operative overall management system ensures optimisation of the whole system. By its direct connections with all subsystems it detects simultaneously and in real time everything which occurs in system 1. It is also notified about activities in system 2. System 3 passes on instructions directly to system 1 via the central command axis. The monitoring and validation of information from the operative units are functions of system 3*.

System 4 represents the strategic system level in analogy with the brainstem, i.e. the external and future orientation. It detects and diagnoses its environment on a system-wide basis and constantly changes the orientation of the overall system on the basis of this information.

The normative level of system 5 (the "cerebral cortex") defines the identity of the system and specifies standards, values and strategies in order to maintain and develop that identity. It maintains a balance between the present and future perspectives.

Implementation of Cybernetic Principles in Production Management The production management system manages the production system by specifying control variables and identifying any intervention by actual comparisons with the control variables which are fed back. This control and monitoring can be described as a control loop (Zäpfel 2001, p. 3 f.). This interpretation is further supported in the control loop concept model in which the production system comprises the management system for planning and control (production management) and the performance

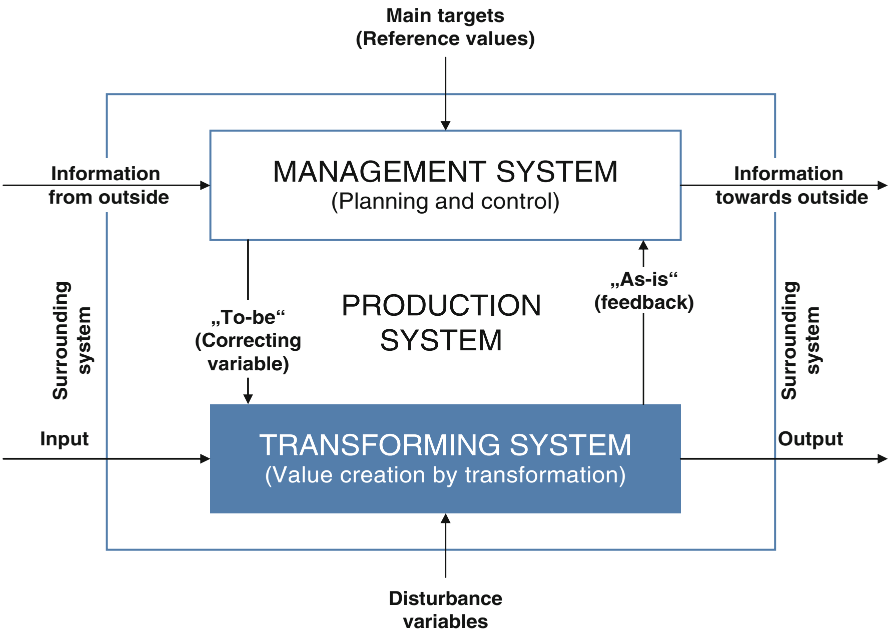


Fig. 6.28 Production management control loop. (From Dyckhoff and Spengler 2007)

system for adding value (see Fig. 6.28) The relationships between the management and transforming system (production system) are illustrated below.

In terms of the control loop, the management system transmits so-called reference values (control variables). These reference values define the quantitative and qualitative specifications for the performance system which in turn affect the economic and other target variables. External disturbance variables such as supply shortages, employee sickness etc. cause the actual results (control variables) of the production to deviate from the specified reference variables. This deviation is analysed with the help of further external information such as the development of procurement and sales markets and taking into account the global aims of the company (e.g. achieving maximum marginal income or profit). The results are transferred to the future reference variables. The central task of the production management system is therefore target-orientated planning and control of production taking into account organisation, human resources, information supply and monitoring (Dyckhoff 2003, p. 6 f., 29).

Approach and Preconditions of High Resolution Supply Chain Management Thanks to technological innovations, real-time operative information is now available at a new level of granularity. Particularly new technologies of information collection such as Radio Frequency Identification (RFID) und and informational integration of companies have paved the way for this revolution (Fleisch 2008). This information is used for production management in HRSCM. The approach combines the structuring of real-time data-based planning and control with an adequate decentralised organisational structure of production management.

Due to the interdisciplinary nature of the approach the preconditions of HRSCM are technological and organisational. The technological preconditions include the availability and usability of actual values for planning (reference value definition). These are the basic preconditions for enabling the transition from control to self-regulation within production management. Only realistic and achievable reference values permit the adjustment of regulated process variables. Therefore the generation of sufficiently detailed information, e.g. by sensors, and their availability through IT systems is a technological precondition for HRSCM.

From an organisational perspective, the complexity of companies as organisational units and their dynamic environment are the fundamental challenges. The organisational structure of the approach must overcome the complexity and be able to maintain internal stability as well as that of the external environment. As a basis for this, the structure must be able to satisfy the information requirements of the individual units which it comprises.

6.3.4.2 Structural Model of High Resolution Supply Chain Management

The aforementioned requirements determine the general requirements of the structure of a High Resolution Supply Chain Management (HRSCM) reference model. The organisation must be able to cope with the complexity of the company and the complexity of its dynamic environment. This can only be achieved by means of structural decentralisation and a high degree of response because decentralised interactions reduce the amount of coordination required (Frank et al. 2004).

Based on the Aachen PPC model and the VSM it is necessary (with regard to modularity) to define different reference perspectives to reflect the different perspectives of the overall model. The Aachen PPC model unifies four perspectives—the task-, process-, function- and process-architecture- perspectives. The function-perspective does not need to be considered as an actual perspective within the context of the model because it is aimed at the specific assignment of IT systems to implement the tasks. The process and task perspectives are adopted directly and extended for the newly established tasks and (management) processes within the metasystem.

Strategic and Normative Production Management System The mechanisms described next ultimately lead to the definition of strategic specifications for the operative production management system via the interface of the configuration channel. The core task of strategic and normative production management is to develop the business model and order processing according to changes in the market while preserving the actual company identity and values and to initiate targeted changes. The spectrum of these changes extends from organisational and structural changes to technical adjustments and innovations.

The external equilibrium pursues the ideal of effective orientation of the order processing with respect to the environment and environmental requirements. The

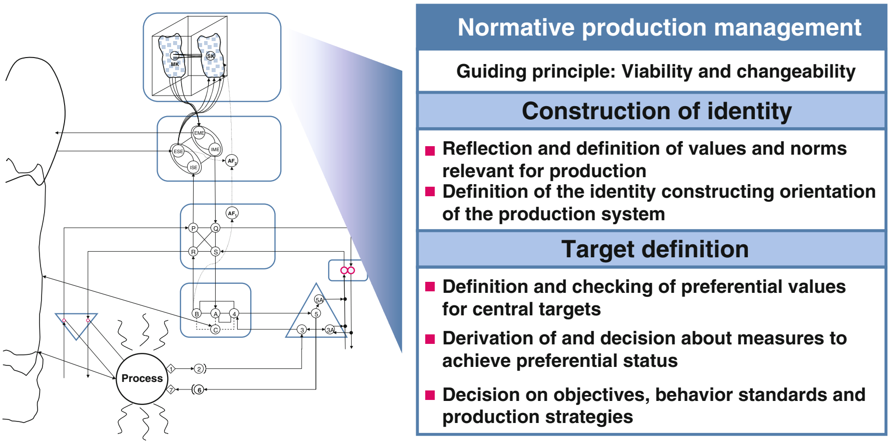


Fig. 6.29 Organisation and overview of the tasks of normative production management

equilibrium of the overall system ultimately involves balancing internal and external stability statuses. This overall equilibrium is monitored by the normative production management system which is responsible for the change, development and viability of order processing.

In the following sections we present the tasks, processes and information flows initially for normative production management. We then describe the tasks, processes and information flows of the strategic production management system.

The normative production management represents the highest level of the production system. Its aim is to ensure that the overall system remains viable in the long term by orientating the system both normatively and in terms of continuous change. As well as the aim of commercial viability, the target system thus engendered is also supplemented with normative ideals and values and requirements for its future scope of activity. The values and standards ensure the orientation of the target system within the identity of the company. Therefore normative production management is responsible, along with identity and values, for defining reference statuses and the initiation of measures to achieve these statuses.

The tasks of normative production management for securing the viability and flexibility of the production system are illustrated in Fig. 6.29.

The information flows for fulfilment of the tasks described above can be divided into the sensory (detecting) and motor function (instructing) information flows. The effectiveness of the system design depends on the interaction of the motor function and sensory mechanisms.

Sensory Components The sensors obtain all information about the internal and external stability statuses from the strategic production management system. The sensors also obtain all algedonic information which has been prepared by the alarm filters. If the actual status in the results varies from the desired status of the overall system,

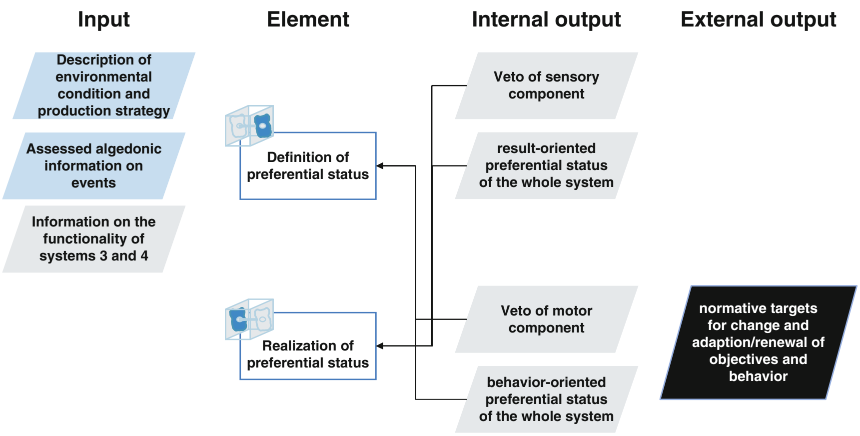


Fig. 6.30 Information flows of normative production management

measures are initiated. The sensors may in this case veto measures initiated by the motor functions.

Motor Function Components The normative instruction mechanism now checks the plausibility of the specifications from the sensors. Decisions are also based on preference statuses which can be assigned to them by the sensor system. The motor functions have corresponding knowledge of which results can be achieved with which responses. The statuses contain actual instructions which imply specific responses at the operative and/or strategic level.

Figure 6.30 illustrates the basic information flows of normative production management.

The task of the strategic production management centre is to anticipate potential futures, to initiate the adjustment of the organisation to the dynamic environment and thereby to adjust the strategic orientation of the production system so as to preserve existing success potentials and create further potentials (Gomez 1978, p. 24; Eversheim and Schuh 1996; p. 5 ff., Rüegg-Stürm 2003, p. 71; Malik 2006, p. 141).

Hill's concept is designed for the configuration of a differentiated production strategy which is particularly involved in the linking of the company and marketing strategy with the production strategy. Coupling is carried out by considering the qualification and differentiation features of the products (Hill 2000, p. 32 ff.). Qualification features make the company fit to compete for a market segment. Thus they enable the company to participate in the market. If qualification features are lost, the customer is no longer included in the decision to purchase the product. Differentiation features however have an impact on the purchasing decision of the customer when faced with comparable products. They differentiate the performances of the company in terms of satisfying a customer requirement from those of the competition (Hill 2000, p. 38).

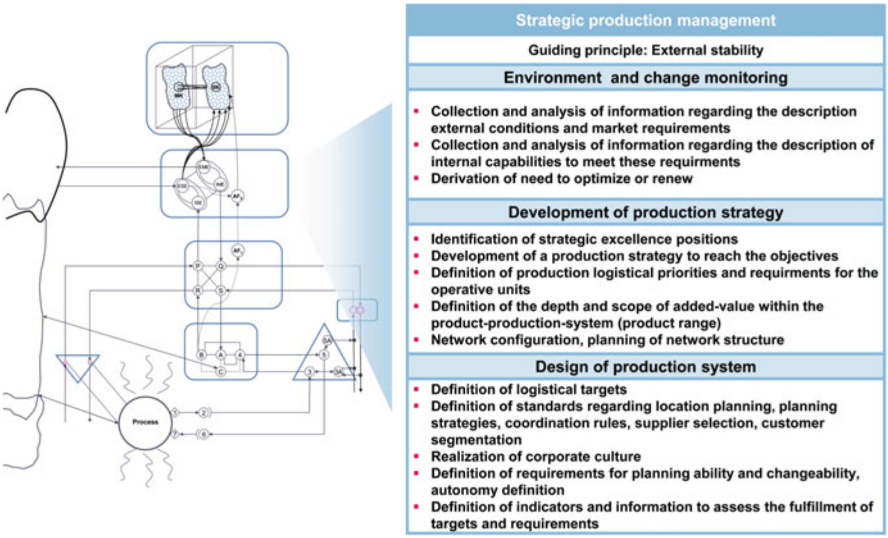


Fig. 6.31 Organisation and overview of the tasks of the strategic production management system

The tasks of the strategic production management system for ensuring the external equilibrium and effectiveness of the production system are illustrated in Fig. 6.31.

Like normative production management, the information flows also travel from the strategic level through a sensor and a motor function component. The essential difference in the structure as compared to the normative level is the presence of interfaces at higher levels of the production management.

Sensory Mechanism The central task of the sensory mechanism is to record and process information. In this case the stability detection mechanism (the internal reporting system) obtains information about the actual status of the equilibrium within the production system. This is information about the actual logistical parameters of use of resources, inventory, schedule-adherence and throughput time but also about the progress of structural and organisational changes.

Motor Function Mechanism The motor function mechanism is the executive organ of strategic production management. The mechanism obtains reference specifications about the external and resultant internal stability statuses in the form of the production strategy. In addition, normative production management is able to change strategic specifications at any time and define new goals for the system.

Alarm Filter 2 Alarm filter 2 passes on exception information to higher control levels. This is firstly the information from alarm filter 1 of the operative and tactical production management systems and additional information from the environment or from the programmes which implement the production strategy. Alarm filter 2 balances the information via internal disturbances with the market analyses of the strategic

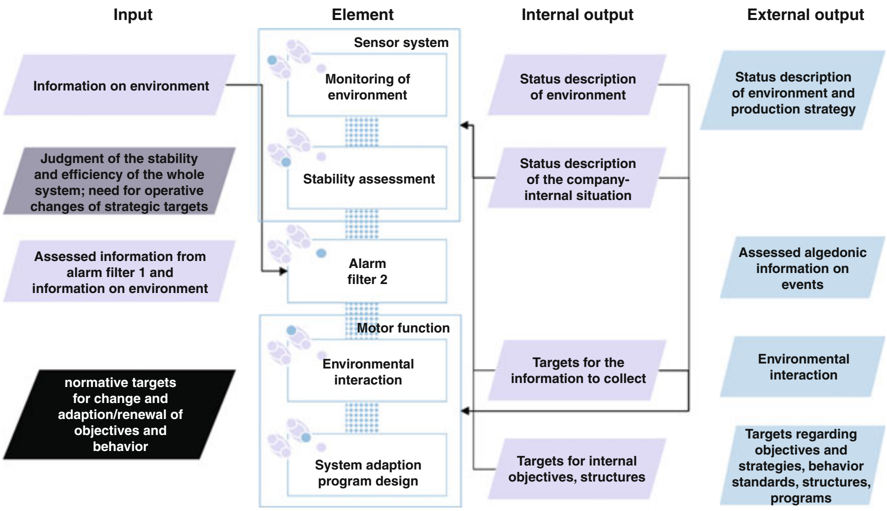


Fig. 6.32 Information flows of the strategic production management system

sensors. The interaction of the disturbed internal stability statuses and the market performance is the output of alarm filter 2. Figure 6.32 illustrates the information flows of the strategic production management system.

Operative and Tactical Production Management System In this section, in the same way as the previous section, we present the tasks, processes and information flows for operative and tactical production management.

Operative production management serves to steer the operative units with the aim of stabilising and optimising the system in the “here in now”. This is carried out by designing the order processing in systems 1, 2 and 3 of the first recursion level (main order-processing processes, process coordination centre, tactical production management and by designing systems 1–5 of the second recursion level (main- and subprocesses, process control centre, process management).

The operative configuration and optimisation of the operative processes is the task of tactical production management (system 3 in VSM-terms) In this case, the primary issues are medium-term configuration according to strategic specifications from the strategic production management system and short-term optimisation of the processes without structural intervention. According to the basic assumption that the processes of order processing, when stable, proceed without intervention from higher management levels, unlike the previous research, order processing management tasks are largely decentralised. The tactical production management controllers can focus on stabilising interventions in the event of major unrest within the system and on continual improvement of the interplay of the subprocesses. To this end, the tactical production management system, unlike the process coordination centre, can intervene via the central control channel directly into the autonomy of the processes. The autonomy of the operative processes is thereby a gradual function and is limited or expanded according to the situation. Indirectly however, tactical production management can also influence the internal procedures via the process coordination centre (coordination channel) or via process monitoring (monitoring channel).

Configuration of the Production System to Ensure External Equilibrium A precondition for the short and medium-term configuration of operative processes is the understanding of external requirements of the logistical performance of the production system. These are identified in the strategic production management system and are input into tactical production management by a feed-forward mechanism. The dimensions of the requirements definition are primarily the reference delivery time (order throughput time), reference schedule-adherence and price level. The task of production management at this point is to check these external customer requirements for feasibility, translate them into internal process requirements and to simultaneously guarantee maximum efficiency of the overall process. This takes place by defining optimum operating points in the overall system and individual main processes. The target conflicts of the process planning in this case are minimised as far as possible by information transparency and organisational networking.

Operating characteristics according to Nyhuis and Wiendahl are a suitable means of choosing operating points. These have been established to illustrate the effective relationships between logistical target variables within production logistics, see for example (Gläßner 1995; Yu 2001; Lutz 2002; Nyhuis and Wiendahl 2003; Wiendahl 2005; Lödding 2008). Production characteristics are a condensed illustration of different operating statuses. Operating statuses are determined by analytical or stochastic monitoring modelling of real systems and by computation by means of simulation (Nyhuis and Wiendahl 2003, p. 36 f.). Figure 6.33 illustrates the target conflicts in production control on the basis of a characteristic curve diagram.

Based on the characteristic profile of the characteristic curves, the task of tactical production management is to position the reference operating points within the target conflicts of the PPC according to the strategic requirements. As a result, the reference variables of schedule-adherence, schedule deviation, throughput times, material and/or order levels, use of resources, quality of the process results and process costs per main process are defined and communicated as specifications.

If the desired reference values of the operating points cannot be achieved within an existing configuration, a series of tactical decisions must be made to determine the framework conditions for the operative processes and to configure the processes accordingly. In this way, the profile of the characteristics curves changes and enables new combinations of operating points. Decisions about the dimensioning of the use of resources are negotiated via process channels.

Depending on the targets in question and the resources available, the process channels then automatically define the effects of these targets in the form of control variables for the assigned subprocesses. This heredity ensures that the process

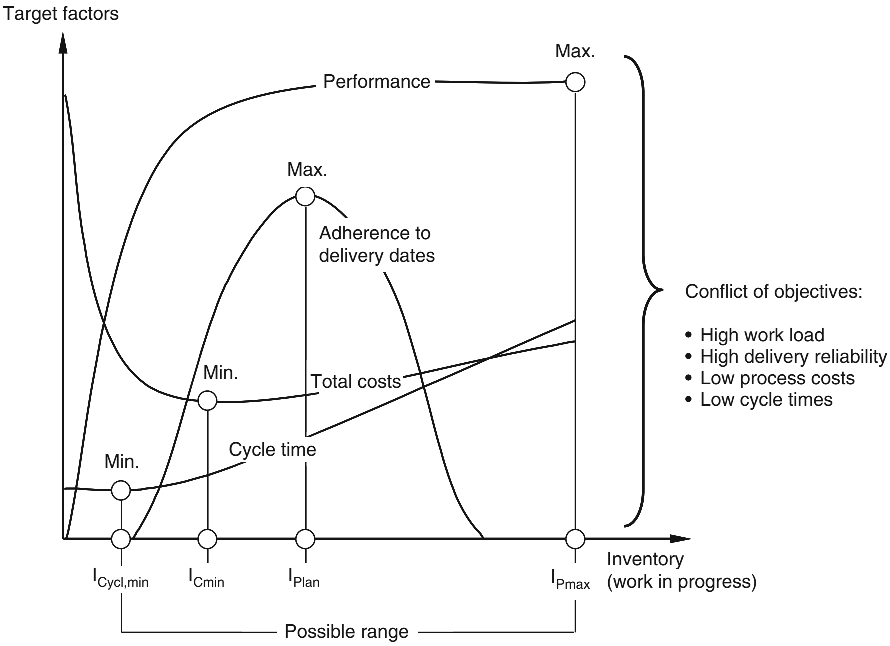


Fig. 6.33 Target conflicts in production control. (Wiendahl 2005, p. 35)

channels design the target corridors of the subprocesses in the recursion RL3 level such that the targets of the metasystem can be maintained at levels RL1 and RL2.

Monitoring and Optimising Horizontal Coordination to Ensure Internal Equilibrium As well as configuring the operative processes to ensure equilibrium with external requirements, the second task of tactical production management lies in monitoring and achieving internal equilibrium. Internal equilibrium is considered as having been achieved when the productive system fulfils the target values of schedule-adherence, throughput time, costs and synergy suggestions. Reference values are also defined based on the definition of escalation levels which describe alternative processes for handling discrepancies inside and outside the target corridors. Threshold values underpin these processes which determine when a particular disturbance should be balanced either locally by the process control centre or by the process coordination centre. If neither damping mechanism is able to find a solution for a problem, it is escalated to the tactical or even to the strategic production management system. The frequency and intensity of the discrepancy from reference values is the measure of the stability and therefore the internal equilibrium of the system or of subsystems.

As a reaction to these discrepancies, medium or short-term and one-off or recurring countermeasures can be instigated. One-off or sporadic measures are used to handle individual exceptions such as for instance major interventions in an individual order (machine use, order sequences, quality or cost compromises), the involvement of external advisers or downsizing or upsizing. This limitation of the autonomy of the decentralised units is implemented via the control channel and should be avoided if at all possible. Recurring countermeasures are incorporated as standards which so to speak pre-equip units with instructions, rules or response instructions. These are reflected amongst other things in changed reference values, changed rules for use of resources, changed procedures, changed decision-making mechanisms or physically changed resources, increased or decreased capacities or changes to responsibilities. Both types of intervention are aimed at ensuring and optimising the fundamental synergy pre-settings in the interaction of the main processes.

Process Monitoring for Ensuring Internal Equilibrium In the case of these processes, the tactical production management system is supplied with information from the process coordination centre and with unfiltered information from the processes. If deviations from reference values are symptoms of a lack of synchronisation, process monitoring is responsible for identifying the causes, in terms of which the actual situation is being interpreted (Herold 1991, p. 135 f.; Thiem 1998, p. 110; Malik 2006, p. 136 f.). Direct interfaces with the areas (IT systems or personal reporting on the part of group representatives) are used to request absolute information about the situation of orders and capacities (type, quantity, schedule, quality, progress etc) and suitable characteristic values for identification of deviations from the stable status of the units are disclosed. These overload manifestations are accompanied by changes in the performance values of use of resources, inventory, schedule-adherence, throughput time, costs and quality and can be expressed in the form of

- exhausted overtime accounts,
- high levels of sickness,
- continuously high order queues,
- low schedule-adherence in the case of subprocesses,
- urgent orders or “fire-fighting” and
- temporary capacity shortages.

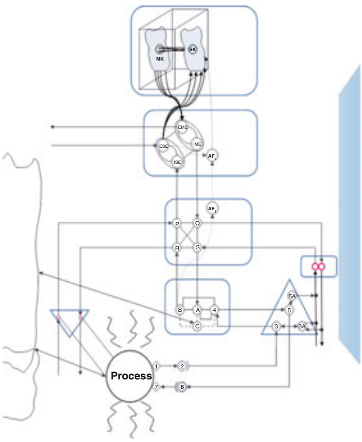
Against this background the information described above is collected by the tactical production management system and used to improve the decision-making basis for all orders. In the case of a disturbance, corrective measures are generated to resolve the permanent or temporary shortages or short term responses to “unforeseen developments”. These measures directly affect the procedures and are controlled by the monitoring channel (Herold 1991, p. 144). Examples of these kinds of immediate measures are the provision of additional capacities by approving overtime, involving external suppliers or hourly-paid labour as well as overriding locally-planned order sequences, batch sizes and resource planning.

Adequate Structuring of Planning Capacity by Implementation of High Resolution The success of the business processes as well as the coordination of the individual processes in the operative control is largely influenced by the planning capacity. This in turn depends on the availability of information and the quality of information flows (Wiendahl 2006, p. 16). This is particularly taken into account in the mathematical description with the time loss factor δi. Thus a core aspect in the tasks of the tactical production management system is to align the time delay for the transfer of information between the decision-making points with requirements. The appropriate interaction of the node points changes the decision-making quality of the controller. Poor availability, susceptibility to error, incompleteness, incomprehensibility and irrelevance of information reduce the decision-making quality of a node point and correspondingly the decision-making quality of the whole management mechanism. The task of the tactical production management system therefore is to define the requirements of granularity, real-time capacity and response time behaviour of the individual control loops and their associated management, control and disturbance variables. In principle, the resolution of information should be as high as necessary but not as high as possible.

The order processing target system must be expanded accordingly by optimising the planning capability. Planning capability (or planning reliability) is generated by a high level of information transparency and equally high planning data quality (Schmidt 2008, p. 63). The quality features of planning information thereby include completeness, currency, level of detail and absence of error of the informational content (Loeffelholz 1991, p. 27 f.; Schotten 1998, p. 44). As with the characteristic curves this target is inherited by the process channels and the coordination system in the form of reference operating points. The tactical production management system must test and optimise the fulfilment of the information quality criteria via the monitoring channel and the process coordination centre.

Adequate Flexibility of the Operative Systems The variability of individual process steps and of the overall production system has become a central aspect of production management (Blecker and Kaluza 2004, p. 4 ff.; Zäh 2005; Drabow 2006; ElMaraghy 2005, p. 264; Wiendahl and Hernàndez 2006, p. 147 f.). The aim of these endeavours is to create adequate product flexibility (to maximise variant variety with a specified infrastructure, process flexibility (to manufacture products using alternative processes, machinery, materials or process sequences) and capacity flexibility (scalability of the quantity of manufactured products) (Chryssolouris 2005). At the operative level, the focus is not the restructuring of the overall order processing or of the product mix because this is now a task for the strategic production management system. The aim here is the appropriate design of convert-ability, reconfigure-ability, flexibility and the versatility of the physical production system along with the logical processes and management mechanisms (Wiendahl et al. 2007, p. 787).

The tactical production management system now has the task of translating the strategic specifications for flexibility in the three main dimensions of product flexibility, process flexibility and capacity flexibility into operative specifications for the decentralised processes and to monitor their compliance. Specifications with respect to convert-ability, reconfigure-ability, flexibility and versatility are thus transferred to the design elements in the main processes where they are implemented. The essential enablers of the implementation are reconfigurable process planning, adaptive



Tactical production management

Guiding principle: internal stability and efficiency

Process configuration

- Assessment and reconciliation concerning feasibility of strategic guidelines
- Adaption of topical structures based on strategic guidelines
- Definition of synergetic operation points for the total system
- Deduction of process-independent objectives, priorities and standards
- Definition of requirements of changing operative processes

Process direction and -control

- Forwarding and realization of strategic guidelines
- Definition and control of target operation points, resource availability and autonomy of all main processes
- Assessment and optimization of methods, operations and tools that should be used
- Definition of the order penetration point and planning mechanism
- Assessment and optimization of changeability and feasibility planning
- Support of operative processes on solving critical disturbances
- Control of middle-term supply chain design, supply chain demand and production program planning

Process coordination

- Definition of norms, priority rules and coordination routines
- Definition of escalation levels for the cooperation of coordinating units
- Assessment of process coordination center effectivity

Process monitoring

- Assessment of internal stability and efficiency of operative processes
- Sporadic interventions to save internal stability

Fig. 6.34 Organisation and overview of the tasks of the tactical production management system

production planning and control, reconfigurable production and assembly systems and flexible fabrication (Wiendahl et al. 2007, p. 789 ff.; ElMaraghy et al. 2009, p. 261 ff.). As with the characteristic curves this target is inherited by the process channels and the coordination system in the form of reference operating points. The operative production management system must test and optimise the fulfilment of flexibility criteria via the monitoring channel and the process coordination centre.

The tactical production management tasks of ensuring internal equilibrium and the efficiency of the production systems are illustrated in Fig. 6.34 and incorporated into the structure of the overall model.

Information about the external stability statuses reach the configuration channel via the connections between strategic and tactical production management. The configuration channel interprets the requirements of strategic production management and information about the current internal stability from the other channels. The balance of external and internal equilibrium results in a definition of reference operating points, target corridors in the characteristic curve field, characteristic values and escalation levels, standard behaviours and basic resource availability levels. These changes of the characteristic curve lead to the specification of new operating points in the main processes and therefore to the expansion of the internal stability via the control channel.

The configuration channel completes the information loop by providing information about the actual status of processes to the strategic production management system.

The Control Channel: Information Loop and Process Channels The tactical production management system intervenes into main processes directly via the central command axis of the control channel. The framework conditions for the reference operating points of the overall system from the configuration channel and information about the internal stability channel are translated via this channel into concrete instructions which are conveyed to the process channels of the main processes. These include particularly basic response instructions about core processes such as working times, maximum use of resources, capacity flexibilities or work organisation regulations and the definition of process related aims, e.g. the optimisation of process costs versus minimisation of throughput times. Thus the negotiation process over the allocation of resources takes place via this channel. The target system specified by the strategic production management system is therefore inherited by the next recursion level and appropriated per main process and/or negotiated with the process channels.

The Coordination Channel: Information Loop Incorporating Process Coordination Centre and Process Control Centre The basis of internal stabilisation is the reconciliation of the main- and subprocesses at synergetic operating points. The tactical production management system specifies the management values and target corridors within whose parameters the actual processes must adjust. The reference characteristic curves with the target corridors and operating points. The KPIs of the escalation model are transferred by the coordination channel to the process coordination centre. When selecting and defining the characteristic values, the mechanism takes particular account of the specifications for external stability on the part of the strategic production management system, feedback from the processes via the monitoring channel and the input via the profile of the actual operating points. The process coordination centre is a non-hierarchical instance. It controls the process operating points and the order status. In the event of a discrepancy it initiates compensating measures between the processes within the context of a predefined escalation model. The reporting (feedback) of the information is also carried out according to the specifications of the tactical production management system. In this case, only filtered information about discrepancies from plans, operating points or planning milestones is transferred to the tactical production management system. Thus no detailed image of the actual situation within the processes is permitted yet but it prevents information overload at this point.

The Monitoring Channel: Information Loop Incorporating Operative Main Processes The tactical production management system obtains only relative information about deviations from the reference state via the coordination and monitoring channels. However, the monitoring channel is designed to reflect the actual status of processes. It firstly identifies overload manifestations and new types of developments. Secondly, this channel provides the tactical production management system as required with detailed information about individual orders, load curves or quality data in order to permit corrective intervention in the case of shortages or problems. In addition to direct interventions, the knowledge obtained from the monitoring system is also used at other decision-making points within the tactical production management system. In order to identify new types of developments promptly, information about possible overloads is transmitted continuously whereas the specific order or capacity data is requested and fed back only as required.

The Alarm Filter: Direct Flow of Information Between All Levels So that neither the operative nor the tactical production management systems are flooded with information, information from process management and from the process coordination centre is filtered according to pre-settings. The set up of an integrated, direct channel is therefore required for the exception information flow (Malik 2006, p. 138 f.; Gomez 1978, p. 68 f.). The information flow here follows the principle of "Management by Exception" and is defined individually for each company and/or for each production system. It is transferred directly to the strategic and normative production management centres and is defined as sensitively as is required. Possible content might be serious breaches of schedules in the case of important customer orders, new major customers, specific quality problems, failure of a line or computer system etc.

Figure 6.35 summarises the information flows of the tactical production management system. The dotted line between the 5 channels symbolises the prevailing information transparency within the tactical production management system.

The tactical production management system, as the central switching point for operative processes and the interface with the strategic production management system, should incorporate representatives of all assigned processes for the exercise of its functions. This is particularly necessary to implement a synchronisation of the target systems. The representatives should participate regularly in the operative processes in order to be able to provide information to the monitoring channel independently from the system as well. The tactical production management system is supported in its activities by a control room which prepares transparent versions of the basic information from all four channels. Acting as a logistical control centre, it transports and visualises the characteristic values and status information of the monitoring, control, coordination and configuration channels together in appropriate time cycles.

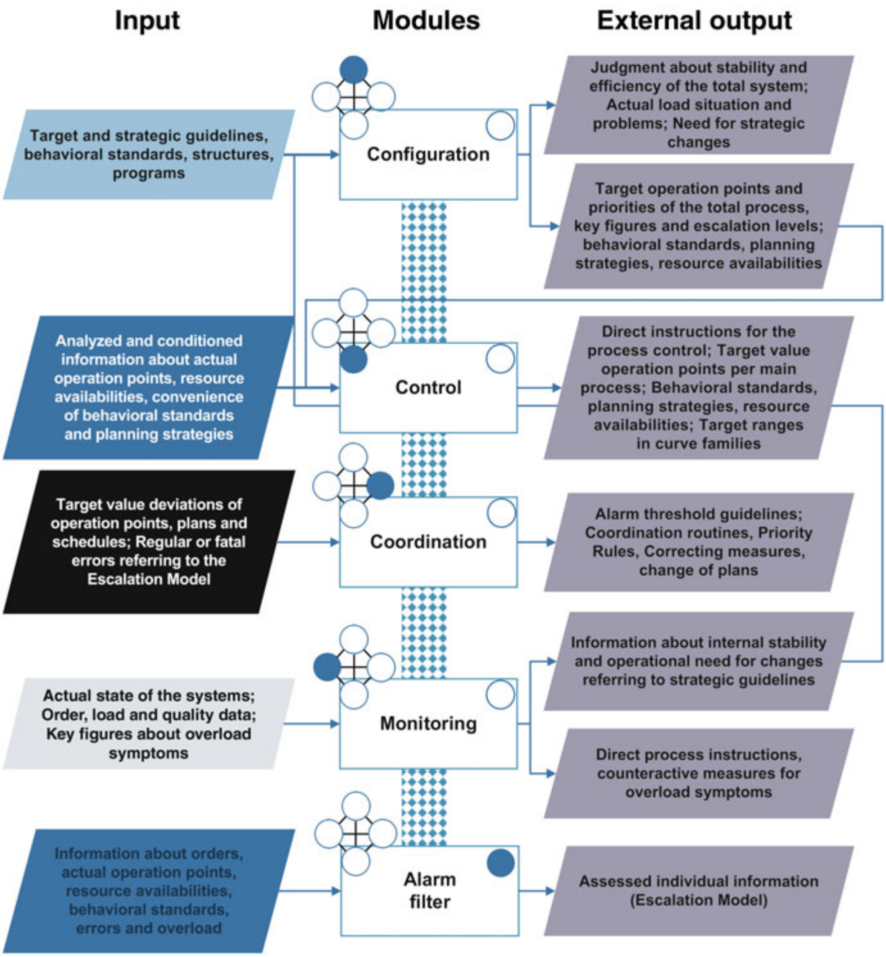


Fig. 6.35 Tactical production management information flows

While the VSM provides a conceptual model and the structural framework for the production management design, the model offers no methods for simulation, analysis and evaluation of the dynamic systems at the process and process control level. At the operative level, the management system can be supplemented by the transfer of specific control engineering components. Since both VSM and control engineering originated within the trans-disciplinary approach of system theory, these can be consistently integrated. However the description of the process control centre and process management defines abstract requirements for the control of the core processes of production management which must now be specified in more detail in terms of control engineering.

The process management systems organise the operative implementation of tasks and processes of technical order processing. The tasks are fulfilled according to the

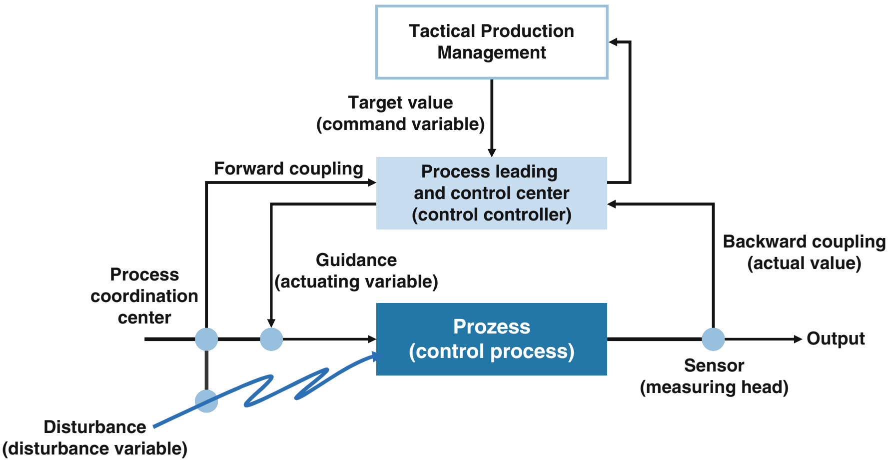


Fig. 6.36 Interaction of operative management units

requirements of viable systems by semi-autonomous value-adding process islands or performance-orientated capacity groups.

Up to four management instances along with the implementing resources are involved in the exercise of process management and control. The process management system of each process obtains reference values for the process management variables from the tactical production management system in the form of a servo mechanism. The process management system obtains information via the process coordination centre about possible and expected problems and the actual status of other systems and/or orders (feed-forward). The process management system receives the status of assigned processes via the process control manager (feedback). The reference variables from the tactical production management system and from the process coordination centre and the status information of the assigned process from the process control centre are translated into control variables (activities, programmes and schedules) for the management of the process. The implementation of these plans is once more initiated by the process control centre. Figure 6.36 illustrates the servo mechanism of the tactical and operative production management systems.

According to the basic logic of the maximum autonomy of the decentralised units, the process management system intervenes only if there is a discrepancy between the process and the reference value which cannot be counteracted by a standard routine. The known repertoire of routines for handling dynamic situations is saved in the process control centre and/or in the process coordination centre. These include particularly project monitoring, control tasks and countermeasures for the regulation of recurring problems. In accordance with the recursive structure of the overall model, the process management system and the process control centre jointly implement the tasks of systems 2–5 for the observed process. Figure 6.37 summarises the process management tasks.

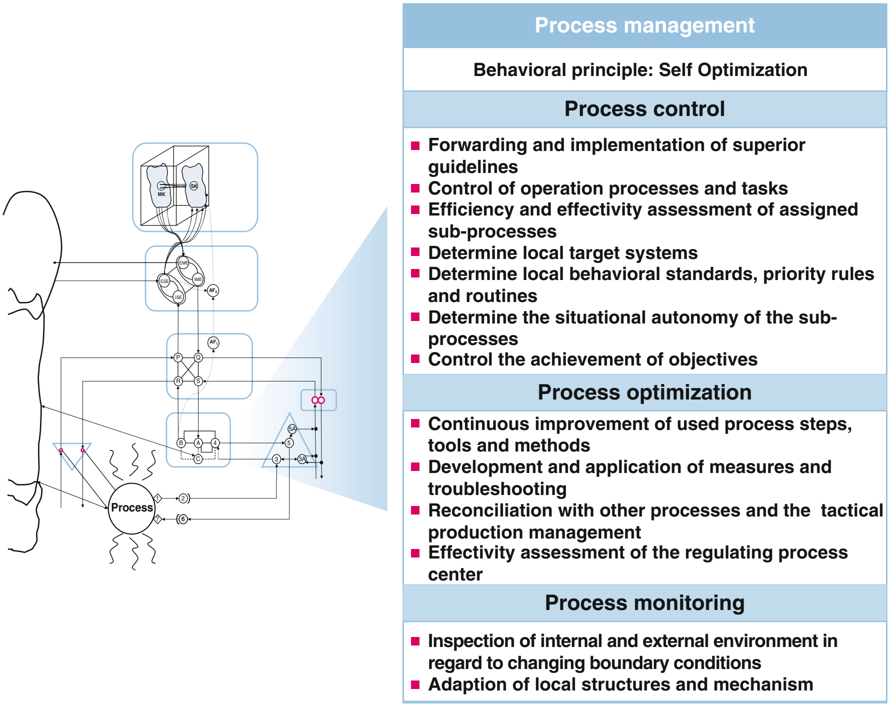


Fig. 6.37 Organisation and overview of process management tasks

The process management systems are supported by the process control centres, which carry out the tasks illustrated in Fig. 6.39.

The process coordination centre undertakes the process control centre tasks at the next recursion level up. It thereby regulates the interplay of the main processes and ensures implementation of the programmes selected by the tactical production management system. The process coordination centre does not play a hierarchal role in relation to the process management systems but rather acts as a link between the main processes in order to stabilise the overall operation. In doing so, the process coordination centre undertakes the following groups of tasks:

- Transfer and monitoring of reference values defined by the tactical production management system for internal stabilisation of the processes
- Construction and assurance of coordination rules and standards
- Construction and assurance of adequate communication channels.

The process coordination centre tasks for supporting self-optimisation and synchronisation of the main processes are summarised in Fig. 6.39.

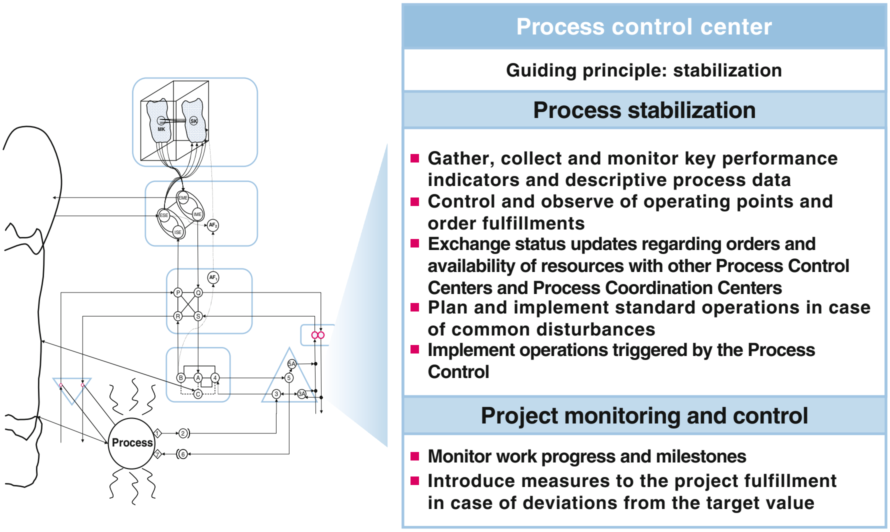


Fig. 6.38 Organisation and overview of the tasks of the process control centre

The tasks of the operative production management system also comprise guaranteeing the internal and external stability of the subprocess (process management system), monitoring and maintaining stability between the subprocesses (process control centre) and guaranteeing synchronisation between the main processes (process coordination centre).

To support the implementation of the aforementioned tasks, the associated management processes and information flows are described below.

The lines between the main processes represent the exchange of production factors which also include information (see Fig. 6.40). All information flows described in the model as internal output and internal input (Oi and Ii) are integrated (“2” in Fig. 6.40) as well as the material flow and access to the labour factor. The external information flows (Oe and Ie) exchange the processes via the connection with the environment (“1” in Fig. 6.40). All project monitoring and control information is fed into the process coordination system and can be requested in real time (“3” in Fig. 6.40). All main processes have interfaces with this data.

The logic of the VSM process images represent the framework for the design of the management processes of the tactical and operative production management system. They place the tasks of the individual controllers into a logical time sequence and structure the information flows through the channels between the controllers. To carry out the tasks described above, the most common interpretation of the VSM processes—according to Gomez—(see Gomez 1978, p. 48 ff.) must be supplemented with specific measures. The enhanced interpretation gives us the VSM process illustration in Fig. 6.41. The names of the node points correspond to the significance of their tasks.

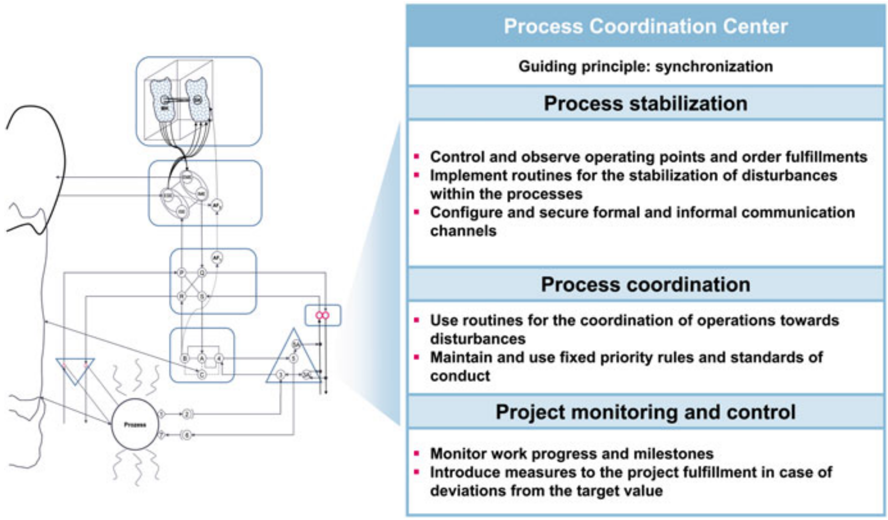


Fig. 6.39 Organisation and overview of the tasks of the process coordination centre

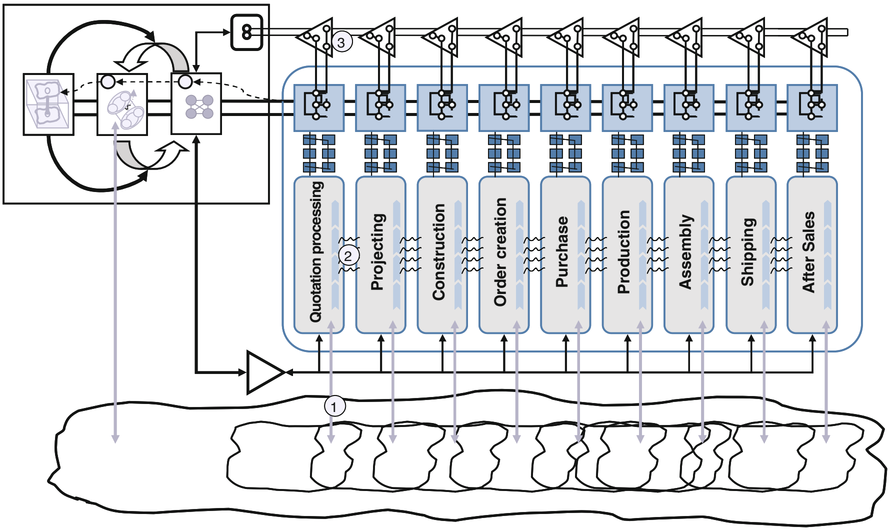


Fig. 6.40 Allocation of information types according to Schmidt (2008) into the structural model

Unlike nodes A, C and 4 the decision-making node B has only a forward-travelling information system. Nodes A, C and 4 reach decisions by iteration and coordination cycles in order to implement the specifications of the metasystem and to manage problems. While B is largely informed by coordination content, its central task is to

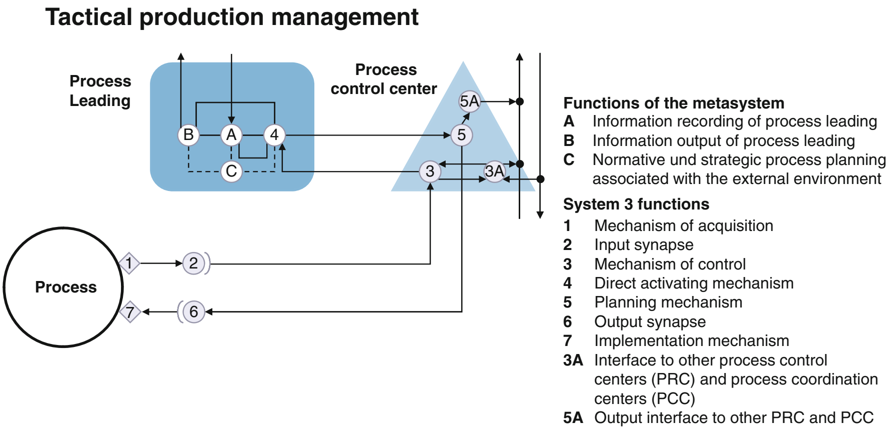


Fig. 6.41 Process illustration of the management mechanism of a main process

dispatch the decisions from A, C and 4 to the metasystem according to the quality criteria of the information.

Node 3A is not only an output but also an input for information into the control mechanism. With this adaptation, a process management system can see information from other processes or from the process coordination centre and can be taken into account in the decision-making process (feed-forward). The output of node 3A is now the relevant information of the monitoring mechanism, i.e. the status information for control of the order processing, availability and if necessary error messages and disturbance warnings. Node 5A however communicates the plans and programmes for process management as a response to problems in the process coordination centre cycle. This dual-differentiation means that status messages from a subprocess are detected by the overall system in real-time before a decision about intervention has been made. At this point it is necessary to decide where this increased transparency can add value and where it would simply cause unnecessary interference.

According to Beer, point C represents systems 4 and 5 of this recursion level. In order to correspond to the structure of the VSM and thereby to fulfil the conditions for viability, node point C allows interactive connection with the process environment. Only by this enhancement point C can undertake the strategic tasks of the local system 4.

The reference variables (reference values) which initiate these servo mechanisms are defined by the tactical production management system. In turn the process management system provides continuous information to the process coordination centre and the tactical production management system about the actual status of the reference values for internal stability. The process management system and the process control centre use the reference variables, available to them, to achieve the required status. Thus they interpret or transfer the global reference values into actual proposals, schedules and standard responses for the individual processes.

The modus operandi of the process coordination centre is based on the coordination measures as well as on an escalation model which is modelled by the tactical production management system and practised by the process control centres and the process coordination centre. In principle each process control centre attempts to compensate for any deviations from the planning milestones or reference variables itself before involving the process coordination centre in reconciliation with neighbouring or affected process units. If no solution can be found for an order or a central reference variable, the tactical production management system is involved to remedy the problems or prioritise and/or reschedule orders. The coordination mechanism is therefore able to absorb a high degree of variety and to induce only limited variety into the tactical production manager.

6.3.4.3 Implementation by Means of Control Logic Within Production Management

Generally, there is a differentiation between two types of control engineering system. In the open control system the incoming control signal is independent of the outgoing signal of the system. On the contrary, in a closed-loop control system, the input signal is influenced by the output signal in the form of feedback and control. The High Resolution Supply Chain Management (HRSCM) project focuses on closed-loop systems because these permit a feedback of information to the controller (e.g. time-dependent stock level, delivery times, restocking times etc) and can therefore react to external influencing factors. In this case the greatest challenge is in the control of dynamic systems which are defined as a functional unit whose input and output parameters change over time. Therefore they represent time-dependent functions. Special features of controllable dynamic systems include linearity, causality and temporal stability. The elements of a closed-loop control system are the dynamic system itself, the controller and a sensor integrated into the feedback loop. Figure 6.43 below illustrates the basic structure of a control system.

The influential variables on a system are its time-dependent input and output variables and equally time-dependent disturbances which can occur at various points along the entire control route and which influence the control variables. The input is manipulated directly via the controller. The measured deviation from the reference variable results from the difference between the reference variable and the fed back signal. The reference variable is the target value of the control system (e.g. target stock level or desired schedule-adherence). In order to achieve this, the control deviation is fed to the controller which in turn supplies the system with an input signal with the result that the reference variable is achieved.

In the context of Supply Chain Management (SCM), state-of-the-art approaches are usually based on optimisation problems. Some approaches are also involved with transferring control loops to the area of planning and controlling of logistical systems. Since the supply chain can be described as a dynamic system, the primary requirements are a flow-orientated and continuous analysis of the time-dependent input and output parameters. Specific requirements must be met in this case.

One of these requirements is expressed by the Nyquist-Shannon Theorem. It states that the measuring frequency must be at least twice the frequency of the input signal.

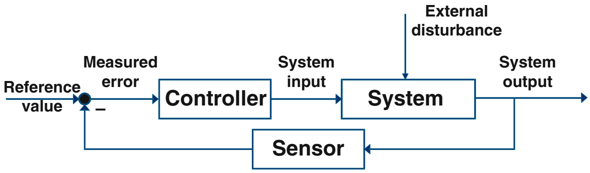


Fig. 6.42 Control system

This is also referred to as the "Nyquist scan rate". Examples for the implementation of a high scan rate in practice include the employment of RFID technology and the application of additional real-time data. Along with the scan rate, the stability and the response behaviour are also critical elements of a control loop. Stability is a qualitative characteristic of the control loop and the essential requirement which is prioritised above all other criteria. A control loop is stable if, after changing any reference or disturbance variable, the control variable assumes a stable value or the status variables achieve equilibrium over time. In order to be able to appraise stability, there is a variety of very different stability criteria including for example the processes according to Nyquist, Hurwitz or Routh. The response of a second-order control system with a small damping factor on the increase of the service level is shown in Fig. 6.43 below.

Typical measured variables for stability and the response behaviour of a system are the adjustment time (T_s) and the rise time (T_r). While the rise time is defined as being the time required to increase from 10–90% of the target value, the adjustment time is the time taken for the fluctuations of the measured value to settle within a predefined corridor.

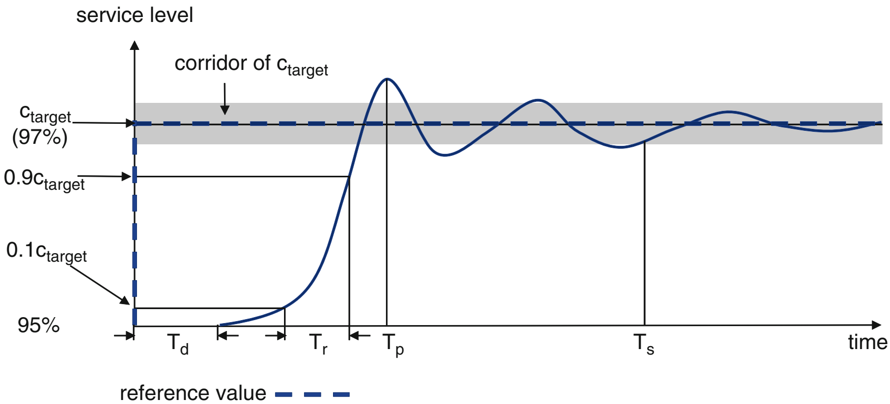


Fig. 6.43 System response in the event of an increase in the service level (92–97%)

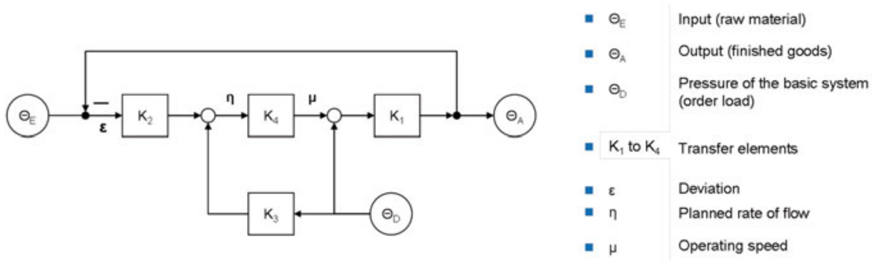


Fig. 6.44 Production control system according to Beer

As described above, the company processes can be considered as a control loop because they incorporate a sequence of target setting, monitoring the target achievement, discrepancy analysis, initiation of improvements and target correction as required.

In his book "Kybernetik und Management" Beer describes how control elements of systems, which do not comprise technical components, can be synthesised with the aid of control engineering. Therefore, the system of industrial production control can also be constructed on the basis of control theory and is described in brief below.

The input (Θ_E) of a production system is the raw material, the output (Θ_A) is the final product. The work between input and output proceeds at a specific speed (μ). The system is under a particular pressure (Θ_D) according to the ongoing orders. There is a primary feedback mechanism between Θ_E and Θ_A which measures the difference between the ideal status (output) and the initial status (input). This deviation is fed back into the system via an operator (K_2), which undertakes the necessary adjustment to the planned flow speed (η). This is transferred by a further operator (K_4) into the actual speed (μ). The order status (Θ_D) has an influence on the actual production speed and is also fed back by a third operator (K_3) in the form of the planned speed. The primary effect of the order load which affects production is immediately obvious at the output. It is represented by the most important operator (K_1) (Beer 1973, pp. 203–204). Figure 6.44 summarises the essential relationships based on a control loop for production control according to Beer.

6.3.4.4 Appropriation of the High Resolution Supply Chain Management Approach

Having described the basic structural model of HRSCM based on the VSM and control engineering, the practical implementation is addressed. For this, two typical processes are designed; "inventory management" and "production control", which represent two central processes of production management, according to the core process management mechanism described above.

Optimised Inventory Management Through High Resolution Supply Chain Management The sub-demonstrator "High Resolution Inventory Management Logic" exemplifies the improved planning quality with simultaneously reduced planning time and costs. Unlike conventional inventory management, the inventory management logic in the context of the above sub-demonstrator uses master data and historical item transaction data along with available real-time data. By now there is no inventory management process which is able to supplement the existing planning with real-time-compatible logic in order to plan order processing more efficiently and more cost-effectively and significantly increase the planning quality. It is expected that "high-resolution inventory management logic" will help inventory managers to achieve greater transparency over the entire range of articles, design and control processes more intelligently and improve the level of supply services to customers.

The central area of activity of inventory management on the part of the stock-based producer involves the three planning components; requirements, stock level and procurement planning. These are implemented in turn in each planning cycle. Requirement planning precedes the other two planning processes and is aimed at predicting as precisely as possible the expected market requirement in the subsequent period to ensure that customer orders can be filled completely and as fast as possible. Requirement forecasting depends on the ability to plan requirements either by programming means, i.e. based on deterministic assumptions of the consumption trends of the respective parts in the future or by consumption, i.e. based on known seasonal consumption levels.

The next component is inventory planning, i.e. defining the necessary safety stock levels and order preparation stock level (reorder level). This requires consistent inventory management and reliable forecasting or requirement values from the requirement planning stage. After forecasting the demand of future planning periods, the inventory management system determines the inventory required to achieve the necessary level of supply readiness. In this case a basic challenge is to integrate the ongoing orders into the forecast-based inventory planning. This indicates reconciling the reorder level (defined on the basis of the forecast) with incoming orders.

The next planning process—procurement planning—is used to define the optimum procurement quantities, taking into account all procurement costs (inventory management, purchasing, transport, etc.) along with storage and capital lock-up costs.

Inventory management faces various challenges in everyday company situations Guaranteeing production and supply while reducing stocks to minimise costs is a basic conflict of aims inherent in logistics. Furthermore, there is a necessity for improving the logistical efficiency through reducing procurement costs and capital commitment while increasing the liquidity of the company. The target system for the economic appraisal of procurement logistics can therefore be described by three quantifiable dimensions: the production supply, liquidity and procurement costs.

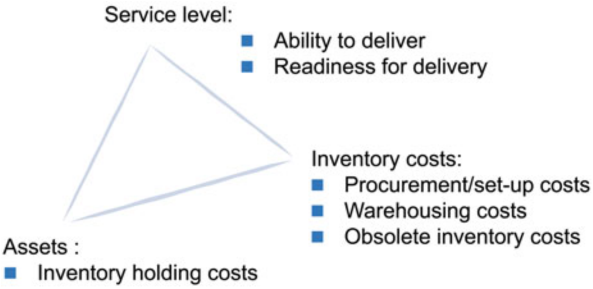


Fig. 6.45 Inventory management target system

The basic task of inventory management is to optimise these mutually-conflicting target dimensions. The aim of inventory management is to increase the production supply reliability (supply service) while simultaneously increasing liquidity and reducing procurement costs.

The supply service dimension comprises all factors required to ensure production with all necessary raw materials, semi-finished products and other materials. It is measured by the supplier's supply capacity in relation to the customer's desired schedule and by the supplier's supply readiness in relation to a delivery schedule previously agreed and confirmed with the customer. Liquidity describes the value-based commitment of capital to supplies and/or stocks. Procurement costs include all relevant ongoing costs, e.g. purchasing, inventory management, ordering and processing, obsolescence costs and the costs of storage (warehouse costs and capital lock-up costs). The operability of the target system variables is guaranteed by logistical parameters which are defined within a characteristic value system assigned to the target system. Figure 6.45 gives an overview of the most important target values of inventory management.

While the reference variables of many planning processes and planning logics of today's inventory management rely on the use of static average values, a key feature of high-resolution inventory management logic is the renunciation of average-based planning in favour of a *real-time-compatible adaptive planning method*. The planning method adapts average values periodically to the actual situation based on real time information. The resulting *real-time planning corridor* enables consideration of dynamic reference and control variables without however excessively increase the volatility of the downstream business processes. Planning complexity, time and costs are kept to a minimum while planning quality substantially increases. To prevent inefficiencies such as excessive stocks, integrated planning processes are supported within the context of high-resolution inventory management with a *differentiated procedure*. A set of procedures is deposited for each planning process in the context of demand, inventory and procurement planning. These are selected on the basis of decision-making logic to optimise the planning purposes of the articles with regard to the target system. The use of differentiated procedures enables more precise and also more efficient control of the company resources.

In order to use the real-time information appropriately within the context of inventory management, it must be prepared, filtered and aggregated previously.

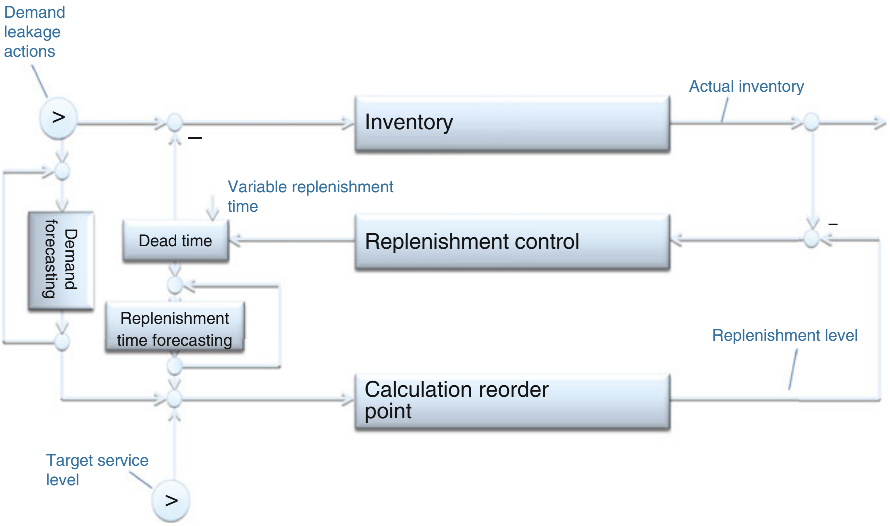


Fig. 6.46 Inventory management process control loop

Preparation and filtering processes ensure that the real-time information is free from error and only the respectively required data is used in inventory management. Hence information overload and increased complexity are avoided. *Aggregation of information* in this context means that the real-time data provided must be transferred to a different aggregation level according to quantity and time in order to be able to use it effectively for inventory management. For example it is not worth initiating procurement after every single sale of an item. For cost-effective dimensioning of the batch size and/or procurement quantity, the overall range of items must be taken into account over a specific period. Figure 6.46 shows a typical cascaded process control loop for inventory management.

In order to validate the added value and benefits of high-resolution inventory management logic for industrial companies, the potential of the developed planning system is demonstrated in a demonstrator the form of a test structure in comparison to the current state of the art. The test structure enables modelling and simulation based on real data. For this purpose, the inventory management process of an ERP system is modelled, in which as a basis products with its specific master and transaction data are provided.

The subsequent simulation analyses the characteristics of the planning systems based on different variation parameters and the formulation of two experiments: conventional inventory management logic and high-resolution inventory management logic. Variation parameters are selected in the way that they allow to simulate the situation as real as possible, e.g. requirement volatility, target available-to-promise level, replacement time, replacement time fluctuation, forecast errors, cost rates, warehousing model, consumption model. The object of the investigation of the simulation based on real data is the central target variables of inventory management illustrated in Fig. 6.45: stock costs, liquidity and service level.

Optimised Production Control Through High Resolution Supply Chain Management The challenges of production control are today not so much the development of tools to solve subtasks but much more due to managing the decision-making complexity. Production control problems have given rise to a variety of solution concepts. However there are no systematic approaches to guarantee the configuration of solution concepts while simultaneously being able to appraise the effects of this configuration on the logistical performance of the production.

Furthermore, current economic environments require a greater focus on the rapid adaptability of production control. Companies are being required to devote increasing time and energy to adjusting part list structures, changing inventory management logic or introducing software-supported fine tuning (e.g. MES systems). These examples can often take more than a year to implement. This shows that the adaptability of production planning and control can no longer keep up with the dynamics of economic change. As a result, targets have been reprioritised. The focus has shifted from use of resources and schedule-adherence to short throughput times and stock levels.

The approach to the configuration of production control that was developed in the context of HRSCM on the one hand comprises a configuration and a simulation model of production control. A four-level model as shown in Fig. 6.47 below is used as a basis.

The aim of the analysis layer is not the evaluation of traditional Key Performance Indicators (KPI's), but to illustrate the time-based processes with the opportunity of displaying data through the interaction with the user in aggregated or even fine grained form. This reveals interrelationships which cannot be illustrated by generating characteristic values which are usually based on averages. Figure 6.48 illustrates examples of an investigation of the process communality and throughput time variance.

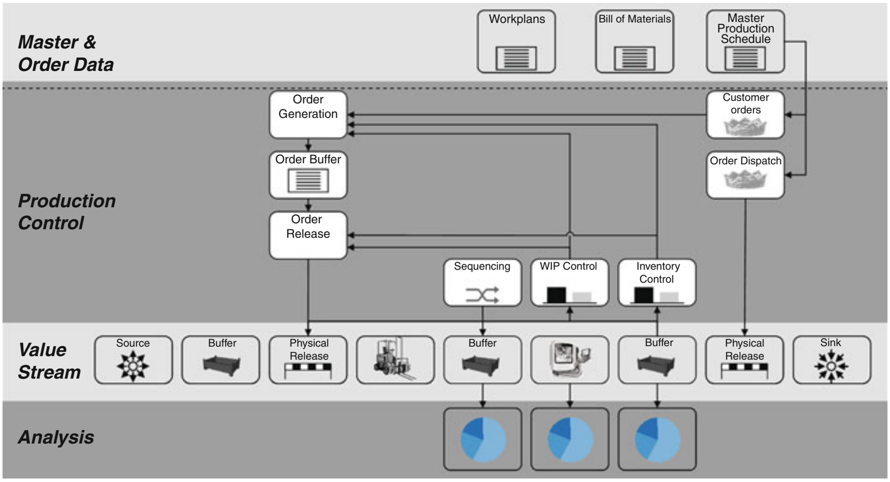


Fig. 6.47 Levels model of value flow-orientated production control

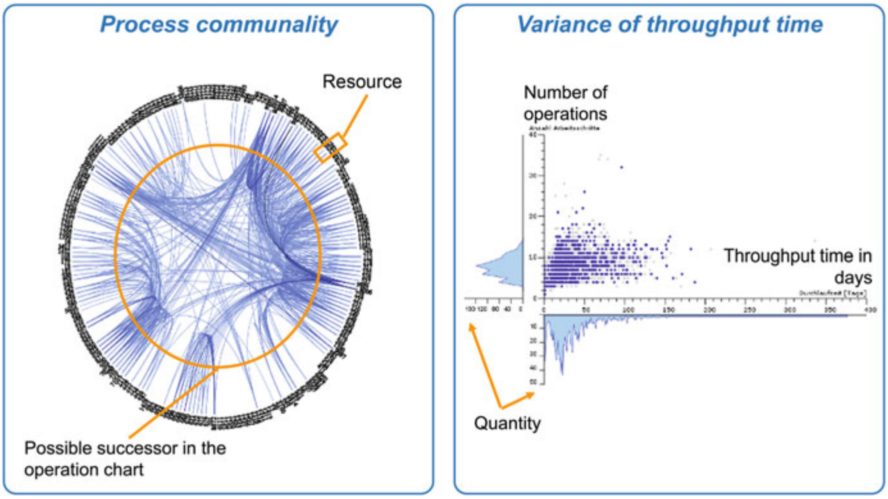


Fig. 6.48 Technical investigation results for process communality and throughput time variance

Process communality is a measurement of the variety of material flow within a production. Each production resource usually has one or more preceding or succeeding processes. If this kind of production system is exposed to a large number of production orders, there are often converging or diverging material flows. This means that there are few uniform material flow routes through the production resources. One measurement for this is process communality which is shown on the left-hand side of the figure. Within the context of a tool specially developed for the analysis, it is possible to select individual production resources and to investigate the number and or frequency of use of predecessors or successors.

In contrast to process communality, the throughput time is a traditional investigation variable. It is generally specified as an average in combination with its variance. In the right-hand part of the figure, the throughput variance is illustrated in the operational schedule depending on the number of operational steps. In the selected type of illustration, the frequency of distribution of the two variables was shown separately on the respective axes to gain additional information. It is possible to select individual periods and/or product groups for which the analysis should be displayed. By using such overviews, the information content increases sharply compared to separate KPIs. This provides a good starting point for analyses and reconfiguration of production control.

The value stream level represents the production process and is used to determine monitoring sections along the value stream. It is connected to the production control level which describes the control configuration.

The level model is in the following presented based on a value stream-orientated approach for the simulation of production control. The aim of the *value stream-orientated simulation* of production control is to enable systematic evaluation of these change processes within production control on the basis of the Enterprise

Dynamics simulation platform and to reveal the limits of existing configurations from the start to enable change processes to be introduced earlier. The basis of the simulation is a combination of the traditional value stream design and production control logic (Lödding 2005).

The work schedule data and the order data are imported into the master data layer (level 1) from files with standardised interfaces. The work schedule incorporates the item number, the workflow number, the resources employed, the processing and setup times, the batch size and the planned throughput time for evaluation. The order data must include the item number, a quantity and the required deadline.

The control layer (level 2) contains the interchangeable modules for order generation, order release and sequencing. They can be assembled in any combination by the user according to requirements. For example, LOR (*Load-Orientated Order Release*) can be combined with MRP order generation and a FiFo (*First In First Out*) rule for processing the queue upstream from a machine. The sequence rules are always applied at a buffer/store upstream from a production process.

The value stream of the control section to be investigated is reflected in the value stream layer (level 3). Here the material flow is modelled by use of source, sink, buffer/stock, release, production and transport elements. The source generates the item to be produced. These then awaits release in a buffer/stock. Once the order release is given according to the selected strategy (control layer in level 2), the item is released with the corresponding batch size for "processing". It is then forwarded via a transport module (facultative), according to the work schedule, to the processing components. Each processing step comprises a buffer/stock atom and the production process atom. The parts first reach the upstream buffer which if necessary initiates a sequential rule calculation (see level 2). They are then relocated to the processing station. After processing, the batches must be re-separated in the UnPack atom before new batches can be packed for the next processing cycle. These new batches then require a new release.

The analysis layer (level 4) contains analysis options. Wait times and inventory projections can be mapped. Strict separation of the information and material flows are required to ensure the modularity of the model. In this way the control modules can be exchanged without adjusting the value stream model. In most companies currently the order release rules ASAP (release the order as soon as it is recognized) or MRP (release the order on the release date) are used.

The value stream-orientated simulation of production control permits the systematic evaluation of the combination of control logics and a specific selection and recommendations for companies. Initial results from MRP control show that even slight turbulence in the order programme have major effects on the stock level. From a certain level of intensity of order exchanges in the order program on, stock levels no longer rise. By implication this means that minor improvements have hardly any effects on companies with high stock levels and MRP control. Only after a breakeven point results improve significantly.

6.3.5 *Industrial Relevance*

The industrial relevance of the research work is now illustrated on the basis of potential applications within industry. Applications in practice can be divided into two aspects. Firstly, the structural model developed offers support in the analysis and design of the production management system and of the business processes. Secondly, the resulting operative control loops of inventory management and production control enable optimisation of the fundamental components of production planning and control.

6.3.5.1 Analysis and Design of the Production Management System and of the Business Processes

The structural model described here can offer valuable support in projects for the analysis and design of the production management system and of business processes in practice. The structural model can already be used in the analysis phase. It provides help in structuring the as-is-analysis and accelerates the analysis of weaknesses. However, the developed structural model eliminates the open search for solutions because the solution approaches are already anchored in the model. It therefore requires a logic which assigns the potential solution approaches within the model to the analysed weaknesses. This takes place in the classification phase. The actions are then evaluated using standardised criteria and submitted to the project management which then decides on the implementation. The model application steps correspond to the design phase (see Fig. 6.49).

Phase 1—Analysis Phase The aim of this first phase is the substantiated analysis of the actual situation of company structures and processes. The reference model provides assistance in that it structures the process. Here the reference model functions as a template which is laid over the actual status for rapid identification of weaknesses. The template need not be known to the whole project team. Target-orientated moderation makes more sense here than overloading project participants with a partially abstract model.

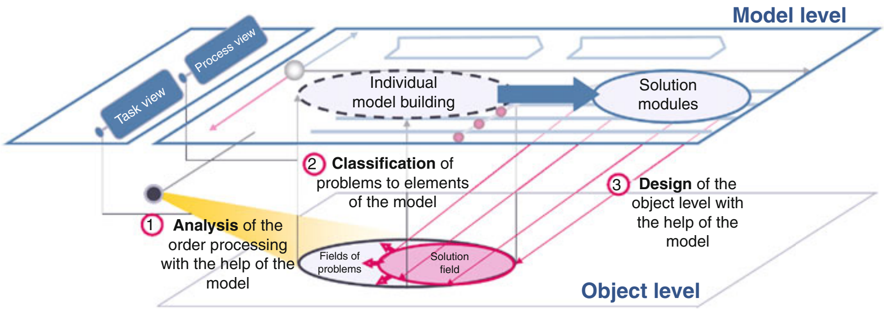


Fig. 6.49 Application logic of the structural model

Due to the different mechanisms and the variety of participants within the company, the process and structural analysis takes place within two analysis streams. The first part of the analysis is concerned with operative business processes and their coordination (part A). The target of the analysis is to improve the efficiency of the business processes. The second part of the analysis is aimed at investigating change mechanisms and at the integrated definition, transfer and monitoring of targets. It is involved with the effectiveness of the production management structures from a cybernetic perspective and takes into account management processes but not business processes (part B).

A—Analysis of the Efficiency of the Business Processes The process follows the logic of the process and structural analysis in the three-phases- concept developed by FIR at RWTH Aachen University. The aim of the module is to record the actual status of the order processing processes and the associated information flows. Methodological tools which can visualise value streams and IT-, process- and structural relationships can be used to document the organisational structure and the IT landscape and to investigate the workflow management and information flows. The record of the IT landscape involves analysing the data and the system architecture which support the order processing. The order processing processes are then collated into company-wide main- and subprocesses. Standard processes with no discrepancies are recorded first. In addition processes, which may or may not be standardised, are described in the case of disturbances or deviations.

B—Analysis of the Effectiveness of the Production Management System Decisions and changes within the company are seldom consciously perceived as structures and so it makes sense to implement this analysis phase with themes that are as concrete as possible. The financial controlling function is an ideal entry point. The detected characteristic values provide information about the control of the operative processes and about the control of the overall system. It also provides more or less structured forms of information about developments within the environment in terms of sales, production development and business management, Corresponding interviews can thus be used to create a map of the reference variables including their dimensions. In the next stage, the question of the availability of information forms the basis of decision-making. The recorded characteristic values are investigated to determine the extent to which they correspond to the criteria of currency, completeness, vulnerability to error, comprehensibility and relevance. If control elements are comprehensible, the individual decision-making processes and their links are then investigated. In this case, an analysis is made of which events can instigate or have instigated structural changes. In this case, particularly the two basic changes in direction from the operative to the strategic and vice versa should be analysed. The integration of management tasks is ultimately expressed in the extent to which changes are expressed over different hierarchical levels in successive characteristic value systems and change processes.

Phase 2—Classification Phase The starting point of this project phase is the assignment of the weaknesses analysed in phase 1 to causes which relate to lack of completeness or lack of integration of elements of the cybernetic production management system. This therefore transfers the reality of the situation to the model level of the cybernetic production management system. For this purpose, clusters of weaknesses with common causes are created and then the causes are assigned to problem-free elements from the reference model. Finally, the benefits and costs of the design of the individual elements are evaluated in order to prioritise the implementation sequence.

Phase 3—Design Phase Corresponding to the structure of the reference model and the approach to analysing weaknesses, actions are once again divided into two dimensions. The actions are detailed and implemented, their success evaluated and the project concluded in the design phase. In the process, the main features of the strategic actions (part B) are defined as they provide the basis for the target-orientated development of the operative processes and their efficiency (part A).

A—Actions for Increasing the Efficiency of the Business Processes The order processing process levels are now reorganised. In this case the project team starts with a reference design of the main processes. The subprocesses are then developed. The weaknesses and their causes are eliminated as far as possible by changing the processes, information flows and control mechanisms. For this purpose the structural model and the operative processes already developed provide a template which is used for the design process.

B—Actions for Increasing the Effectiveness of the Production Management System The actions for increasing the system effectiveness are implemented within the strategic production management system. The creation of a production system for implementing a production strategy through programmes should be the first step of the implementation. The requirements of the operative processes and their logistical performance, planning and flexibility are also defined at this level. In the process, the design of the actual production strategy is only the first step and must be reflected in the cybernetics-orientated design of the management mechanisms, information flows and physical value chain processes.

6.3.5.2 The Potential of Optimised Inventory Management Through High Resolution Supply Chain Management

The competitive pressure on companies is steadily increasing as they have to bring new and increasingly customer-specific products to the market in shorter lifecycle times. Customers are also demanding shorter delivery times. Implicit in these framework conditions amongst other things are greater demands on logistics systems. Particularly affected are the areas of procurement and distribution logistics along with inventory management at the material and also the finished product level.

The BVL study “Trends und Strategien in der Logistik” (2005) addresses this problem in its survey and proves that reducing stock levels and minimising capital

Table 6.4 Design of control loop modules in the case of classical logic and HRSCM logic

Controller	Classical logistics	Logic of high resolution supply chain management
Forecast of the replenishment time	Global average value	Moving average
Demand forecast	Simple exponential smoothing	Simple exponential smoothing
Calculation of the reorder level	Dynamic calculation	Dynamic calculation (current replenishment time considered)
Procurement control	Economic order quantity	Economic order quantity

lock-up costs are of great significance both for industry as well as for commerce. Against this background, the industrial relevance of support for inventory managers by situative automation of consumption-controlled scheduling is clear. In particular the flexible adjustment of control loop-based logic to changing framework conditions of procurement (for example, varying replenishment times and customer requirements) increases their usefulness and performance capacity as compared to classical processes.

In a study of potential an initial development stage of the HRSCM scheduling logic illustrated in Sect. 6.3.4.4 was compared with traditional methods currently used in industry. It displayed clear and considerable potential for optimisation. The investigation compared a traditional dynamic scheduling logic based on average values for replenishment times with the HRSCM logic described above (see Table 6.4). The HRSCM logic is distinguished in this expansion stage by the regular update of the replenishment time and its method of forecasting based on moving averages. The study accordingly reflects the potential of continuously recording the replenishment time and the use of actual values in planning in relation to the service and stock levels.

The investigation included determining the service level and the accumulated stocks achieved using the scheduling logic for different scenarios. The scenarios were formed by the combination of different requirement and replenishment time profiles (see Fig. 6.50) There was clear optimisation potential in the use of HRSCM logic in the form of an 8% reduction in stock level with the same service level and/or a 46% increase in the service level with a stock level of just 12% more. This represents a clear increase in the logistical performance capability.

6.3.5.3 Potential of Optimised Production Control Through High Resolution Supply Chain Management

Due to the variety of improvement actions in the area of production control, potentials are very difficult to quantify in general. In addition, the rate of improvement greatly depends on the starting point of the respective control system. Therefore it is possible only to indicate typical potentials. In the context of a case study,

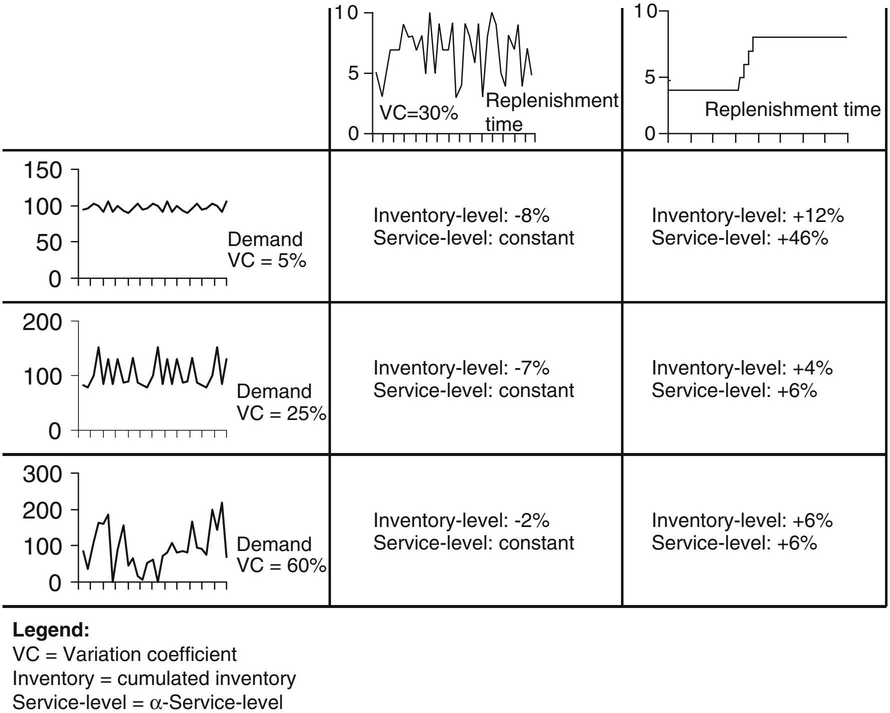


Fig. 6.50 Optimisation potential of HRSCM logic in different scenarios

the proposed concept for the configuration of production control was applied to a job shop production system. This was characterised at the start by poor schedule adherence, long and particularly varied throughput times and high rotating stock levels. An Advanced Planning & Scheduling System (APS) was used to carry out all the tasks of production control. On the basis of the above analyses, the system was switched to FiFo (First in First out) control with shortage-orientated order release.

Measure variables were generated over the three operating statuses of starting status, transition phase and improved status in order to permit a comparison. Firstly, the rotating stock or WIP (Work in Process) was designated as a measurement for management of order releases which has a direct impact on the throughput time. Secondly the performance of the system was measured by totalling the processing minutes of the production orders leaving the system each day. Figure 6.51 illustrates the results of the measurements. It is obvious that there is an initial clear drop in performance just after the starting status. After this short fluctuation phase however a greater average performance is achieved along with a lower level of stock rotation.

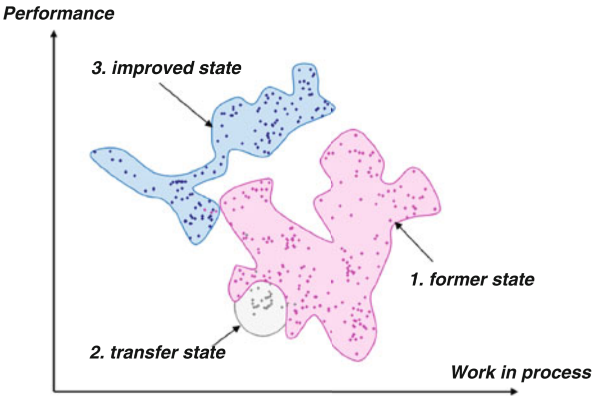


Fig. 6.51 Measurement results in the production control optimisation process

6.3.6 Future Research Topics

So far amongst other things the requirements of company viability of have been investigated in the “High Resolution Supply Chain Management” project. These include the following aspects:

- Real-time capacity of systems
- Vertical integration of information flows
- Horizontal integration of information flows
- Self-optimisation of systems

On the basis of the findings obtained, research into innovative implementation options emerges as a relevant future research theme. Therefore in the following possible future research themes related to this issue are presented.

Real-time capability in companies means having the right information at the right time at the right place for the right purpose (Martin 2003). Idealized real-time systems use information immediately after it has been generated. They automatically detect data as it is generated, avoid buffer and exhibit no medial or semantic breaks. Idealized real-time systems also select the most important information for a decision. They make and implement decisions automatically at the decision-making point (Fleisch and Osterle 2004, p. 3 ff.). Real-time capability requires a high degree of currency and an equally high granularity of data. This real-time data is generated partly within the actual company and partly externally.

So-called *vertical integration* is required to be able to use the real-time data generated within the actual company for the purposes of maximised information transparency. Vertical integration is the synchronisation and harmonisation of the different IT systems available within companies. In this way, for example, signals from operative hardware units are projected into the Operating Data Acquisition System (ODA), the Manufacturing Execution System (MES) and the Enterprise Resource Planning System (ERP). These operative hardware units include for example sensors and readers from auto ID technology e.g. 2D-Barcodes or RFID tags. It is also possible for manipulations in the ERP system to have direct effects on vertically integrated operative hardware units. For example, the entry of a customer order into the ERP system could lead directly to changes in an automated scheduling system.

As with vertical integration which relates to the reconciliation and synchronisation of internal information flows, *horizontal integration* offers potential for future research work. Horizontal information in this context is understood as the synchronisation and reconciliation of different IT systems located within the supply chain. The research theme of horizontal integration is highly innovative and of particular scientific interest because relevant information for the real-time compatible design of companies is not only generated within the company itself but can also come from cooperating concerns. Furthermore these cooperating concerns not only include direct customers and suppliers but also all the companies involved in the corresponding added value chains, from raw materials suppliers to consumers. Industry-wide order processing is already possible today with the use of standardised EDI interfaces. These interfaces simplify electronic order processing. It is possible to exchange further data in real time within a supply chain along with data about electronic order processing. In this way, logistical cooperation concepts can be implemented with minimum outlay and a high level of transparency. Future research themes might be concerned for example with the implementation of logistical corporation concepts as Web-based services. It is worth considering the IT implementation of the following logistical cooperation concepts and/or SCM concepts.

- Collaborative Planning, Forecasting and Replenishment
- Continuous Replenishment
- Efficient Consumer Response
- Industry-wide eKanBan
- Just-in-Time/Just-in-Sequence
- Quick Response
- Supplier Relationship Management
- Value Added Partnership
- Vendor Managed Inventory.

As well as the aspects described above, self-optimisation is a central requirement of the viability of companies. A system is self-optimising when the actual situation of the system is continually detected by the interaction of the system elements, when system aims are defined and prioritised within the context of a defined level of autonomy and when the system behaviour is adjusted accordingly (Frank et al. 2004, p. 22). In the context of the analysis of the actual situation of the system, the current internal system status and the systems environment are investigated and described in detail. The status description can also incorporate information from other systems and existing empirical values.

Changing the target system can be achieved by selecting targets from a prespecified, discreet, finite number of possible aims, by changing existing aims or by generating new independent aims. Adjusting the system behaviour, which represents the feedback of the system-optimising control loop, can affect the parameters, structure or behaviour of the system (Frank et al. 2004, p. 22).

These self-optimisation findings lead to two further questions which relate to open research requirements:

- How can the self-optimisation of companies be achieved by the homogenous use of control logics?
- How can the reference value corridor of the operative unit be dynamically reconciled with a company's target system?

The *integral use of control logics for the self-optimising control of a company* follows the cascading principle. At the lowest hierarchical level, there are many control loops for continuous self-optimisation of the operative units which have already been intensively researched in the context of "system 1" within the project. These control loops work at a hierarchical level and are interconnected by defined interfaces. In case that it is not possible to control the operative units, pulses must be generated from a higher hierarchical level. These points were defined in the previous project as "system 3". From a control perspective, the pulse generator corresponds to a higher level control loop which can control subordinate control loops by varying additional control variables. This type of enmeshing of control loops can be organised according to the St. Galler Management hierarchical company structure. The highest cascade would accordingly represent the normative control of the company which was described in the previous process as "system 5".

Modelling this kind of complex system of enmeshed control loops must start at the lowest level to ensure that all relevant input and output variables are taken into account. The sales forecast and requirements planning are core tasks of production planning and control (PPC). They are particularly well suited to initial modelling and are analysed more closely below.

Amongst other things the following information flows must be taken into account in the control of sales forecasting:

- Historical trends of item- and customer-specific sales
- Internal company disturbance factors e.g. price fluctuations and sales promotion activities
- External disturbance factors e.g. sale of competing products and customers' promotional activities
- Sales forecasts of direct customers and indirect customers up to the consumer

Amongst other things the following information flows must be taken into account in the control of requirement planning.

- Item- and customer specific gross requirements
- Item- and customer specific stock levels
- Item- and customer specific sales forecasts
- Stock level status (booked/available)

The second question generated by the requirement of self-optimisation describes the *reconciliation of the reference value corridor* with the dynamically changing target system of a company. The target system of a self-optimising system plays a decisive role in the system optimisation and comprises internal targets which can adjust and change the system within the constraints of its autonomy based on external targets that are imposed by the direct environment or by other systems or inherent targets which results from the purpose of the design of the system (Frank et al. 2004, p. 30).

The need to adjust targets results from the localised analysis of preset targets prevailing within companies. In the case of non-uniform target setting, individual targets can cancel one another out with the result that preset priority targets cannot be met (Mälck 2001, p. 30 f.). Production economy target systems are the basis of financial production control whose task it is to reconcile production management with the production system (Hoitsch 1993, p. 27 f.). The literature contains a variety of approaches for company target-setting systems. These include target systems with a monetary focus (DuPont system, ZVEI system, RL system (Becker 2008, p. 166 ff.) or comprehensive target system concepts e.g. Balanced Scorecard according to Kaplan and Norton (2006)).

In order for a company to operate in accordance with its target system, a link must be created between this target system and the control mechanisms. Input variables for the system of interlinked control loops are specific parameters which in turn define reference value corridors within the control loops. The synchronisation of precisely these parameters with a previously-defined target system could be considered as a future research theme. To answer the research question mentioned above one might address the following secondary questions:

- How is the target system of a company structured?
- What interdependencies exist between the individual targets of the target system?
- At what points does the target system directly influence decision-making within the production management system? What mathematical relationship can be created?
- Is it possible to connect the target system with the enmeshed control loops of the production planning and control systems such that changes to the target system influence the reference value corridors of the PPC?

6.4 The Road to Self-optimising Production Technologies

Ulrich Thombansen, Thomas Auerbach, Jens Schüttler, Marion Beckers, Guido Buchholz, Urs Eppelt, Yves-Simon Gloy, Peter Fritz, Stephan Kratz, Juliane Lose, Thomas Molitor, Axel Reßmann, Heiko Schenuit, Konrad Willms, Thomas Gries, Walter Michaeli, Dirk Petring, Reinhard Poprawe, Uwe Reisgen, Robert Schmitt, Wolfgang Schulz, Dražen Veselovac and Fritz Klocke

6.4.1 Abstract

A product's success in its market is influenced by both technical and economic factors. If its fitness for purpose is given on the technical side, production costs come into

focus. In an age of short-lived products and fast innovation cycles, manufacturing is coerced into being able to produce economically small production runs and individual items. While in mass production mostly costs relating to processing time dominate, the critical factor for production of individual items is often the time for production setup. In both cases, the resource-efficient achievement of product quality is a basic requirement.

Developments in the past have often aimed to increase performance of production systems by means of increasing processing speed. Advances in control technology as well as in drives have allowed the construction of machines where processes are pushed closer towards their physical performance limits. Processing at the limit of these process domains requires precise knowledge of the relationships between the setting parameters and their effects on process stability and processing result. The precondition for this kind of process control is knowledge about the conditions at the current operating point, in order to decide on the adaptation of values for setting parameters.

The described relationships clarify, that a multidimensional and highly non-linear task has to be mastered, which imposes genuinely new challenges to control systems. Existing approaches have only partly achieved industrial application so far. Many expert systems and Artificial Neural Networks, for example, do not yet have the stability and clarity of response needed for use on production machines running over multiple shifts.

Research in the area of self-optimisation aims to enable technical systems to adapt themselves to changing process conditions. To achieve this, research is needed in the area of process modelling and the development of reduced models to optimise setting parameters, the development of cognitive systems to determine the current operating point, and the creation of control and adjustment systems which, based on machine-readable models, decide and adjust setting parameters for themselves.

The concept of Model-Based Self-Optimisation opens the path to the integration of process models into technical systems which, by using cognitive components, are able to determine for themselves the current operating point within the parameter space, and use this to start carrying out process control. The basis for this is the methodical collection of information about the production process and the creation of reduced models, which can be used to define the target values and how they depend on the parameters with sufficient precision. On the basis of this process knowledge optimisation systems are created, in combination with decision algorithms, which enable an autonomous control of the production system achieving the defined target values, even when random disturbances occur during processing.

For this kind of functionality it is necessary to collect information continuously from the process and process it. The formulation of models to describe the product quality requires knowledge of quality relevant relationships between setting parameters. Even if the initial level of implementation does not provide autonomous process control, this knowledge can already be transferred into application. Validation of the research results on real life examples has led to the concept of information-processing sensor-actuator-systems (ISA-systems). These offer a way for Model-Based Self-Optimisation to realise flexible and re-configurable control platforms which can translate externally prescribed objectives into internal setting parameters using reduced models. This creates an abstraction layer between optimisation and machine control, which also permits implementation of stability and safety strategies.

The continuation of the research project addresses the extension of the parameter space under consideration. Where models are already based on multi-dimensional input parameters, a direct application as a control system is often rendered impossible by the complexity and the resulting scale of calculation resources required. As well as issues of process stability it is also necessary to model the process-machine interactions. But set-up assistants which exist already, and the first controls using scalar parameters are indicating today the potential and the sustainability of the concepts developed.

The solutions achieved in the area of self-optimisation of technical systems, and the documented progress in the area of production processes is based on work carried out by five research partners, who all contributed to the generic solution with diverse requirements, which in turn advanced the specific solution for each application. Many of these results have been transferred to industry, and a common platform was created as a basis for the second phase of the project. Under the title "self-optimising manufacturing systems in *verbund*"[1] open issues are addressed and the area impacted by self-optimisation is extended. With this approach, the whole production site comes into the design scope. This opens up the perspective onto an extended approach to manufacturing planning, in which the entire chain from product design through to manufacturing can be considered. Changes in specifications can hence be assessed with respect to their impact on the manufacturing process.

6.4.2 State of the Art

6.4.2.1 Artificial Intelligence

The concept of Artificial Intelligence (AI) was established in the 1950s by McCarthy (1959). Today, there is no distinct area of impact of the AI due to its diversification and transfer to a multitude of topics. Nevertheless, a certain set of AI methods can be named, which have been used in industry for process monitoring and control purposes.

Originally the first artificial systems aimed at the development of so called expert systems (Savory 1985). Expert systems are numeric systems, which collect and store information about a certain task, draw conclusions based on the gained knowledge and offer solutions to specific problems within the task (Engesser 1993). The stored knowledge, the rules and the input information that is to be processed—the facts—create the knowledge base of an expert system. This knowledge base contains two

[1] Verbund in this context describes the union of cooperating manufacturing systems which themselves react autonomously to changing requirements such that all entities within the union together achieve the global objective of the union.

types of knowledge: static, field-specific knowledge, which does not change during a consultation, and case-specific, dynamic knowledge, which is created using an inference engine or can be entered by an experienced user. This knowledge is structured in the form of relatively simple rules and facts, which as a whole represent the expert knowledge. Expert systems have been developed for the most diverse areas of interest. Examples are: informative and advising systems for design and synthesis in construction or fault diagnosis systems in assembly. (Mattke et al. 1993; Puppe et al. 1996; Cremers 1991). However, the use of expert systems often failed due their high acquisition and maintenance costs as well as due to the fact, that the heuristic knowledge bases and the programmed applications itself were error-prone. In most cases, the sensitivity of the overall system towards partly inherent knowledge gaps and faulty assumptions had a significant influence on the system's overall functionality. Therefore, such expert systems were not able to replace professional experts, but can be used to support their activities.

Artificial Neural Networks (ANNs) model the principle functionality of a brain, in that units which work in parallel, the neurons, are linked to each other via directional connectors. The structure of the network needs to be recreated for each individual problem, where it is impossible to predict its necessary topology in advance. The adaptation of an ANN to the problem is executed subsequently, for example by changing the values of a weight matrix using the back-propagation algorithm (Borgelt et al. 2003). The result of the learning process is an analytical function, which treats the given input information as a variable vector and the received output information as a result vector (Grauel 1992). One of the disadvantages of applying ANNs lies in the sensitivity of the network to the input data. Therefore, the training data and the system topology itself need to be selected in tight reference to the specific issue, in order to avoid overmatching or faulty decision bases. In addition, it is never certain that with an optimisation the global optimum is found.

Genetic algorithms (GA) belong to the class of evolutionary algorithms and are based on the principle of "survival of the fittest". This procedure uses populations of potential solutions and selects promising individuals based on their fitness. Based on the selection, new populations are created by combining the properties of the selected candidates. The candidates in the new population are changed by stochastic mutation, and an additional selection is performed. Evolutionary methods are suitable to handle any problems that are presentable as character strings or tree diagrams (Heidrich 2001).

6.4.2.2 Self-optimisation

Research in the area of self-optimisation covers the development of methods and systems for its implementation, as well as research on the processes to generate the required detail in process understanding. CoTeSys (Cognition for Technical Systems) for example is a project that investigates cognitive solutions for different technical systems such as factories, vehicles or robots. In the case of robots, the corresponding system is a human being and the cognition of these systems is implemented as a classic humanoid form. Hereby, the perception and orientation of the system is achieved using cameras for visual and microphones for audio input, in order to perceive the request that is imposed on the technical system. Alongside the technical aspect, Bannat et al. are researching the basis of humanoid coordination as part of this project. Regarding self-optimisation the project's contribution to self-optimisation consists of investigations on a robot's reaction to unforeseen events (Bannat et al. 2008; Laue and Röfer 2006).

In the field of communication networks, strategies for self-optimisation are already implemented (Litoiu et al. 2005). For example Witkowski et al. use self-optimising systems for the automatic creation of mobile ad-hoc communication networks between robots based on swarming behaviour (Witkowski et. al. 2008). Systems like these generate humanoid robot systems for disaster scenarios which improve their skills by learning.

Self-optimisation is also used for mechatronic systems such as the guidance module of rail-based vehicles. In the area of mechatronic systems, research was done on the needed system structure to achieve internal objectives with the assistance of classic control algorithms (Gausemeier 2005). This work was done in the context of the German collaborative research centre 614 "Self-optimising systems in mechanical engineering", which is currently being worked on at the University of Paderborn. Along with fundamental research on the general understanding of self-optimisation, methods and tools are developed to realise self-optimising products. The developed solutions employ sensors and actuators to control mechanical systems (Isermann 2002). In most cases, these have the advantage of being able to measure and evaluate their operating point directly in order to draw conclusions about their current status without having to build a model.

Production systems must meet the requirement to control machine components and a physical process simultaneously. Therefore, the realisation of self-optimisation for manufacturing systems often has to deal with the management of multi-dimensional parameter spaces and non-linear relations between setting parameters, process variables and process result. It requires solutions for handling complexity within the very short time-scales which are available for numerical calculation and optimisation algorithms during the production process. Nevertheless, the potential of self-optimisation for manufacturing processes in terms of reducing set-up time and improving a process's stability is large (Schmitt and Beaujean 2007). The expected added value from turning this potential into practical use in manufacturing scenarios in high-wage countries is therefore a strong driver for research activities in this field. The state of the art in technology shows gaps and starting points for the focussed manufacturing processes.

6.4.2.3 Meta-modelling and Model Evaluation

"Meta-models" in this context are numeric surrogate models that describe relationships between input variables and output variables on the basis of data sets. They therefore take the place of the complete physical description of the manufacturing process, also known as a white box model, and in this respect describe a process' behaviour, without having to be process-specific with respect to the modelling procedure. Suitable approaches to multi-dimensional data interpolation and data approximation are offered by regression models, Artificial Neural Networks (ANNs) or Radial Basis Functions (RBF), which from the modelling point of view are named black-box-models (Stork et al. 2007; Buhmann 2003). If process knowledge is included in the selection of the base functions or in the implementation of the model structure, then these models are considered to be grey-box-models. The process of creating a meta-model includes a series of steps which determine the quality of the resulting model. The decision which parameters are taken into account, the data sources, as well as the data structure and model structure, are just some of the possible sources of influence which require a high degree of expert knowledge when creating a meta-model and defining the configuration parameters (Coit et al. 1998). While these steps are usually carried out manually and case by case, there do exist common tools for the implementation of models, in the form of application programmes such as Matlab, SNNS (Zell et al. 1991), DesParO (Stork et al. 2007) or SUMO (Gorrisen et al. 2009). DesParO, programmed by the Fraunhofer Institute for Algorithms and Scientific Computing (SCAI) very rapidly creates surrogate models with the help of Radial Basis Functions (RBF). It is distinctive because it has a sensitivity analysis for input values and output values which means that irrelevant parameters can be identified and eliminated. This allows models to be downsized, meaning a reduction in the amount of numerical calculations to be executed. The SUMO product is sold by the Interdisciplinary Institute for BroadBand Technology (IBBT) at the University of Gent, as a toolbox for Matlab. Besides Radial Basis Functions, this platform also offers rational functions, splines, neural networks and support vector machines. In addition to standards like the squared error, algorithms such as pattern searching or particle swarm optimisation can be used to optimise models. SUMO offers an open system structure which allows the export of all model parameters and model functions. This allows the meta-models developed to be transferred to other platforms, such as real-time systems (Efron and Gong 1983).

Cross-validation is a procedure which allows the determination of the forecasting quality of a model with respect to process behaviour which can also be used to compare different models. For this, a data set with N data records is split into k subsamples ($k \leq N$), then, while removing one subsample each time, the single error rate from the training with the remaining samples is calculated. A special case in cross-validation is the leave-one-out, where each time one single data record is held back, so that if the data set consists of N data records then $k = N$. This also allows evaluation of the local error for the data point held back.

6.4.2.4 Manufacturing Processes

Milling Monitoring and control systems available in the market are limited in their field of application. Hereby, in most cases the recognition of a process disturbance is accomplished comparing measured process signals to reference curves acquired at stable process conditions. Therefore, recognition, evaluation and avoidance of disturbances are closely connected to specific operations. Due to their lack of transferability these strategies are not suitable for use in single item or small lot manufacturing. Here, system solutions are needed that enable a reliable process control at varying process conditions. This requirement applies at an even greater extent to machining freeforms, due to a continuously changing tool engagement (Reuber 2001).

The continuously changing boundary conditions and the special characteristics of hard to machine materials often require compromises in planning 5-axis milling processes. Therefore, in most of the cases conservative parameters are chosen, that lead to an overall drop in productivity (Li et al. 2004), increased costs and machining time, which are especially a constraint on competitiveness for production facilities in high-wage countries.

In the past, system-based approaches to control the manufacturing of complex free forms have been researched, but show substantial disadvantages in their industrial applicability. Ismail developed a system solution to suppress process vibrations when milling turbine blades, applying an offline adjustment of the feed-rate in combination with an online control of the cutting speed. Hereby, the feed-rate scheduling approach was more successful than the online control of the cutting speed, which at this point seemed as technically not feasible (Ismail and Ziaei 2002). The approach of controlling the feed-rate was also investigated by Erdim et al., where the feed-rate adjusted is non-proportional to the material removal rate. This approach can only be applied during the process planning steps, and not in process control, as the calculation required is too time consuming for online applications (Erdim et al. 2006). Dohner et al. developed an active control system which was able to manipulate the vibration behaviour of the tool and the spindle as a function of the dominant system vibrations. Using sensors and actuators, they eliminated process disturbances caused by vibration, such as chatter vibration. Although the implementation of a prototype was successful, the transfer of this solution in industrial applications has not been realized until today (Dohner et al. 2004). Mane et al. researched the control of process stability during finish milling thin-walled components by adjusting the cutting speed. Here, it was possible to determine the stability limits of the process offline, and to carry out an upstream process planning, reducing the number of cost- and time-intensive test runs. As this approach requires too much calculation time its applicability towards an online control was proven to be unsuitable for industrial practice (Mane et al. 2008). In 2008, Ferry et al. presented an approach to the optimisation of the feed rate where operating domains and the maximum allowable feed-rate for predefined operating points are calculated (Ferry and Altintas 2008). An approach that was also not applicable as an online solution, because of the required computing resources.

In summary, the current state of research covers certain approaches regarding process optimisation and process control, but did not reach the level of industrial applicability.

Plastics Injection Moulding The precise manufacture of complex and high-quality formed parts from thermoplastic substances produced by injection moulding, places high demands on the design of its tools, as well as on the process itself (Kudlik 1997; Menges et al. 1971). Conventional process control, now established for decades in injection moulding, is based on controlling the screw speed during the injection phase, together with the hydraulic pressure in the injection cylinder or the force acting on the screw during the holding pressure phase. The influence of the screw speed on the forming process in the mould cannot be clearly defined, because of various disturbances, such as variations in materials, pressure and temperature, as well as requirements on environmental conditions (Menges et al. 1971). Variations in quality of moulded parts from one run to the next are the result. These variations occur mainly after process set-up, since with conventional process control, at the start of production steady thermodynamic conditions need to be achieved.

After injection of the melt into the cavity, the holding pressure phase begins, where the material begins to shrink as it cools. By applying pressure, additional material is inserted into the cavity in order to compensate shrinkage. The pressure and temperature within the cavity have a particular importance here, as these values define the shaping of the moulded part. In an injection moulding machine, both electric motors and the hydraulic cylinders are used as actuators during the hold pressure phase. This means that by using the position signals on the servo-inverter or on the hydraulic vent, the pressure plot within the cavity, the so-called cavity pressure, can be adjusted. Based on the relationship between pressure, temperature and specific volume, or pvT ratio, which varies by material, by measuring the temperature a pressure can then be calculated to give a constant specific volume.

Both the direct adjustment of the cavity pressure and the adjustment of the target pressure course according to changes in boundary conditions of the process are included in some existing approaches in science and industry (Michaeli et al. 1992; Pramujati et al. 2006; Smud et al. 1991; Arburg 2010). Due to the behaviour of the controlled system, which is sharply non-linear, time-variable, and includes a time delay, the solutions provided so far which manage the cavity pressure settings and which are mainly based on PID controllers with an adaptation of the parameters, have not led to satisfactory results (Michaeli et al. 2004). The realisation of a freely selectable cavity pressure with a high level of precision and reliability, while being highly dynamic and with simple set-up and adjustment of the control system has not been solved so far.

To adjust the planned course of the cavity pressure to the boundary conditions of the process, the strategy of pvT-optimisation is well known, which allows a calculation of the necessary pressure course for the requested quality of moulding, depending on the melt's temperature course (Johannaber and Michaeli 2004; Michaeli and Lauterbach 1989; Menges and Thienel 1977). Based on the unsolved challenge of the control of the cavity pressure, the effects of pvT-optimisation on the quality of the mouldings could only be partially demonstrated.

Welding In the field of automated complex gas metal arc welding (GMAW) processes, the selection and setting of suitable welding parameters for specific welding tasks today is still done by the welding equipment operator. This implies well-founded technical experience and welding expertise of the operator in order to set the required welding parameters optimally for the given boundary conditions. The identification of optimal welding parameters for each application requires time-consuming, empirical welding tests. Since the mid-nineties some approaches have been developed for monitoring and controlling welding processes using Artificial Intelligence (AI). One application using Artificial Neural Networks (ANNs) aimed to predict welding seam geometry based on machine setting parameters that were entered off-line (Middle and Li 1993). Furthermore, the prediction of weld seam quality based on statistically processed transient data was also a focus in many developments in the field of on-line quality surveillance (Maqbool et al. 1998; Heidrich 2001; Kim et al. 2004, 2005). In addition, the use of Neural Networks for the detection of weld seam imperfections was investigated to determine weld seam quality. For this, process stability and the occurrence of spatter were determined using the transient process measurands welding current and welding voltage, wire stick-out lengths and torch angle (Kim et al. 2005). While Martin has used statistical data based on the electric process measurands for the detection of spatter and weld seam imperfections (Martin 1994), Gladkov et al. used the acoustic power density spectrum to identify weld seam imperfections (Gladkov et al. 1998). The optimal use of neural networks demands a sufficient quantity of training data to appropriately cover the process domain to be modelled, which results in enormous investments of time and economic expenditure to carry out the required welding tests (Matteson et al. 1992; Hichri 2005).

In another approach an expert system was used for on-line process optimisation of GMAW processes. For this, a dialogue-based, off-line expert system to determine parameters and optimise processes was modified, extended and linked to the welding process via a process computer. Several sensors were used to collect information about the welding process. To calculate the target seam geometry, the seam surface was measured by an optical sensor located behind the welding torch. By means of fuzzy logic the weld seam geometry, including the penetration depth was classified. At the same time, the electrical process measurands of the welding process were recorded which, together with the result of the evaluation, were used as input values to the on-line process optimisation by the expert system. The spatial distance between the sensor and the welding torch proved problematic in these works, as this spatial distance resulted in a time lag for the sensor information limiting the controller capabilities (Roosen 1997).

Laser Beam Cutting Sheet metal processing industry has, due to globalised markets, an increasing demand for cost efficient and flexible manufacturing systems. CO_2-Laser beam cutting machines enable the processing of a wide range of different types of materials, and thicknesses of material, which ranges from a few tenths of a millimetre thick stator sheets for electrical motors, to 30 mm thick stainless steel for power plant construction. With machines of this type, either individual items can be produced for medium-sized production companies or mass production parts for automobile construction. The local relevance of this is shown by the share of 5,200 laser flat-bed cutting machines sold in 2008, of which 37% were installed in Europe (Belforte 2009).

Along with the wide range of different materials, high cut quality and the ability to reproduce it are both significant factors for competitive production and machines. A smooth and dross-free cut saves expensive post-processing. Under laboratory conditions, the use of photo diodes or acoustic detectors was experimentally investigated to capture signals for controlling the cutting process (Keuster 2005). Along with the non-spatially-resolving sensors, cameras were used to capture geometric and thermal process emissions. For example, Haferkamp was able to use a heat-sensitive camera to show a correlation between temperature distribution in the interaction zone and the quality of the cut (Haferkamp 1999). Control of the setting parameters based on process changes such as, for example, variable properties between batches of material, or ageing and dirt on optical components which affect the laser beam, has so far only been accomplished in the laboratory (Keuster 2007; Schneider 2004; Zheng 1990).

In laser cutting systems today, devices such as penetration sensor systems, automatic nozzle changers, laser power controllers or the distance control of the cutting head have become state of the art. On the other hand, values for processing parameters are taken from empirically determined technology tables and adjusted by the user as necessary. The anticipated product quality is therefore dependent, among other things, on values such as the composition of alloys and the drift in focus position, which in an industrial setting leads to a loss of quality, or to increased costs for set-up and maintenance. Controlling the product quality today is still not state of the art.

Weaving The weaving process is noted for its very high efficiency. Once set up correctly for a certain fabric, a loom produces reliably for a long period. To further improve productivity, textile producers attempt to achieve this stable condition faster than before. Shortening set-up times after a change of fabric therefore allows a further increase in productivity.

Generally, when setting up a loom for a fabric, setting parameters are used from an in-house operations database. For new fabrics, setting parameters must be established experimentally. These tests involve additional costs for weaving mills. Costs are caused by the loss in productivity on the machine, plus the materials consumed during the set-up phase. Weaving mills in Germany in particular have high costs caused by set-up due to the intensive personnel time involved. In addition, the fabrics which are manufactured are generally technical fabrics. The raw material for technical fabrics, as for example aramide, glass and carbon, are expensive compared to raw materials for the clothing industry.

To shorten the time needed to find correct setting parameters for unknown fabrics, the company Picanol NV Ypres, Belgium introduced the EasyStyle-System in 2006. This system suggests machine settings based on a database of weaving parameters. The database contains setting parameters from many different weaving mills who have released their data for this system. The proposed machine settings in this system are not correlated to resulting characteristics in product quality.

At the Institut für Textiltechnik at RWTH Aachen University (ITA) an intelligent set-up assistant was developed to reduce the set-up time. The core of this set-up assistant uses the warp tension as the process measurand. Warp tension is a decisive process variable for weaving processes. It influences the machine's running behaviour, for example the number of warp breakages, or the cleanness of the front shed and therefore the weft breakages. The sum of these factors defines the economic efficiency of the weaving process.

By the combination of an Artificial Neural Network and an Evolutionary Algorithm, the set-up assistant defines the optimum course of the warp tension. This course is evaluated using defined performance criteria. A disadvantage of the set-up assistant is the extensive effort to train the Artificial Neural Network. In addition, no parameters for product quality are taken into account here.

6.4.3 Motivation and Research Question

The efficiency of production in high-wage countries has a significant impact on the total costs of a fit-for-purpose product. The further development of manufacturing systems therefore has to contribute to reducing costs per item through maximum productivity.

Classic controller structures are in most cases not suitable for controlling and optimising modern manufacturing processes. Processes are increasingly highly dynamic, and the interaction between input values and the processing result are mostly non-linear. Stability limits are therefore determined by a number of parameters and their correlations. Because of an increased number of materials to be processed, and because of batch variations, there is an increasing number of variations in the process input values which have a significant influence on the processing result. In addition, alongside geometric precision, in particular for safety critical components, requirements also apply for the creation of particular properties such as, for example, surface integrity. All this taken together shows clearly that for high performance processes a large number of functional criteria need to be produced within narrow tolerances. Consequently, this means that optimisation is also only possible in small process windows. Particularly in on piece flow production, this increases the production and testing effort because barely any consistent process models are available to describe the dependencies. The applied techniques for setting up a stable process are predominantly based on empirical methods.

It is from these circumstances that the research question is derived, whether technical systems can be enabled such that, under variable process conditions, they can autonomously steer the process at the performance limits, and reliably achieve the prescribed objectives. A pre-condition for this is that the system must know about the process at any time, and generate suitable values for setting parameters itself. This can also be stated more generically as follows:

> Can complex and variable manufacturing processes be set up, monitored, controlled and adjusted by systems based on reduced physical models while taking the manufacturing history into account?

The research question addresses basically two areas. The creation of reduced physical models and in progressive steps the set-up, monitoring, as well as controlling and adjusting of a manufacturing system. The last part of the research question draws the production history, which contains information about relevant upstream manufacturing processes, into the design frame and optimisation frame. The necessary cognition of the technical system must be built such that all parameters which influence the processing result can be determined during processing, and the manufacturing system can make the required changes in the process behaviour in an adequate manner. What is principally needed for this, is that all relevant information about the current and the upstream manufacturing steps can be made available in a usable way.

As a first approach, the ability to model manufacturing processes is a significant task. Only once setting parameters and target values are combined in causal relationships and formally linked, the process can be modelled.

The following *research hypothesis* is derived from the research question: It is postulated that "Model-Based Self-Optimisation", consisting of a process model focused on product quality, and the necessary cognition, will be able to run a manufacturing process autonomously at the technical limits of the process domain, in such a way that consistent quality is achieved through all disturbances.

To verify the research hypothesis the following topics need to be studied:

- *Modelling:* The transfer of process knowledge into machine-readable models for product quality oriented process control and process adjustments
- *Cognition:* The development of suitable cognitive components for technical systems to determine the status of a process
- *Control and adjustment:* The control and adjustment of highly dynamic manufacturing systems, where the course of the process is unknown at the start of processing.

To create a generic solution in these areas, research will be carried out on five representative and practically relevant processes, while the model evaluation is viewed horizontally.

- *Gas metal arc welding* Welding and Joining Institute (ISF) of the RWTH Aachen University (ISF) with a manufacturing system from the industry partner CLOOS.
- *5-axis milling* Laboratory for Machine Tools and Production Engineering of RWTH Aachen University, Chair of Manufacturing Technology (WZL-TF) with a manufacturing system from the industry partner MAZAK
- *Laser beam cutting* Fraunhofer Institute for Laser Technology (ILT), Chair of Laser Technology of RWTH Aachen University (LLT)/Department Nonlinear Dynamics of Laser Processing (NLD) of RWTH Aachen University (NLD) with a manufacturing system from the industry partner TRUMPF
- *Plastic injection moulding* Institute of Plastics Processing at RWTH Aachen (IKV) with a manufacturing system from the industry partner ARBURG

- *Weaving* Institut für Textiltechnik der RWTH Aachen University (ITA) with a manufacturing system from the industry partner PICANOL
- *Model evaluation* Laboratory for Machine Tools and Production Engineering of RWTH Aachen University, Chair of Metrology and Quality Management (WZL-MQ)

6.4.3.1 Research Questions for Partners

Model Evaluation (WZL-MQ) Meta-models in the current context represent data records of measured values in the form of mathematical functions by means of interpolation and approximation. The creation of the models includes a series of steps which have an influence on the quality of the predictions. The choice of the setting parameters considered, the data sources, as well as the design of the model structure require apart from expert knowledge often also manual selection to achieve a good representation of the process by the model. The approach consists of visualising and evaluating the model quality and deriving strategies for efficient model creation, taking the influencing parameters into account. The derived question is as follows:

> How can surrogate models for diverse processes be efficiently created and evaluated with respect to the model quality, taking into account the influencing variables of model purpose, process robustness and complexity, choice of parameters and their density, spread and distribution?

Milling (WZL-TF) In 5-axis milling operations of free-form surfaces, the cutting conditions and therefore the process behaviour changes along the processing path. This results in a constant variation of the control-path behaviour of the production system itself, which prevents the implementation of a standard control system. The approach is the development of a model-based control. If it can be achieved to describe the control path behaviour in relation to the respective cutting conditions, then a stable process control can be realised that additionally includes the current cutting condition. This would be a substantial step on the path towards the overall goal of self-optimisation for a 5-axis milling process. From this context, the following research question for the 5-axis milling process is derived:

> Is it possible to realise a model-based, self-optimising control of the 5-axis milling process, that takes the current time- and position-dependent cutting conditions into account?

Welding (ISF) Automating welding systems that use the gas metal arc welding process (GMAW), often results in problems which negatively affect the required weld quality due to changes in process boundary conditions. The main causes of these problems are tolerances in component size, tolerances in positioning of components, as well as tolerances in previous finishing steps and tolerances due to thermal distortion caused by the welding process itself. Monitoring the welding process and applying an automatic correction of the welding parameters in case of deviations is often necessary to ensure the required weld quality. In the past, these problems have been approached in various ways, but these were always application specific solutions.

These required a time consuming collection of empiric data records by a welding expert, both for the initial set-up process and for carrying out the welding experiments. Therefore, the following research questions are posited for GMAW-welding:

> What optimisation methodology allows the generation of application-independent and model-based correction parameters, without having to carry out time-consuming and cost intensive welding tests?
> How, on the basis of information supplied by sensors during the welding process, can the current weld seam quality be determined using a model? What further requirements must be met by self-optimisation strategies in order to create optimum correction parameters based on the required weld seam quality characteristics?

Plastics Injection Moulding (IKV) For plastic injection moulding the product quality from a business point of view is determined by the resources required for set-up, and the cycle time, while from a manufacturing technology point of view, the weight and dimensions of the finished part are the major determinants of quality. The self-optimisation approach for plastic injection moulding consists of adapting the planned cavity pressure course, using the boundary conditions captured by metrology, such as for example the temperature in the cavity. The pvT-behaviour of the plastic can be used as a quality model, in order to enable the self-optimisation to compensate disturbances, such as temperature fluctuations. The research question therefore is:

> How can a self-optimising system be developed, built and applied, which is capable of ensuring constant product quality during injection moulding, by means of direct control of the cavity pressure, taking into account the pvT-behaviour?

Laser Beam Cutting (ILT/LLT/NLD) The laser beam cutting process is being investigated with respect to self-optimisation of steel sheet cutting, particularly focussing on stainless steel alloys. As the composition of the alloys itself has a major impact on product quality, the modelling activities in this area are supported by another sub-project (see Sect. 4.4) addressing metallurgical investigations. The quality of a laser cut product is determined economically by the time for set-up of the manfuacturing system, as well as the time and resources needed for the processing itself. In particular, the set-up of both the machine and the process for manufacturing of individual parts using special materials, have a significant impact on the economic efficiency of manufacturing. On the manufacturing technology side, the roughness of the cut edge and the amount of dross are the factors which determine quality.

For the laser beam cutting process there are challenges in the area of sensors and control, as well as process modelling. Measuring the values which characterise the process is currently an unsolved problem, due to the inaccessibility of the cut front. Due to the highly dynamic nature of the process, the question of how to execute control embraces the selection of setting parameters as well as suitable strategies to influence them a short time scale. Therefrom, three central questions arise for the laser cutting process:

> How can the laser cutting process be described such that machine-readable models for self-optimisation can be derived? How can process variables be measured which allow deducing product quality during laser beam cutting, so enabling self-optimising control of the laser

cutting machine? How can a laser cutting machine be controlled so that the cutting process itself can be optimised?

Weaving (ITA) The number of possible settings on a loom is very high. The individual components of a loom which are adjustable cannot be considered in isolation from each other. In most cases there is a complex interaction of the individual components.

The result of weaving is a product, the fabric, whose quality needs to be measured and evaluated. A direct transfer of the target value, product quality, to the multitude of setting parameters can currently not be expressed in closed form. The need for high-quality products and high-quality manufacturing in high-wage countries requires exactly this closed-form expression, and above all the feedback of target oriented setting parameters into the machine. A model-based approach is required to capture multi-parameter systems, to intelligently process and directly align to the target values. For the weaving process thus the following question exists:

Is a multi-parameter system with the target product quality able to autonomously set multiple parameters on-line, by using a model-based evaluation algorithm in order to enable sustainable, intelligent production in high-wage countries?

6.4.4 Results

Increase in productivity entails not only the reduction of processing time, but also considerable challanges. In materials processing, higher performance of tools and faster drive technology with higher precision create new requirements for process control. If a process is to be run at its technical limits, then to ensure its stability it must be robust against disturbances. Static setting parameters meet this requirement through holding off the stability limit by experimentally determined safety margins. This limitation can be overcome if the manufacturing system is able to recognise critical operating areas within the process, and if a multiplicity of parameters can then be adjusted as needed within a very short time-frame. The prerequisite for this are technical systems which can determine their own operating point and can adjust their behaviour on-line. These are the fundamental characteristics of Model-Based Self-Optimisation.

Model-Based Self-Optimisation for manufacturing systems offers methods and strategies which can be used to run highly complex and highly dynamic manufacturing processes safely at their performance limits. In this area of design, the development and qualification of cognitive entities for the determination of operating points, and the application of reduced models to predict product quality is a major challenge. The high-level methodology for creating the model and the design for self-optimisation are stated generically. The process models developed, the cognitive components created, and the set-up assistants implemented demonstrate the approach to Model-Based Self-Optimisation.

The concept of self-optimisation in the current context is defined as: *Self-optimisation of a technical system consists of its adaptation to changing process* *inputs and boundary conditions based on embedded expert knowledge and direct process information without external intervention, so that the required output values are achieved optimally.*

The realisation of this concept is divided into two functional areas, the implementation of a system to optimise the process to match externally specified target values, and a system to transform internal objectives into control of setting parameters. Research in this area can be summarised under three main headings:

- Execution of Model-Based Self-Optimisation for manufacturing systems.
- Cognitive functions of technical systems for self-optimisation.
- Description of processes by models.

6.4.4.1 Model-based Self-optimisation for Manufacturing Systems

In product design customer needs are defined as functional and economic requirements for an end product. As part of the construction process these give rise to technical demands on components, which are each produced by a manufacturing system. Manufacturing systems fulfil the task of taking an input product and creating from it the requested output product.

In the simplest case, the required values of setting parameters for the processing task are set by an operator based on his knowledge and experience. Today, the values are taken from a technology table which is defined empirically from a large number of tests. These values define an operating point which generally offers adequate stability reserves against undetected and uncompensated disturbances. The actuators of the manufacturing system actualise the values of the setting parameters, while sensors check the work of the actuators and the result. In the event of a deviation between the predefined value of a setting parameter and the actual value, the control deviation is minimised using fixed classic controllers. This functionality can be described as a sensor-actuator (SA) system.

If all disturbances remain within the limits which also underlie the technology tables, the likelihood of achieving the product quality is high. In those cases where the scale of a disturbance exceeds the predefined limit, or if the input product does not meet the required characteristics, then it is likely that product quality will not be achieved. This needs to be precluded for mass production, and in the manufacturing of small batches or manufacturing of individual pieces it can lead to significant economic loss and delivery delays.

Figure 6.52 shows the principle of Model-Based Self-Optimisation for manufacturing processes. From suitable sensor signals, a meta model is used to establish the current operating point which is being compared to the operating point that would apply to optimally achieve the defined external objectives. From this, the optimisation potential is determined. This optimisation potential is used by the model-based optimisation system (MO-system) to autonomously generate a solution in the form of new internal objectives. The SA system is extended to include the ability to handle these internal objectives so that it becomes an information-processing sensor-actuator system

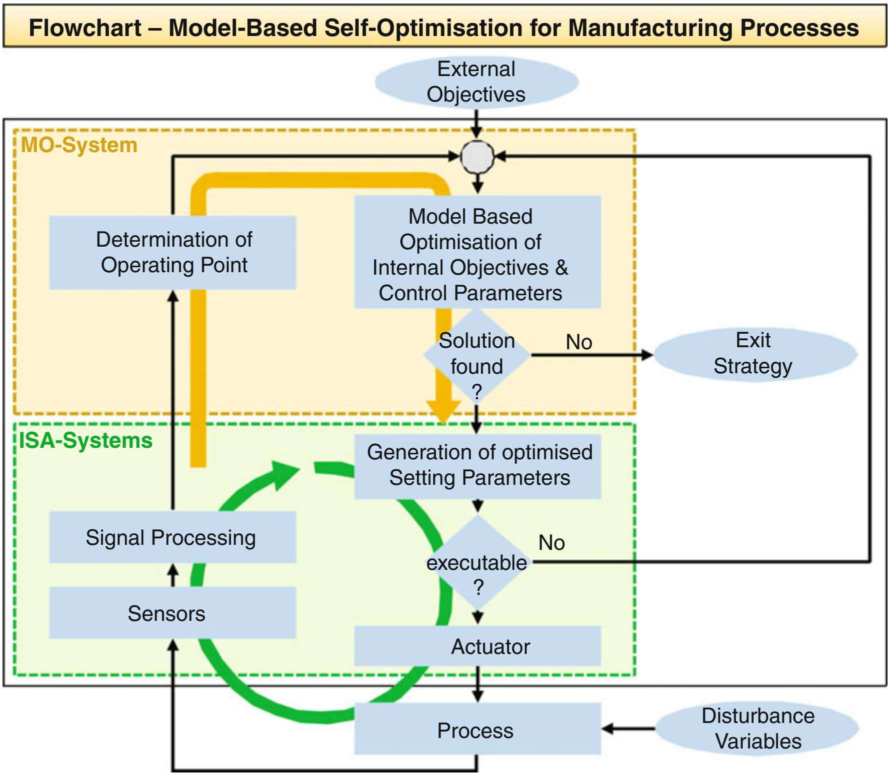


Fig. 6.52 Schematic presentation of the sequence of Model-Based Self-Optimisation for manufacturing processes

(ISA system). As well as translating the internal objectives into setting parameters, it also controls them and can be reconfigured in real-time during the process.

External Objectives of Manufacturing Systems The fitness for purpose of a product is achieved by matching the predefined characteristics for product quality (PQ). In most cases there are several quality characteristics to be taken into account. For example the product must meet predefined dimensions, have a certain level of roughness on its surfaces, or must meet requirements relating to rigidity and structure. This also applies to buisness requirements such as machining time and manufacturing costs. If resource efficiency is handled separately, then this parameter must also be included in the optimisation.

Process Variables Product quality generally correlates with a number of process variables. Model-based self-optimisation requires sufficient data, both in quality and distribution, about these variables, in order to be able to describe the conditions at the operating point with sufficient precision. Often suitable sensors are not available, or there is no access to the source of information. In many cases it is therefore necessary to develop new sensors, or to adapt existing systems to the current task.

An important process variable in the weaving process is the warp tension, whose temporal behaviour has a significant influence on product quality. To measure the warp tension with a high temporal resolution a new sensor has been developed which does not add any measurable friction to the process.

When finish milling thin-walled components the product quality is mainly defined by the surface quality and dimensional accuracy. In order to keep both targets in predescribed tolerances, the values of the significant process variables must also be kept within predefined limits. Here, the main variables are the behaviour of the component during processing, the tool used, the machine behaviour and the process with the significant setting parameters cutting speed and feed rate. Latter can be read directly from the machine control. This information is not sufficient for Model-Based Self-Optimisation, why further sensors have to measure additional process variables which are being integrated into the manufacturing system to enable self-optimisation. In the present example, further process variables which are relevant for quality are the vibration of the component and of the tool. It has been shown that neither force sensors, whose eigen frequency is too low, nor acoustic emission sensors, whose frequency range lies outside the observed frequency range, are applicable. Besides this criterion, other requirements need to be met by the sensor system, such as, for example, a sufficiently high signal-to-noise ratio, robustness against influence of coolant and material removal, as well as being independent of the mounting position. As tactile and optical sensors do not meet these boundary conditions, a multi-sensor system was developed. Two acceleration sensors were attached to the machining table and spindle housing, as well as an eddy current sensor facing the tool. The multi-sensor system delivers consistent information about the vibration behaviour of thin-walled components during finish milling. By using signal processing algorithms the measurement signals are converted into values for quality related process variables.

A further example is the selection of sensors for laser beam cutting. Most of the quality related process variables exist in the interaction zone of the laser beam and the material. Due to the high temperature of the molten material and the small width of the cut, of just a few hundred micrometres, it is difficult to acquire measurands to provide information about the roughness of the material at the solidified cutting surface. The process model however allows to derive the correlation between process measurands and process variables. Using the modeled relationship between the width of the cut and the focal position, the technical system determines the values of the inaccessible but quality relevant parameter, focal position, by measuring the width of the cut. In laser beam cutting an imaging sensor system observes the processing zone coaxially through the laser optics of the cutting head to acquire information from the interaction zone which has a causal relationship to product quality. The signal processing analyses the acquired sensor information and passes the data on for self-optimisation.

Determining the Operating Point: The operating point of a manufacturing process is given by the current values of the process variables. If measurands cannot be determined directly, then the dependencies of process variables on measurable variables must be known and described by models. In process models the functional relationships between setting parameters, process variables and process measurands are defined. Meta-models describe these relationships in a reduced mathematical form, and thus allow the determination of the current operating point during the manufacturing process.

In welding, the model to depict process variables and setting parameters with respect to product quality, for example the weld geometry, is also built up on fundamental physical relationships. The values of the dynamic process variables voltage, current and gap-width are determined by sensors. Other setting parameters, such as shielding gas quantity, which influences the shape and boundaries of the working areas, are predefined by the manufacturing task. The necessary calculations to determine the operating point are carried out numerically in real-time. This enables defining and tracking the operating point during the process. The process domain shown in Fig. 6.53 defines the parameter limits where the process has the same physical properties. Within this area the operating domain represents all values which are capable of delivering the pre-defined product quality.

Model-Based Optimisation The optimisation of the manufacturing process takes place in a number of steps. These include the determination of the process status at the current operating point, and the model-based assessment in relation to the external objectives. The externally defined objectives such as, for example, machining time, manufacturing costs, resource efficiency or surface roughness, dimensions, weight must be correlated to significant process variables in order to be transformed to internal objectives. Achieving the internal objectives is done by adjusting the setting parameters.

The models of the manufacturing processes enable prediction of the product quality characteristics from the setting parameters. The predictive use of the models

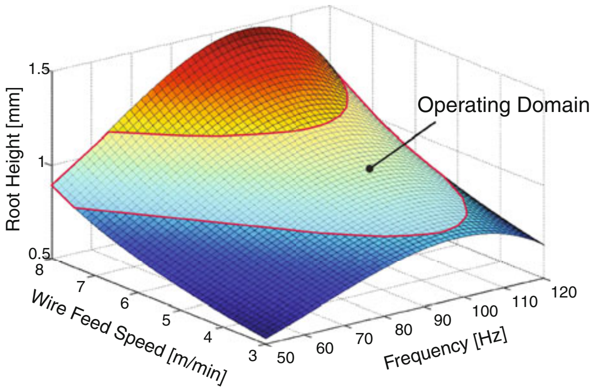


Fig. 6.53 Process domain of a gas metal arc welding process

supplies an unambiguous solution. If the models were bijective, then the inversion could be carried out using a mathematical operation. Because of the complexity of the processes, and the dimensionality of the parameter ranges, process models are in fact mostly not bijective. To determine the values for the setting parameters for a given product quality using a non-bijective model, an inverse solver is needed, which usually uses iterative methods.

A suitable solution generally includes meeting various objectives. Usually individual features of product quality, such as surface roughness, geometrical properties and process efficiency are competing with each other (Pareto optimisation). In addition, boundary conditions such as machine performance limits or process limits must be taken into account. For manufacturing processes, the dependencies between process variables which are part of the optimisation task and the process outcome are strongly non-linear. These types of problems are described as multi-criteria optimisation with boundary conditions. One approach for solving problems like these is transformation to a scalar optimisation function, with weightings for each of the targets and optionally the formulation of boundary conditions. The optimisation task that results from this can be solved with non-linear programming methods, such as gradient methods or Newton's method. Depending on the precision required, solutions can be created which are suitable for use in self-optimisation with regard to their calculation time.

In the context of Model-Based Self-Optimisation for manufacturing systems, there are two central goals. The first comprises achieving product quality by applying suitable operating points. If all operating points which deliver a required product quality are collected into the operating domain, then the minimum such domain is the optimum operating point.

The second target comprises the stability of the operating point. Meta-models are able to carry out a continuous calculation of target values within the parameter space. This enables optimisation solutions relating to stability to be assessed, by taking into account information about gradients and discontinuities.

Even if all operating points within the operating domain achieve the defined quality, some operating points will lie closer to instabilities than others. To evaluate robustness against disturbances, the area surrounding the operating point is also tested for instabilities. With this information, adjustments of the setting parameter values can be achieved so that the stability zone is increased.

Reaching an optimised operating point may require significant adjustments of setting parameters, which in turn can impact the stability of the process. In addition, because of the dynamic response of the process the adjustment can often not be made within one step of the control cycle, which makes it necessary to calculate intermediate steps. All the points on the trajectory, which are defined with the help of the meta-model, must work in a way that guarantees continuous quality and a sufficiently large operating domain to ensure robustness against disturbances.

The following condition could be interpreted as a safety risk or stability problem: If the system finds a operating point which meets the external objectives, but is in a operating domain which is not connected to the current operating domain, then no trajectory can be calculated for the transition. Therefore, this case must be handled separately, depending on the process and the application. But it is still possible that depending on the size of the new operating domain and the dynamic answer to the changes in values of the setting parameters, that this operating point is indeed the better choice for achieving the external objectives. The consideration of such decisions can be handled by adding additional terms into the target function for optimisation, or it can be solved with special decision algorithms. The result of optimisation is a new set of values for the internal objectives combined with sensitivities and validity areas.

Information-Processing Sensor-Actuator Systems Model-based optimisation, which is represented in the upper dashed box in Fig. 6.54 (MO-system), generates internal objectives and control parameters, which are used by the manufacturing system to achieve the external objectives. This information package is passed on to the system, which contains an information processing unit, a sensor and an actuator (and ISA system). The information-processing unit contains a static part with electronic interfaces to the sensor and the actuator as well as a variable part with programmable logic to reconfigure the control behaviour.

In a simple case, the information package handed over consists of a control value as an internal objective and the description of a simple proportional control element as a control strategy. If the relationships between the setting parameters and the process variables are non-linear, and if there are dynamic reactions to be taken into account, then the generation of the internal objectives and their control strategy becomes more complex. In addition, the information processing unit must even be able to be reconfigured at run-time if the internal objectives each depend on

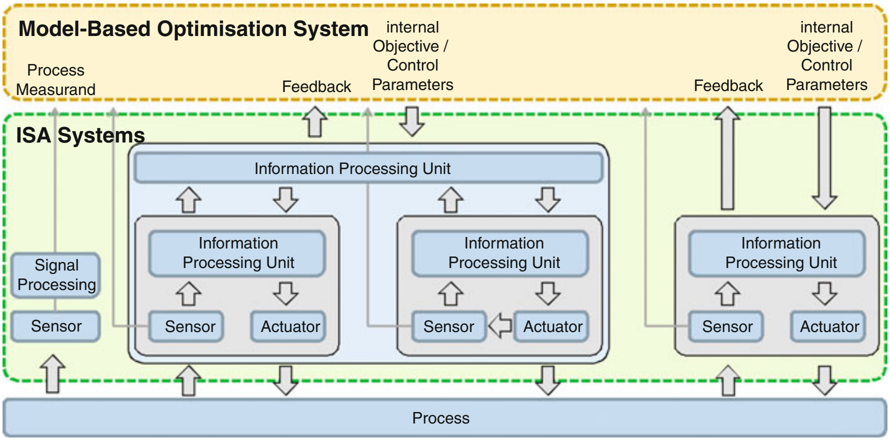


Fig. 6.54 ISA systems as a platform for controlling manufacturing systems with highly dynamic processes in variable configurations

the current operating point. On this point, the model-based approach offers precisely the possibility of generating dedicated meta-models which can be evaluated in the information processing unit in real time. This split between generating internal objectives and controlling the actuators introduces an abstraction layer between optimisation and control, which also enables decoupling the individual problems timewise.

The demands on the information processing unit increase with the complexity of the modelled relationships between process measurands and values for the setting parameters. For example, in weaving, the density of the fabric is an internal objective. Optimised values for the internal objective are transferred to the subordinate system and there translated into warp let-off and fabric take-off.

In welding, the external objective of the energy per unit length can be controlled via a system which consists of three ISA systems, combined as shown in Fig. 6.54. To meet the metallurgical demands of the product, the energy introduced into the product has to meet certain values over time and distance. Conversion of the external objective to internal objectives takes the form of electrical current and voltage values as the welding robot moves along its defined path. In this case, the values for welding current and voltage are managed by a superordinate information processing unit as an ISA system which controls the resulting arc power. This system receives instructions and constrains for the energy per unit length, as well as information about the pulse characteristics, like peak current or peak time, as well as pulse frequency. The ISA system for the arc power is parameterised by the characteristics of the welding power source, the welding current and voltage are measured continuously, while the ISA system for the robot controls the movement.

For high performance laser beam cutting, maintaining a given focal position is important, as this value has a significant impact on the quality of the cut. The control of the focal position is a very complex task as its position varies with the power of the beam and the diameter of the raw beam. In addition, the behaviour changes over time because of the optical system warming up. While processing, the focal position cannot be measured directly, but it has a strong correlation to the width of the cut. Therefore, a meta-model provides the width of the cut as an internal objective which needs to be met, in order to achieve the external objective. As this relationship is neither linear nor unambiguous, the information for the ISA system has to be extended to include control parameters and sensitivities before it can be transmitted. Using this information, the ISA system can pursue the internal objective. The coaxially mounted, imaging sensor observes the processing zone and determines the width of the cut using algorithms for image processing and signal analysis. This process measurand is used by the information-processing unit to activate the focussing unit in closed loop control. The transmission of the internal objective, the width of the cut, together with the control parameters such as dynamics and sensitivity, allow the ISA system to complete the task on a very short time-scale.

The generation of optimum values for the setting parameters in the injection moulding process, aims to achieve consistent geometric properties and weight of the moulded parts. The correlation between pressure, temperature and the specific volume of the plastic melt is used to calculate the optimum course of the pressure in the cavity, using the help of a model predictive controller. On the basis of a process model, which describes the process behaviour of the manufacturing system, the control input (control voltage in the servo-inverter or proportional vent) is compared to the controlled variables. The model predictive controller uses this model to predict the course of the pressure in the cavity, given the prescribed course for the control signal (cavity pressure course). The control signal is adjusted using an optimisation function, which is based on this model prediction. The injection moulding machine, the type of plastic granules, and the geometry of the cavity together define the system to be controlled. If one of these components changes then the process model is adapted accordingly to the new configuration.

The design of the ISA systems establishes an abstraction layer between optimisation and control, which decouples both areas in terms of time and content. The information-processing sensor-actuator systems (ISA systems) are able to track their internal objectives at a high control frequency while the model-based optimisation system (MO-system) monitors the existence of optimisation potential. Subsequently, new internal objectives are generated and applied to the manufacturing system using parameterisation or reconfiguration of the ISA systems.

6.4.4.2 Meta-modelling

Modelling Manufacturing Processes Manufacturing systems use a technical process to process materials, for example by joining, cutting or forming. If a physical process model exists for a manufacturing process, then the values for the setting parameters can be used to predict product quality. If such a model is also available in machine-readable form, then it can be used as the basis for self-optimisation. Models of this kind, often described as white box models, cannot be evaluated automatically and sufficiently fast in most cases. This means that, although they can be used to generate initial values for setting parameters, they cannot be used for process control.

An alternative approach is the use of meta-models, which in this context are seen as a reduced mathematical representation of the functional relationships between multiple input parameters and output parameters. These models are based on process knowledge which is gained from either experimental or simulated data and which describes the process in sufficient detail. Applying meta-models to the solution of a dedicated task allows the reduction of the complete system analysis to a lean numeric solution that can be automatically evaluated which enables control and adjustment of the manufacturing process. Possible implementations of meta-models are grey box models, where for example, the basic functions or the model structure are defined in advance, based on process knowledge. Black box models do not contain any process-specific information and only use data-based adjustment or training to adapt themselves to the behaviour of the process (Stork et al. 2007; Buhmann 2003).

The use of meta-models can refine or even generate process knowledge. Through the closed-form representation of the data, a direct comparison between experiment and theory is enabled, which allows examination of the sensitivity of setting parameters (Schüttler et al. 2009). A general disadvantage of approximation methods is their poor ability to describe discontinuities within a process domain. Hence, sudden changes in the behaviour of a process, as occur at the edge of process domains, usually have to be considered separately.

The creation of meta-models for manufacturing processes comprises several tasks that determine the quality of the model. Significant factors comprise the selection of the model structure, adjusting the model parameters to the data set and handling of discontinuities. Solutions of this task are supported by a process-independent methodology, which was developed as part of this study on model generation.

Methodology for the Development of Metamodels The methodology for creating models was developed as part of the project and evaluated against five different manufacturing processes. It includes, alongside the general approach, the necessary resources and tools as well as steps for optimising the models.

In Fig. 6.55 the column headed "B" describes the systematic process for creating a model. The first steps (*B1*) *Definition of process and model requirements*, (*B2*) *Selection of parameters*, (*B3*) *Determination of the process domain* and (*B4*) *Selection of data source* define the requirements and limits of the model. The second group of activities (*B5*) *Planning and setup of experiments* and (*B6*) *Generation of data* prepare the data for the calculations in the model. The last activities are for

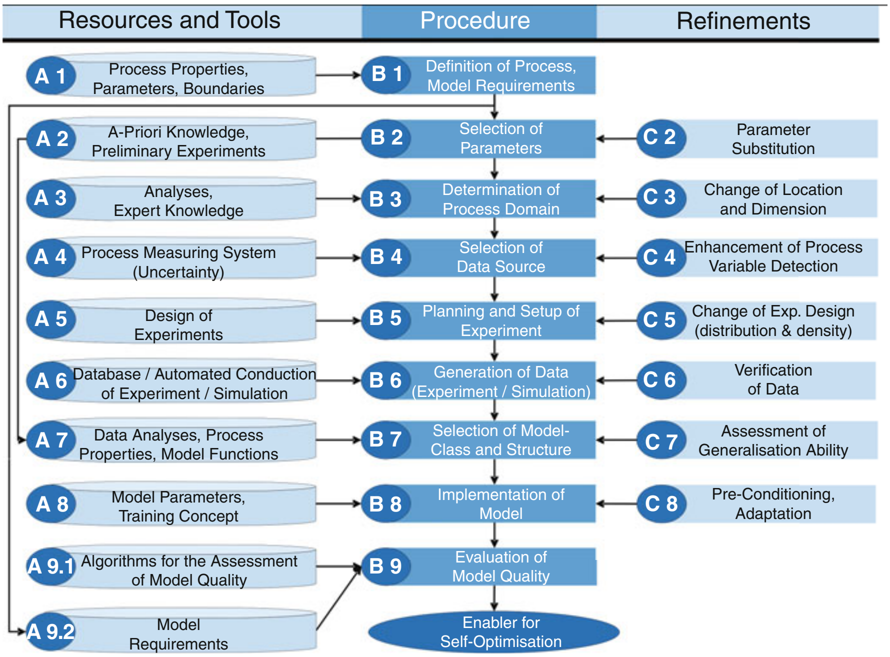


Fig. 6.55 Methodology for creating a meta model

the development of the model itself. Therefore, in (*B7*) *Selection of model class and structure* a suitable model type is selected, the model is created based on the acquired data in (*B8*) *Implementation of model* and finally the models are evaluated in (*B9*) *Evaluation of model quality*. The necessary supporting resources and tools are indicated with "A" and are assigned to the appropriate processes of creating the model. Since many different factors influence the overall model creation process, the required model quality is not always achieved after the first iteration of the model creation process. For this reason the method offers measures to optimise the model quality, which are shown with a "C" in Fig. 6.55.

Definition of Process and Model Requirements The generation of meta-models starts with the analysis of the process to be modelled. Input, output and influencing variables which affect product quality need to be identified and defined. In addition, on the basis of prior knowledge about the process, decisions are made, for example whether to use a dynamic or a static model, linear or non-linear correlations. The necessary requirements on the model, such as the maximum permitted computation time, the necessary model quality, or the differentiability of model functions must be taken into account. Adequate definition of the boundary conditions has a significant impact on achieving the required model quality, and minimises the number of necessary iterative steps right from the start.

Selection of Parameters The manufacturing process and the resulting product quality are influenced by multiple parameters, such as properties of the materials, machine settings and the prevailing environmental conditions. Existing process knowledge can be used to identify important parameters, or these can be identified by preliminary tests. Significances and correlations can be further detailed by experiments using design of experiments.

Determination of Process Domain After the selection of the parameters to be taken into account in the model their ranges of values are defined. This covers the process domain, the range of setting parameters and product quality in which the process has the same physical properties. A subrange of the process domain is the operating domain. This comprises all operating points at which the required product quality will be achieved, and does not necessarily have to be one continuous area. The self-optimisation of the manufacturing process therefore occurs within the overall process domain, taking into account the stability at the edges of the various operating domains.

Selection of Data Source Once the relevant parameters and their value ranges have been defined the data for the generation of the meta-model need to be acquired. In relation to the required model quality and the sensitivity of the setting parameters with respect to product quality, it is important to minimise noise, deviation and uncertainties in the input data and either adjust for or exclude the impact of disturbances. The selection and application of sensors, together with the conduction of experiments, have a significant influence on the quality of the model (Witt 2007). Particular attention must be paid here to the signal-to-noise ratio and the choice of a suitable time resolution for the acquisition of measured values and their average. In all cases, it must be taken into account that the sensors themselves have no significant influence on the process. For example, contactless distance sensors are suitable for measuring vibrations (Wibbeler 2002). The same considerations apply for data which is generated with the help of numeric simulations. Here, sensitivity and precision can usually be determined or estimated in advance.

Planning and Setup of Experiment If surrogate models are based on interpolation, then their quality depends substantially on the density and the distribution of the data. For weakly non-linear parameter correlations a suitable balance can be found between data density and model quality by having well distributed data points. This can be achieved using design of experiments. But in many areas, the modelling of highly dynamic and strongly non-linear manufacturing processes requires an individual distribution of data points within the parameter space in order to get reliable model predictions. A method which is often used to generate equal distributions is the Latin Hypercube Sampling (LHS) (Steinberg and Lin 2006). For an LHS-distribution first of all, the number of levels per parameter and the parameter range are defined and then, the parameter space is divided up using a hyper-grid into the corresponding levels. Then, test points are inserted so that exactly one point per row exists within each dimension. A more equal distribution in the upper dimensional space is provided by the Centroidal Voronoi Tessellation Method (CVT) (Romero et al. 2006). For this method, the parameter space is divided up around the randomly distributed initial data points into regions, such that each region contains exactly one data point. In addition, the distance between each data point and the centroid within a region is shorter than the distance to the centroid of any other region. A subsequent alignment of the data points provides that each of them is exactly the centroid of their region. The Latinised Centroidal Voronoi Tessellation Method (LCVT) is a combination of the two methods described above and of their advantages and was used successfully during the project to generate test points (Reisgen et al. 2010).

Generation of Data After defining the test plan, the necessary data for producing the meta-model is generated. Right from the beginning, the automated acquisition of measured data has to be a goal to minimise the manual experimental effort. The same applies for data that is generated using simulation. With the initial design of the simulation, interfaces should be implemented which support later parameterisation. In both cases, datasets are available after generation, from which the meta-model can be built up in order to predict the process outcome in a much shorter time or even in real-time.

Selection of Model Class and Structure Meta-models are defined by their class and structure. White box models in the current context describe the physical relationships of a manufacturing process using mathematical formulae, with the aim of providing an exact description of the real processes. While the precision of this class of model is very high, due to the use of very detailed process knowledge, the time required for implementation is also high, and its complexity often is associated with a high demand for computational resources. One approach to manage this problem is the use of reduced physical models that decrease the computational effort needed for a sufficiently precise solution, for example by applying an asymptotic expansion (Schulz et al. 2009). Grey box models combine process knowledge and adjustable structures. While the structure of the model is derived from the process knowledge, the coefficients are adjusted to the underlying data using a model fitting techniques such as linear regression with least squares method (Draper and Smith 1998). Other grey box implementations of this kind use symbolic regression with evolutionary algorithms. The most flexible and most general class of model generation is the black box model. They are adjusted to the underlying data records without predetermined structures based on process knowledge. One of the most widely used black box modelling tools is Artificial Neural Networks (ANNs) (Rojas 1996). ANNs consist of networked neurons which are defined as having inputs and outputs as well as an activating function. Several layers of such neurons can be used for the approximation of the underlying dataset, with the neurons being trained using the data. Another approach, which follows the paradigm of the black box model, is the use of Radial Basis Functions (RBF) (Jurecka 2007). These models carry out an interpolation into the underlying data records using a set of functions which are positioned at the sampling points and are weighted, and which are also scalable by means of a shape parameter. In areas which are highly dynamic and highly non-linear, the main risk is an under-adjustment. For self-optimising manufacturing processes, the importance of identifying and separately assessing the model quality in these areas is decisive. In grey box models this can be handled explicitly.

Implementation of Model After generating the data and choosing the type of model the meta-model is created. A series of application programmes such as SNNS (Zell et al. 1991), Matlab, DesParO (Stork et al. 2007) or SUMO (Gorrisen et al. 2009) offer modules to implement Radial Basis Functions, Artificial Neural Networks or a wide variety of regression models. When using Artificial Neural Networks, the main factor that influences the quality of the model is the number of training passes, when using Radial Basis Functions it is the definition of weightings and shape parameters. In many cases, there may be serious losses of model quality due to under-adjustments and over-adjustments. In addition, depending on the choice of model parameters and the number of underlying data points, the calculation time for creating the model can grow exponentially. In this case it needs to be checked whether a reduction of the model parameters or number of data records would compromise model quality.

Evaluation of Model Quality The model quality describes the usability of the created model. A suitable method is the determination of model error using cross-validation. In this, data records that were not used in creating the model are applied to validate the model by comparing them to the result of a prediction by the model. Various techniques of cross-validation are holdout, random sub-sampling, K-fold and leave-one-out (Efron and Gong 1983).

Classic error criteria are the mean absolute error or the mean square error. Neither of these take into account the global scaling of the model parameters, which means that after parameter transformation by pre-conditioning or scaling the relevant model error criteria are no longer comparable. The criterion of the coefficient of determination R^2, which defines the ratio between the mean square error and the standard deviation of the parameter is, for its part, not affected by scaling (Efron and Gong 1983). The values of R^2 range from 0 for a bad model to 1 for a "perfect" model and so enable comparison.

Models which describe the response to transitions between stationary process states require the application of a time horizon in order to evaluate their dynamic prediction quality. A k-step-ahead prediction (Ljung 1999; Nørgaard et al. 2000) allows looking at several discrete time steps, where each solution from the model is used as the input to the next iteration. By comparison of the prediction to a measured validation value, the predictive quality is obtained.

6.4.4.3 Results Model Evaluation (WZL-MQ)

The generated models for meta-modelling were tested for their model quality. The work on model evaluation contains additional research into the selection of a suitable evaluation criterion and the creation of software for model evaluation. For this, the AX-Workshop software was developed, which creates meta-models from the recorded process data, and reads in existing models in order to evaluate them.

Evaluation Criteria As first criterion to evaluate static models the criterion of the mean square error (MSE) was investigated. Starting from real output values y_i a comparison is made using the output value calculated by the model $\hat{y}_i$. The MSE is calculated by

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{6.2}$$

Over time, it appeared that the MSE-criterion only offers a relative comparison of models from the same manufacturing process. It offers neither a process-independent nor an absolute assessment of the model quality. Figure 6.56 shows as an example an evaluation of the laser cutting process using the MSE-criterion. It can be deduced that Artificial Neural Networks are best suited for the first two output parameters and reciprocal quadratic regression terms for the third output parameter.

To be able to make absolute statements about the quality of the model the suitability of the coefficient of determination R^2 was assessed. This is derived from the ratio of the square error

$$SS_{err} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{6.3}$$

and the absolute sum of the squares

$$SS_{err} = \sum_{i=1}^{n}(y_i - \overline{y})^2 \tag{6.4}$$

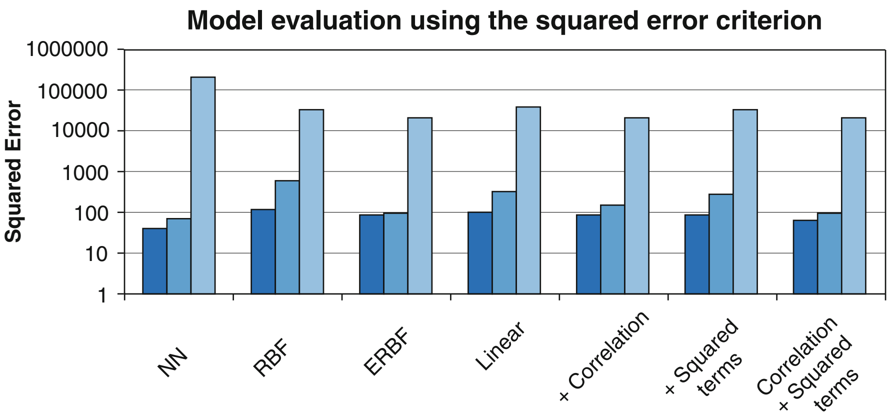


Fig. 6.56 Example of model evaluation on laser cutting using MSE-criterion

as a ratio to the mean value

$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i. \tag{6.5}$$

R^2 is then defined as follows (Fig. 6.57 shows an R^2-evaluation of static models):

$$R^2 = 1 - \frac{SS_{err}}{SS_{tot}} \tag{6.6}$$

Particularly Artificial Neural Networks and the Radial Basis Functions generated by DesParO achieved overall a good quality of model (Buhmann 2003). The question as to which minimum R^2 value is necessary in order to obtain satisfactory rigging, monitoring and control of manufacturing processes within the area of self-optimisation still remains open, however.

Investigation of Distribution and Pre-conditioning of Setting Parameters An important criterion that determines the quality is the number of data points used to create a stable and precise meta-model. Response Surface Designs (Circumscribed, Inscribed, Faced) distributed in the parameter space were applied to the selection of data from the manufacturing process for meta-modelling. By the gradual reduction of the number of data points for the creation of the models and for their evaluation, the model quality was analysed for its dependency on the density of data and the data distribution. Using the circumscribed and inscribed distribution a model containing 593 data points could be reduced to just 25 data points, with only a minor loss of model quality (Fig. 6.58).

It was also examined how far pre-conditioning or normalising the data used contributes to improved model quality. The input data was transformed such that the

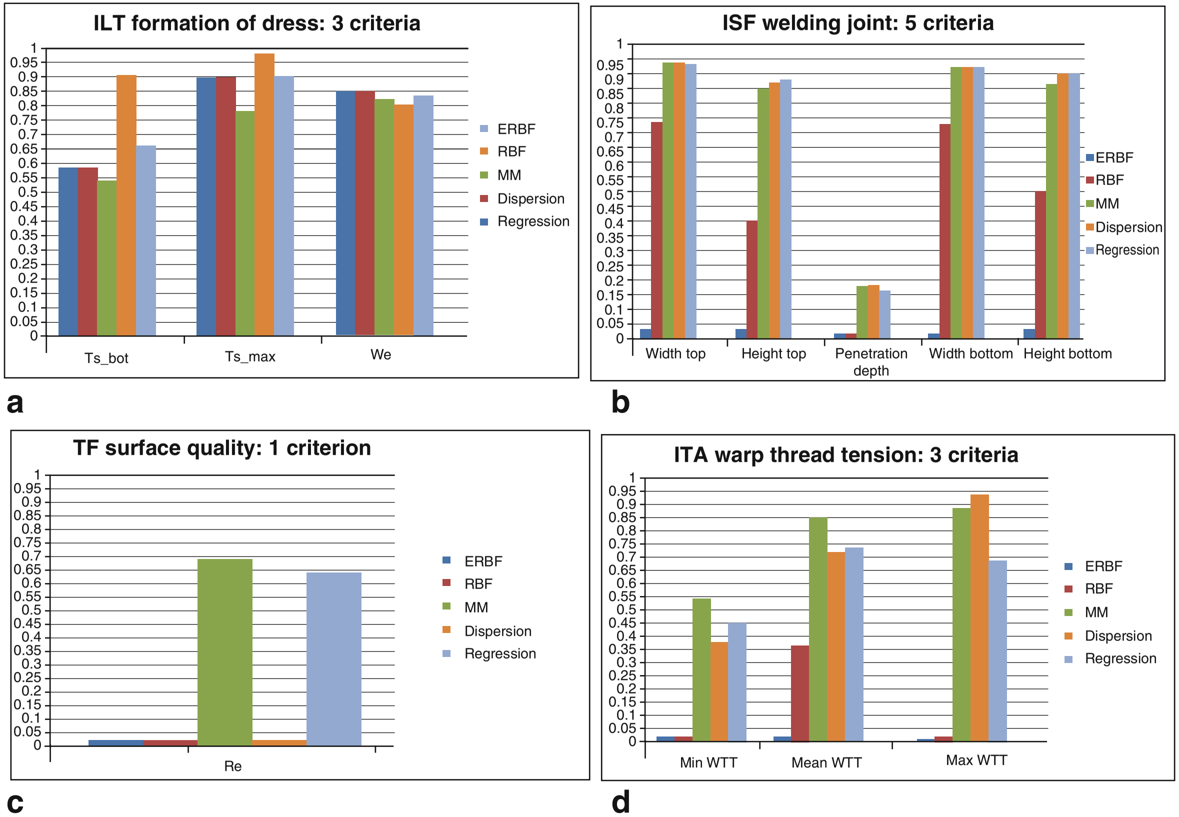


Fig. 6.57 R^2-evaluation of the models. **a** Laser cutting (ILT/LLT/NLD). **b** Welding (ISF). **c** Milling (WZL-TF). **d** Weaving (ITA)

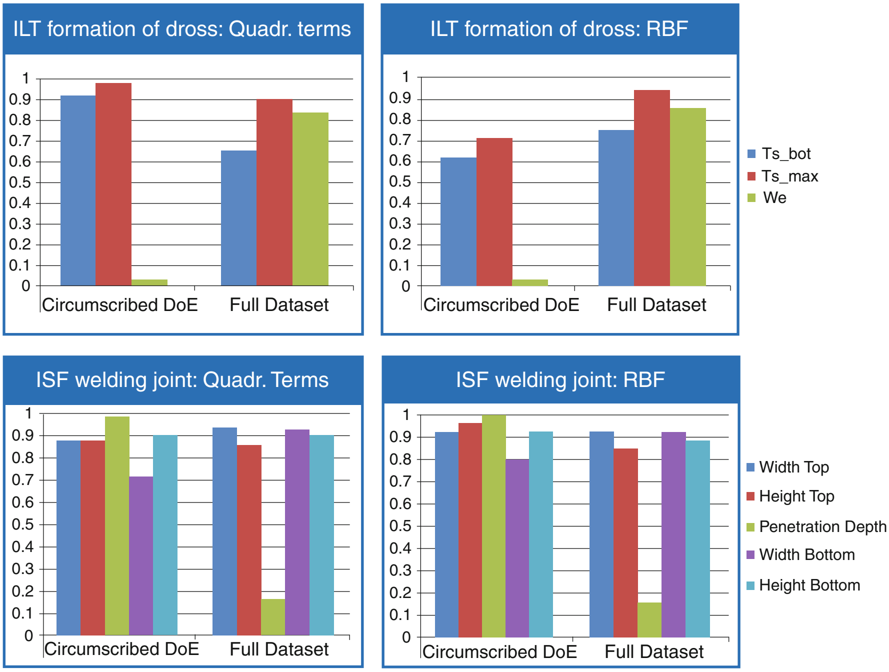
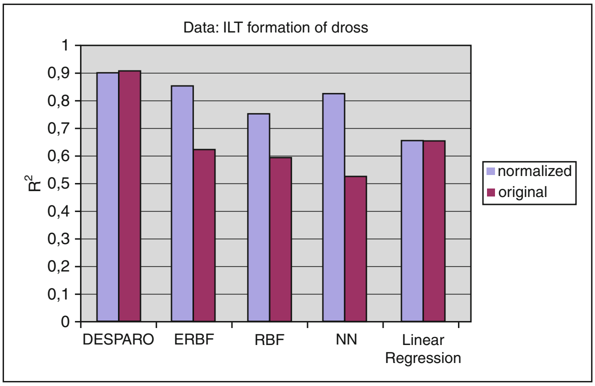


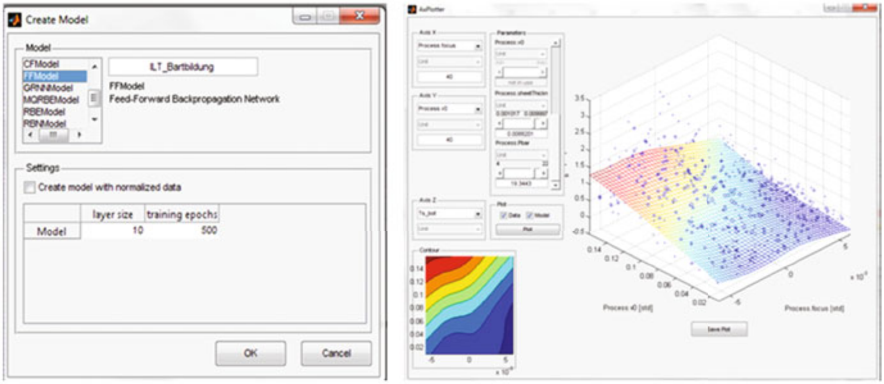
Fig. 6.58 Comparison of model quality of reduced test plans with full test scope

Fig. 6.59 Comparison of models with and without pre-treatment



mean value μ for each dimension was μ = 0. Using a principal component axes transformation the standard deviations σ were scaled to $\sigma = 1$ and correlations between the different dimensions were removed. This delivered a clear improvement in quality of those models based on neural networks (Fig. 6.59).

Software for Creating and Evaluating Models As part of the research, a software system for meta-modelling was developed and further enhanced (Fig. 6.60). It



a b

Fig. 6.60 Software for creation, visualisation and evaluation of meta-models. **a** User display. **b** Visualisation of the model

allows the efficient creation, visualisation and evaluation of different model structures. Radial Basis Functions, Neural Networks, regression functions, splines and other models are available as model structures. Alongside the visualisation, the models can be compared with the underlying data records, and can be evaluated using the implemented validation criteria (R^2, MSE).

A 20-fold cross-validation (Efron and Gong 1983) can be used as a good approximation for the global error using the coefficient of determination R^2. For conclusions about local errors in a meta-model, on the other hand, it is the Leave-One-Out cross-validation which is most useful, as each data point is verified (Efron and Gong 1983).

6.4.4.4 Results Milling Process (WZL-TF)

To ensure the reliability of a component's functionality and quality, from a manufacturing viewpoint the most important aspect in milling is keeping the geometrical tolerances. Here, dimensional and geometrical accuracy, roughness and surface quality were identified as the most important criteria of finishing operations. For roughing operations, the tool life and the removal rates are the most relevant quality measures.

For the analysis of dependencies between the setting parameter (SP), the process variables (PV) and the product quality (PQ) the three process measurands (PM) position, vibration and cutting force were selected. The selection of these process measurands is justified by the fact that the vibration behaviour of the part, and the mechanical load on the tool vary as a function of the position of cut. Based on this, there is a need for a position-oriented analysis and evaluation of the current operating point during the machining operation. This includes direct linking of sensor information to the position of the tool in the work piece coordinate system. Via this position a

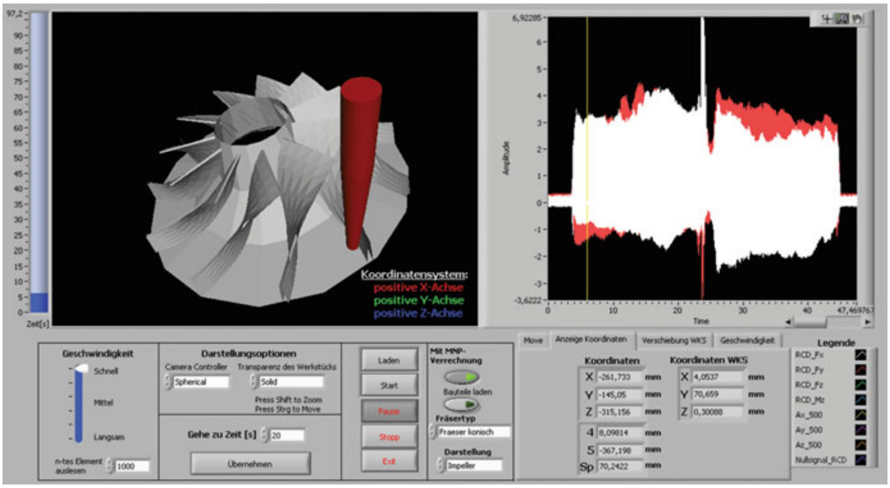


Fig. 6.61 Position-oriented monitoring during 5-axis milling

linkage of process and component parameters can be realised. This enables the analysis and evaluation of the process behaviour during changing engagement situations, as well as depending on the process disturbances and changes of material characteristics. Based on this information, suitable strategies can be derived for the design of a suitable control strategy. This approach can be applied to roughing and finishing operations. Due to this analysis, the whole machining sequence of 5-axis milling is structured more transparently and the general process understanding increases. The result of the position-oriented monitoring in 5-axis milling was demonstrated using a freeform geometry of an impeller and is shown in Fig. 6.61.

The appropriate use of a position-oriented monitoring system is based on the selection of suitable sensors to record the process measures. As an example of the finishing process, where according to Insperger the main causes of unsatisfactory process results are process instabilities, various sensors were tested in the laboratory for their useability (Insperger et al. 2003). Taking the position-oriented boundary conditions into account and the resultant requirements on the sensors, it was possible to demonstrate that the on the market available power measurement systems are not suitable for the investigation of finish milling operations. They offer too low eigen frequencies, so that the dominant process vibrations could not be resolved (Klocke et al. 2008). The measurement of the vibration behaviour with displacement sensors was done using only contactless operating sensor solutions. By doing so a direct linkage of the sensors to the machining operation system was established without affecting the system. Fibre optic sensors and eddy current sensors were used, both of which allowed the capture of dominant process vibrations. Although, looking for an industrial operable integration, both measurement systems are not suitable to be integrated in a manufacturing unit. In further investigations acceleration sensors were identified as suitable measurement systems for a position-oriented monitoring

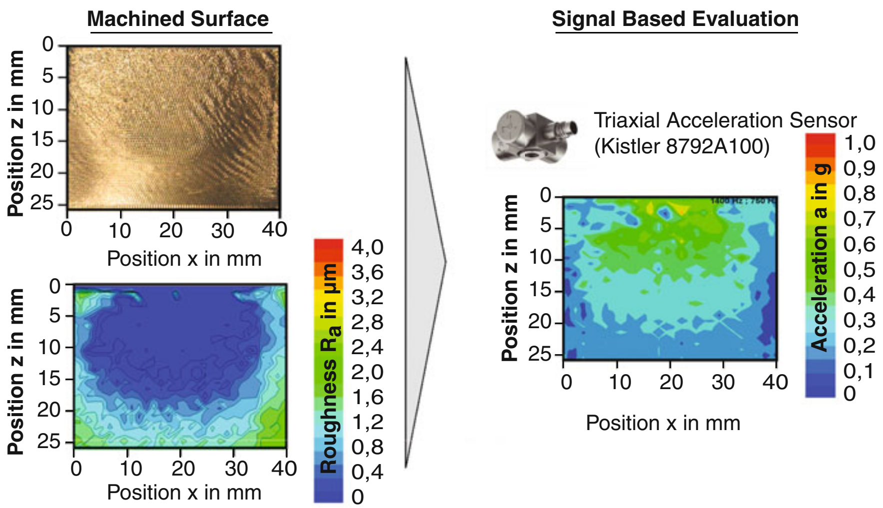


Fig. 6.62 Evaluation of finished surface based on acceleration signal

of the machining process stability. Regarding the finish milling process, the sensor signals were used to evaluate the surface quality manufactured by analysing the process stability. Hereby, the position-oriented analysis of the acceleration signals indicate to what extent the geometrical tolerance and surface quality meet the quality requirements, Fig. 6.62. In future work, the developed monitoring system will be used for rough milling operations, to enable the detection and control of instabilities in this area of the milling process too.

For the application of a model-based process control, solutions were developed for efficient model creation, to be able to provide the changing transfer behaviour in a machine-readable knowledge (Klocke et al. 2009). As the number of influencing factors during milling operations and therefore the time required to generate the data for a model creation, is very large, there was a need to develop a system which can be flexible configured. Furthermore, this system should allow an automatic evaluation of machining tests. In addition, the reproducibility of the test performance and data evaluation must be ensured, as this is the only way to generate reliable models. Knowing this, a system was designed and built which carries out automated machining tests, processes and analyses the signals received from the process-integrated measurement system (Klocke et al. 2010). Based on the analysis of the sensor signals, process-specific characteristics can be identified which correlate to the current engagement situation at any time, and are stored in a suitable database. Using this data, further work was done to develop process models to describe the transfer response behaviour in rough milling and finish milling operations and to integrate these into suitable process control concepts.

The basic models for a process behaviour description and those, which are to be developed for an application of process control, are a first step towards

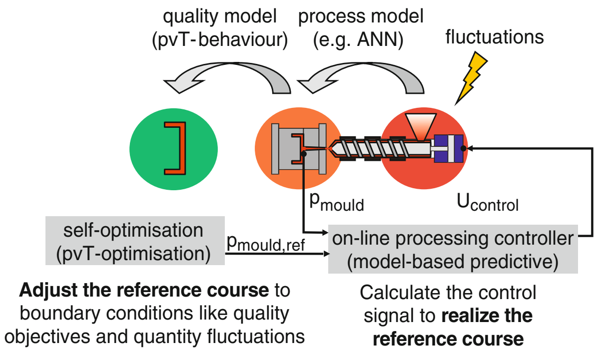


Fig. 6.63 Self-optimisation during the plastic injection moulding process

self-optimisation. A self-optimising 5-axis milling manufacturing system should be enabled by the use of these models in order to change its setting parameters autonomously and to meet the requested product quality demands at each position of operation.

6.4.4.5 Results Plastics Injection Moulding Process (IKV)

In injection moulding the assurance of a constant and high product quality requires a stable process control at an ideal operating point. The aim of such a process operation underlies the demand for a process which runs identically in each production cycle. The process is represented by different process variables, in particular pressure and the temperature of the melt in the cavity.

The effects of pressure and temperature on the specific volume (pvT-behaviour) of the material is used here as the basis for calculations to define the operating point and optimum process progress (Fig. 6.63). One requirement for the optimum operating point is a constant specific volume during cooling in the holding pressure phase. After measuring the temperature inside the cavity, it is possible to determine the set point for the control of the cavity pressure on the basis of the pvT-optimisation using the quality model. As actuating variable for process control a voltage is used, which enables the control of the servo inverter and hydraulic valve and thus the movement of the srew during the injection and holding pressure phase.

Typical disturbances in injection moulding are, for example, temperature variations of the melt and the tool, as well as fluctuations in the viscosity of the plastics material. Varying temperatures can occur during the manufacturing process, on the one hand during start-up or after breaks in production, when the stable thermal state has not yet been reached. On the other hand, there can also be control fluctuations or disturbances in the heating system which can lead to temperature variations. Viscosity fluctuations, which influence the pressure transfer to the tool cavity, typically occur due to variations in the batch of plastic or to the widespread use of recycled materials.

Table 6.5 Model prediction error and control error (Root Mean Square Error (bar) of linear and non-linear model structures with and without time information

	Linear (Regression)/bar	Non-linear (ANN)/bar
Model prediction error		
Without time input	5.2	3.2
With time input	5.1	2.7
Adjustment error		
Without time input	2.3	1.6
With time input	2.3	1.3

These disturbances should be compensated by a model predictive control as an ISA system, which has a cycle time of 8 ms using a dynamic process model. Here, the process model contains the correlation between the control voltage, which actuates the servo inverter or the proportional valve, and the cavity pressure as the control variable. Using identification tests, measured values are recorded which represent the non-linear and time-variable relationship between actuating and control variable. After calculating the time delay, which can be attributed to transporting the plastic melt when increasing and reducing the pressure, to the compressibility of the melt as well as to the inertia of the screw, the process models are created (Normeey-Rico and Camacho 2007). As the control path behaviour depends on the cavity used, the injection moulding machine and the material used, and thus often varies, the process modeling is automated as far as possible. This enables a simple identification of the control path behaviour. Using black box modelling, both linear and non-linear dynamic process models are created based on the measured data and evaluated by using k-step ahead prediction (Table 6.5). ANNs are used for non-linear modeling. These increase the quality of the model prediction. The quality of the model can be further improved by including the time as an input variable for the dynamic model. The prediction error of the process model then correlates with the control deviation of the tool cavity pressure control. In addition, the improvement in the model reduces the control effort, which helps to extend the machine's life.

A challenge when implementing the ISA-system is guaranteeing adherence to the control cycle of 8 ms on a micro-computer (Intel Core 2 Duo processor, 3 GHz with 3.25 GB RAM and Microsoft Windows XP operating system). The implementation of the ISA-system on a real-time system (NI 8353 RT from National Instruments Corporation, Austin, USA, running with an Intel Core 2 Quad Processor, 2.4 GHz with 2 GB RAM) with multi-processors even allowed reducing the control cycle to 4 ms. Future work will also include linking the real-time system to the MO-system.

Alongside the creation of the process model, a quality model is used to determine the internal objectives. The determination of the pvT-behaviour is being done today in laboratory devices. There are, however, limitations on how well the data derived in the laboratory could be transferred to an industrial injection moulding process. Determining the pvT-behaviour directly in the mould means that, although physically

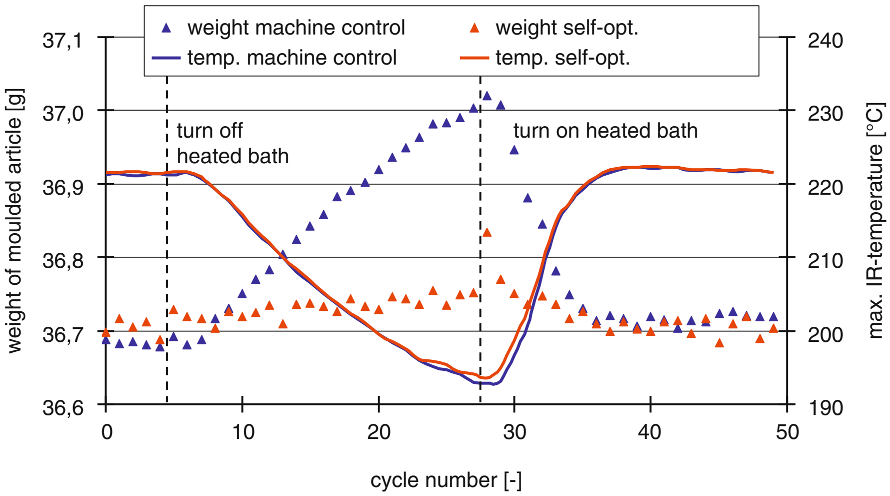


Fig. 6.64 Reaction of the self-optimising system to disturbances

correct pvT-diagrams cannot be obtained, instead characteristic diagrams of materials which are optimally aligned to the individual conditions can be gained (Michaeli and Schreiber 2010). The described process uses grey box modelling which is used to create the quality model. After capturing the temperature a reference course of the cavity pressure creates the internal objective which is implemented by an ISA-system. This cavity pressure set point is adapted by the quality model to the present temperature in the cavity. Within this project various options to acquire the temperature were researched. On the one hand, the melt temperature in the cavity was deduced indirectly using a mathematical approximation, and taking into account the temperature of the tool wall and the melt temperature at the tip of the nozzle. On the other hand, the melt temperature within the cavity can also be measured directly using an infrared temperature sensor. The expensive implementation of direct measurement competes with the inaccuracy of the mathematical approximation in the selection of method.

Model-Based Self-Optimisation should guarantee consistent, high product quality, even with disturbances in the process. Exemplarily introduced disturbances show the response of the Model-Based Self-Optimisation system. Figure 6.64 shows the response of the self-optimising system compared to conventional machine controls when both front heating zones in the plasticising unit are switched off. By measuring the quality values, such as the moulding weight or the actual moulding dimensions, it can be shown that the self-optimising system compensates the thermal fluctuations; the changes in temperature only have a negligible effect, and thus the quality consistency is increased. The active adjustment of the holding pressure course to the changed thermal conditions in the process therefore ensures a clearly improved consistency of quality of the moulded components. Thus, the system developed provides the basis for a complex self-optimising system that can work across cycles to intervene in the thermal balance of the injection moulding process. Additionally, the system ensures the compensation of thermodynamic fluctuations during process set-up and thus minimises set-up time. To prove its applicability in practice, the system was tested on the complex geometry of sports goggles, which are manufactured in a hybrid, multi-component injection moulding process.

6.4.4.6 Results Welding Process (ISF)

For monitoring and self-optimising on-line control of GMAW processes, it is absolutely necessary due to the dynamic of the process, that not only the unchanging initialisation values such as the base material, filler material, wire diameter, shielding gas, seam preparation, torch orientation etc. are taken into account, but also dynamic, sensor-based values. This is necessary since geometric changes on the component can occur due to the thermal influence of the welding process itself. Various sensors are used to gather this information from the current welding process. Electrical sensors measure the transient electrical current and voltage of the welding process. Using an optical sensor system the work piece geometry in the seam area is captured, allowing the determination of the width of the gap and contact tube distance.

In addition, a module of the simulation programme SimWeld was developed to calculate welded joint geometry as a stand-alone geometry solver, and optimised to be used during on-line operation. Using the initialisation values and the sensor-based data, the geometry solver calculates, on a continuous basis and dynamically, the current weld joint geometry, which is defined as the relevant product quality variable (Fig. 6.65). So even during the creation of the seam, it is already possible to react to any deviations from the prescribed seam geometry (seam height, seam width, root height, root width and weld depth). This information is used to monitor the weld seam geometry as well as to control the GMAW process, and therefore to guarantee the required product quality by selection of optimised setting parameters. This comprehensive provision of process and product information is a technological innovation for the welding process. It allows making the planning and control of a production line more transparent, and therefore more flexibly configurable.

When creating the models for self-optimisation the welding simulation programme SimWeld was a key tool. In cooperation with a research project on the production and processing of materials, SimWeld is used to simulate submerged arc welding processes as a component of the virtual process chain. The models used in this context are based on the developments needed for the simulation of GMAW processes. This makes it possible to amend the programme used for self-optimisation in such a way that it could be used to generate data and carry out on-line process simulation. The validation of models using real welding tests allowed further detail regarding the process behaviour to be added to the models used in SimWeld.

Generally, the SimWeld simulation system is a time-efficient way of providing process data to create models. This allows a considerable reduction of time and

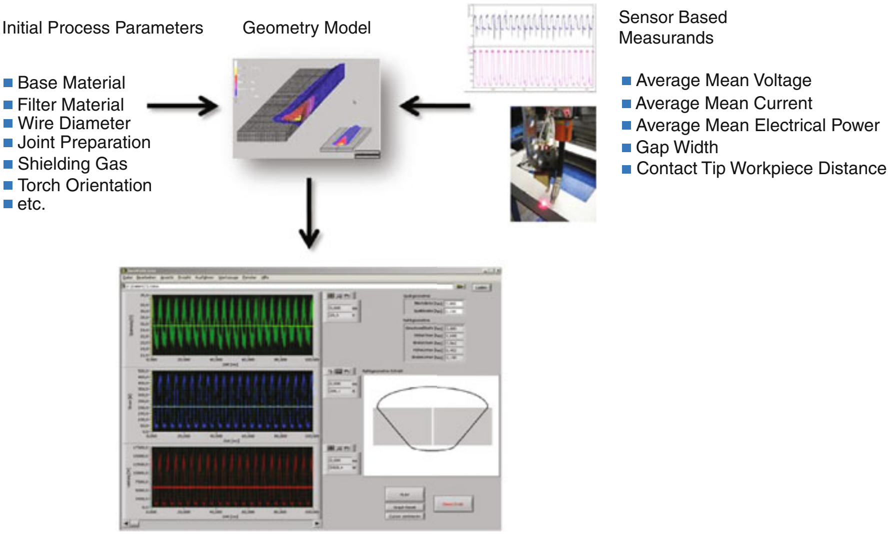


Fig. 6.65 On-line monitoring of weld seam geometry

technical efforts. Models to describe the behaviour of the overall process, and of each of the process domains for the different application areas, were created as the basis for implementing a method for Model-Based Self-Optimisation in the event of variations in the boundary conditions. Figure 6.66 documents exemplarily the welding process model created for two dependencies in the impulse welding process. In the left-hand part of Fig. 6.66 the height of the seam is displayed as a function of the wire feed speed and of the gap width for a square butt weld. The right-hand side shows the relationship between the gap width and the welding speed for a square butt weld.

With the help of the model relating to the overall process behaviour, it was possible to visualise in iso-lines the relationships between the setting parameters which led to an optimisation of the set-up procedure.

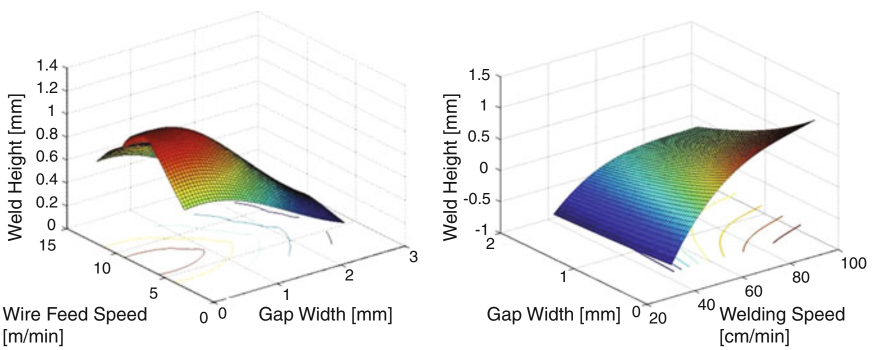


Fig. 6.66 Global surrogate model for impulse welding process

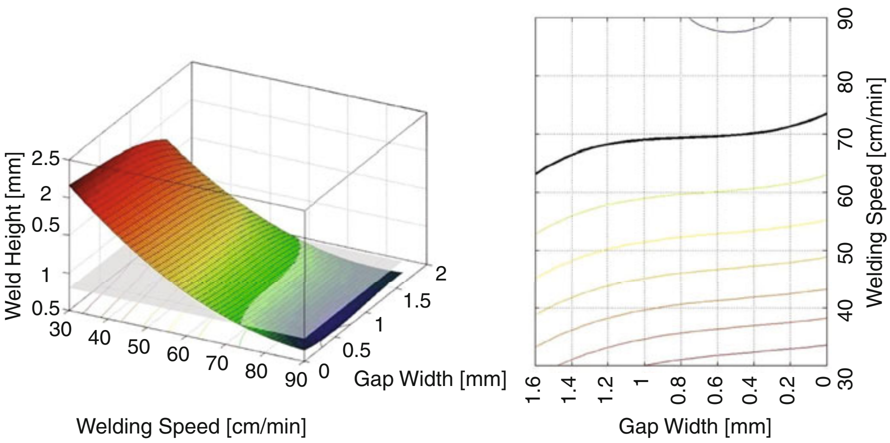


Fig. 6.67 Model-based description of root height

Figure 6.67 shows, for example, a model for the root height (left) and the derived relationship between the gap width and the welding speed with a constant seam height of 0.8 mm (on right). The other four setting parameters are kept constant. By analysing the iso-lines it was possible to derive the maximum possible weld speed at which a pre-defined product quality could be achieved. This functional relationship previously had to be determined by the operator applying his technical knowledge in time-consuming test welds. With the help of the model, this can now be done in the shortest possible time. This enables current on-line adaptation strategies, under which normally only the welding speed is adjusted during the welding process as a function of the measured gap width, a considerable reduction in both time and technical input.

A transfer of this approach to all six setting parameters is, however, not feasible, as the process models created do not show any strict monotonicity, so there is no functional relationship. Therefore, procedures had to be evaluated which enable inverse use of these models. In this context, various optimisers (stochastic and deterministic) were tested on the existing process models with regard to time efficiency and deviation from the optimum. For the set-up process this procedure did lead to an initial optimisation and selection of setting parameters. Future work will aim to qualify optimisation methods for on-line generation of setting parameters during the welding process.

6.4.4.7 Results Laser Cutting Process (ILT/LLT/NLD)

As with the other manufacturing processes studied, for laser beam cutting the key to self-optimisation is a model-based prediction of the processing result in relation to externally defined objectives. For laser beam cutting a large number of setting parameters need to be taken into account, which show strongly non-linear interactions.

If a meta-model supplies the relationships between setting parameters and processing results in order to control the process, then the cognitive component of a self-optimising manufacturing system can use the modelled process knowledge to determine the current operating point based on the acquired process measurands. In either case, the generation of a suitable meta-model can only succeed, if the parameter space is defined with high resolution. This requires large amounts of data, which can only be created at reasonable effort by using simulations.

Model Creation A particular challenge in the case of laser beam cutting consists in the fact that the actual “tool”, the interaction zone between the laser beam and the work piece, changes depending on the process variables resulting from the process dynamics. For example, the absorption profile on the surface of the emerging melt creates a so-called free boundary, whose shape cannot be known in advance, but which is part of the equations describing the process. These free-boundary-value-problems can only be solved very poorly using conventional simulation tools or commercial FEM solvers. What is more, the relevant time-scales and length-scales cover multiple orders of magnitude, which very quickly makes the computational cost become uneconomic, if a conventional numeric procedure is used.

For this reason, the focus was set to the development of a numerically efficient process model, which is capable of describing the cut quality—and in particular the formation of striations on the cut surface—as a function of the process parameters. To do this, the cutting process was described using a reduced model where, by separation of the time-scales and length-scales involved in the process, a hierarchy of subordinated sub-processes is first created (for example, heat conduction, fluid dynamics of the melt and absorption of the laser beam) and this is then further simplified by reduction of dimensions (Schulz et al. 2009, 2010).

This procedure was then expanded to describe the non-linear dynamics of laser beam cutting at high-resolution in both space and time, such that the creation of striations, resulting from melting and recrystallisation, on the cut surface can be calculated. For this, the dynamics of cutting were reduced to two fundamental differential equations for the position of the melt surface $M(z, t)$ and the melt film thickness $h(z, t)$, which can be solved with minor computational cost (compare Fig. 6.68a):

$$\frac{\partial h}{\partial t} + 2h\frac{\partial h}{\partial z} = v_p, \quad \frac{\partial M}{\partial t} = v_p - 1, \quad v_p = \frac{1}{Pe \cdot h_m}(Q_A - Pe) \tag{6.7}$$

$$Q_A = \gamma \mu A(\mu) f, \quad \mu = \varepsilon \left(\frac{\partial h}{\partial z} - \frac{\partial M}{\partial z} \right), \quad A(\mu) = \frac{4\mu \cdot i}{2\mu^2 + 2\mu \cdot i + i^2} \tag{6.8}$$

for $z, t \geq 0$, where $v_p = v_p(z, t)$ describes the dimensionless speed of the phase boundary for melting. The dimensionless energy flow density of the laser beam at the area of absorption $Q_A = Q_A(z, t)$ is described by the Fresnel absorption $A = A(\mu)$ as a function of its polarisation and the cosine of the angle of incidence $\mu = \mu(z, t)$. The spatial distribution of the laser beam is indicated as f; the scaling value for

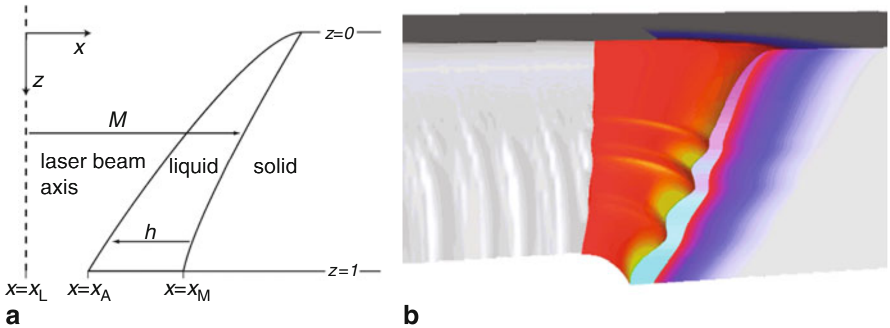


Fig. 6.68 Schematic 2D representation of the laser cutting process (**a**) and numeric simulation of the reduced model (**b**)

maximum intensity γ, the melt enthalpy h_m and the Peclet number *Pe* are explained in more detail in (Schulz et al. 1997).

The smallness parameter $\varepsilon = d_{m/d} \leq 1$ is the ratio of the typical melt film thickness d_m to the sheet thickness d. A detailed description of the reduced model can be found in (Vossen and Schüttler 2010). In the course of the studies it was demonstrated that the recrystallisation striations relevant for cut quality arise as the result of a self-organised melt dynamics.

Using the “QuCut” software which was developed (Vossen and Schüttler 2010), it is possible to calculate the virtual cut surface and quality criteria values for any parameter set, with strong correlation to the experimental results. Using the reduced model the calculation resources required for this could be reduced to the point where large-scale parameter variations are possible with high resolution.

These data form the basis for the meta-model, as they provide the sensitivities to parameter dependencies necessary for self-optimisation. The ability to calculate the relevant criteria values for the quality of a cut edge during laser beam cutting at this level of efficiency, with the help of simulations, has never been achieved before and is a significant result from the work carried out in this sub-project.

For one of the quality criteria, the virtually derived depth of roughness of the cut face R_z, an extract from the parameter space in Fig. 6.69 is shown as an example. Here, it becomes clear how the cut quality depends on the input parameters, and that significant local and global minima exist. The data produced through simulations are suitable for determining local and global optimum values, with the help of iterative optimisation algorithms, and for providing the corresponding internal objectives, whereas normally, in practice, both more parameters and also more elements (indicators) of PQ need to be taken into account than are shown in the example. For example, a global optimum value in relation to a single criterion may not be achieved, since infringement of another criterion or limits in the available ranges of individual setting parameters prohibit it.

Cognition The determination of the operating point of a manufacturing system for laser beam cutting is a significant step along the road to implementing

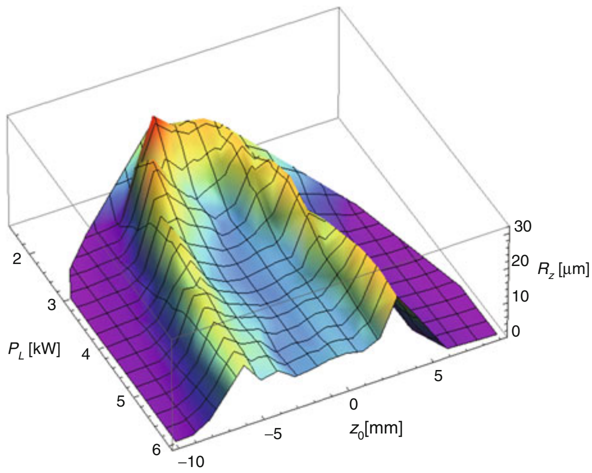


Fig. 6.69 Calculated quality indicators (depth of roughness R_z) as a function of two selected parameters (here: laser power P_L and focal position z_0)

self-optimisation. It is the pre-requisite for deriving decisions that can influence system behaviour. The focal position during laser cutting has a major influence on product quality but, particularly in systems which use CO_2 beam source, can only be measured with great effort. Special instruments to measure the beam enable determination of the energy distribution by using a scanning procedure, also in the focal position at energy densities of more than 5 MW/cm^2. But these measurements can only be made stationary within the machine, and generating measurement data during the process is not possible using this approach. Alternatively, the focal position can be determined experimentally using a so-called comb-cut. A cut is made in a thin sheet of metal at various focal positions, is taken out of the machine and measured with a feeler gauge. By determining the smallest width of all cutting kerfs the focal position can be determined. This value will be wrong if the optical components of the machine heat up differently in operation mode, or if the measuring equipment or manual evaluation were incorrect (Fig. 6.70).

If the manufacturing system is to become self-optimising, then it has to recognise the focal position itself and adjust it to the externally defined objectives and the requirements of the processing task, without external intervention. To achieve this, a cognitive unit is needed, which consists of a sensor to measure the width of the cut, a meta-model to translate the cutting kerf width into the focal position, and a controller to manage the focal position.

To determine the width of the cut when laser beam cutting, the heat conduction equation is discretised in one-dimensional space. This is done using two dotted stripes (see Fig. 6.71), in which the 1-dimensional heat conduction task is solved, in order to determine two data points for the melting isotherm. Because the melt film thickness next to the surface of the metal sheet is very small and the cut face in most relevant cases in practice can be assumed to be almost vertical, the melting isotherm on the sheet surface is a good measure for the expected cut width.

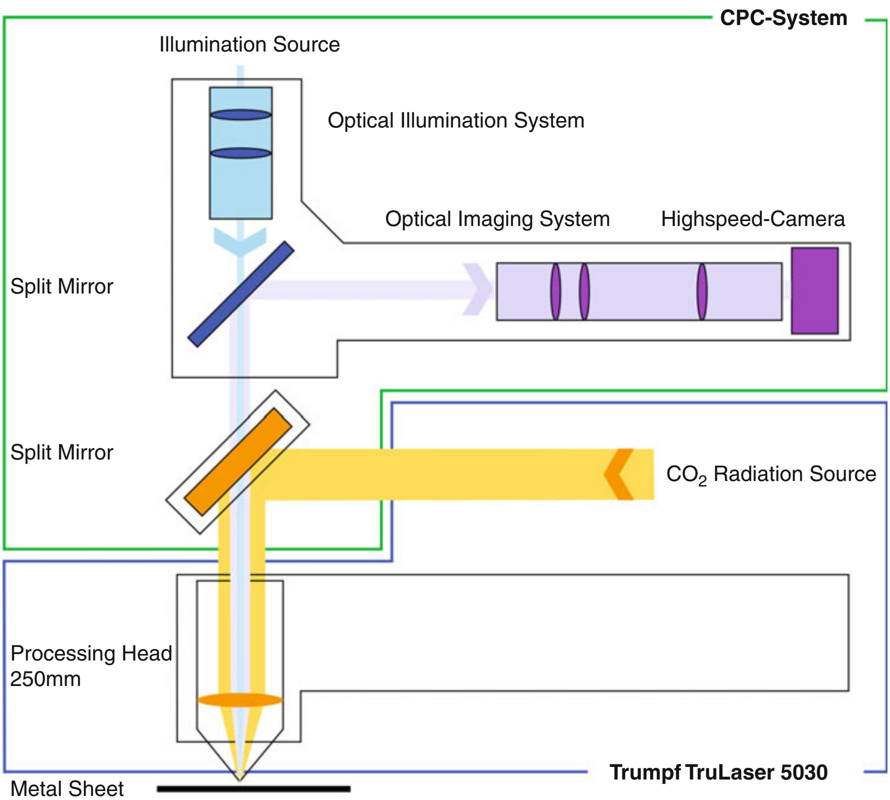


Fig. 6.70 System diagram of sensor placement in the TRULASER 5030

The point (x_0,y_0) or (x_1,y_1) within a dotted stripe where the melt temperature T_m is present as surface temperature T_S will be limited to the space between two data points and is approximated using linear interpolation. The surface temperature T_S and the depth of thermal penetration δ_{heat} are determined locally at the data point as a function of the local distribution of the intensity by using the following differential equations:

$$\frac{\partial T_S}{\partial t} = \frac{\kappa_T}{(1-b_1)\delta_{heat}}\left(\frac{Qa}{\lambda_W} - b_1\frac{T_S - T_0}{\delta_{heat}}\right) \tag{6.9}$$

$$\frac{\partial \delta_{heat}}{\partial t} = \frac{1}{T_S - T_0}\left(\frac{Q_a}{\lambda_W} - \frac{\partial T_S}{\partial t}\delta_{heat}\right) \tag{6.10}$$

with $b_1 = 3/5$, T_0 as ambient temperature, κ_T as temperature conductivity, λ_W as heat conductivity and $Q_a(r, t)$ as the heat flux due to absorption of the laser beam on the surface of the metal sheet.

Using the two points (x_0,y_0) or (x_1,y_1), which can be determined by resolving the above differential equations, and the x-axis as the axis of symmetry, a circle can be

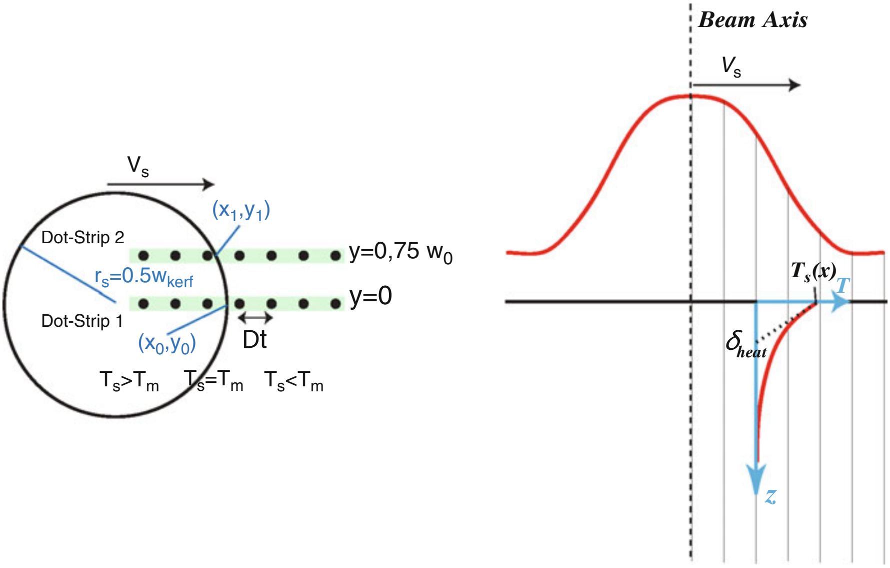


Fig. 6.71 Principle of the 1D pre-heating model to determine the cut width during laser cutting

defined, which is then used in the calculation of the cutting kerf width. This results in:

$$w_{Kerf} = 2 \cdot r_S = 2 \cdot \frac{(x_0 - x_1)^2 + 1.5 \cdot w_0}{2(x_0 - x_1)} \tag{6.11}$$

With this calculation procedure to determine the cutting kerf width, and by varying the parameters, base data can be generated from which a meta-model can then be created to predict the cutting kerf width from the focal position.

In Fig. 6.70 the sensor system is shown which was developed for the coaxial process monitoring which enables the problem to be solved. The processing beam is combined via beam-splitters to the process monitoring, what allows monitoring of the working area during the process. Information about the process is imaged using optical elements on a 2-dimensional detector and this gives spatially resolved measurands. Signal processing allows extraction of process variables and makes these available in machine-readable form for optimisation.

The heating of the beam splitter has a major influence on the information content of the sensor data (Fig. 6.72b). Figure 6.72a shows an image of the cut without processing, so without the processing beam. In this case the beam splitter is at its nominal temperature and achieves the image quality calculated previously using optical simulation. In Fig. 6.72c the influence of heating can be seen with the resulting reduction in the information content of the sensor signals. The aberrations relating to power input can only partially be compensated in the algorithms which process the signals and which extract the width of the cut from the sensor signals. If this point had not already been covered in the design phase, this would have been a case for interating in the methodology entering the creation of a meta-model at Point C4 as shown in Fig. 6.55. In the case under consideration, a module was

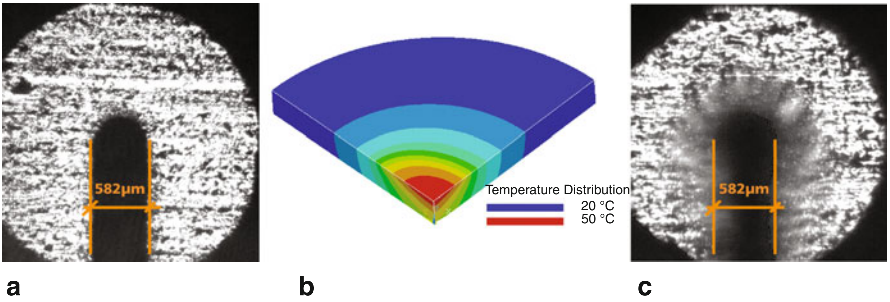


Fig. 6.72 Temperature behaviour of sensors in laser cutting. **a** Image quality at homogeneous temperature beam splitter **b** temperature distribution of beam splitter **c** Image quality at inhomogeneously heated beam splitter

developed to optimise the image quality, which receives data about the heating of the beam splitter from FEM data and makes this data available to the ZEMAX optical simulation. This allows imaging systems to be optimised based on power, energy distribution and cooling. The knowledge gained from these simulations, on the one hand feeds into the optimisation of the process monitoring, but on the other hand also allows the evaluation of signals collected from the process, and so the development of new algorithms. This module not only makes a contribution to the following steps to expand the cognition of laser cutting machines, but also makes a universal contribution to the simulation of thermally affected optical components.

In relation to the focal position, the signals for process monitoring are evaluated using algorithms which extract the characteristics of cut width. The image of the processing area shows the cutting kerf with its head and both side edges. The exact cut width is extracted using multi-stage image processing. The relationship between cut width and focal position is described by a meta-model. The input parameters for the model are the caustic, the power and power distribution of the laser beam, and the feed rate. Output parameter is the width of cut. By comparing the cut width calculated by the model to the measured value a control deviation is found. This is controlled by the ISA-system for the focal position.

The set-up assistant demonstrates the function of focal position measurement during processing, and therefore replaces the measurement by means of a comb-cut. This considerably reduces the time needed for rigging of the manufacturing system, and allows an instant, repeatable measurement, with the prospective possibility of closed loop control of the focus position during processing. The transfer of this parameter into self-optimisation is an example of gradual integration of self-optimising components in manufacturing systems.

6.4.4.8 Results Weaving Process (ITA)

The values used to set-up a loom are normally gained from the weaver's experience. In addition, tests are needed to check whether the required product quality is being achieved. The aim of the study of the weaving processes is therefore self-optimisation in the loom set-up for a given product quality. With the help of a self-optimising set-up assistant, the test effort and its associated employee and materials costs are reduced. Important quality variables for fabrics include its mass per unit area and its tensile strength. The quality variables are depending on the characteristics of the used material, as well as the weave construction and the manufacturing parameters of the weaving process. An important factor is warp tension. It influences the crimp of the textile and hence the mass per unit area and the tensile strength. Basically the following context applies for the mass per unit area:

$$m_G = m_K + m_S \tag{6.12}$$

$$m_K = Fe_K \cdot Fd_K \cdot \left(1 + \frac{E_K\%}{100}\right) \tag{6.13}$$

$$m_S = Fe_S \cdot Fd_S \cdot \left(1 + \frac{E_S\%}{100}\right) \tag{6.14}$$

(with a mass per unit area of m_G, mass of warp yarn m_K, mass of weft yarn m_S, warp density Fd_K,warp yarn count Fe_K, crimp in of the warp yarn E_K, weft density Fd_S, weft yarn count Fe_S and crimp of the weft yarn E_S). For the crimp, the relationship of

$$E\% = \frac{(l - r)}{r} \cdot 100. \tag{6.15}$$

is applied where l is the length of the straight yarn and r is the length of the woven yarn (Latzke and Hesse 1974). The length of the woven yarn depends on the warp tension (Akgün et al. 2010).

For further research the minimum, maximum and mean traction of a warp during the weaving process were defined as process variables. The weft density, the loom speed, the warp tension of the entire warp system and the positioning of the warp stop motion (angle, vertical and horizontal positioning, Fig. 6.73) were selected as setting parameters.

The setting parameters have a significant influence on the time course of the warp yarn tension. The position of the warp stop motion changes the so-called rear-shed geometry of the loom. It has a determining influence on the course of the warp yarn traction during the weaving process. In comparison with other loom elements such as the back rest or the shafts, the effort needed to adjust the warp stop motion requires only a small manual effort.

To create the data base for modelling, the following parameters were varied using a factorial test plan:

- Loom speed [turns/min]
- Weft density [threads/cm]
- X-position warp stop motion [cm]
- Y-position warp stop motion [cm]
- Angle of warp stop motion [degrees]
- Warp tension of the entire warp system [kN]

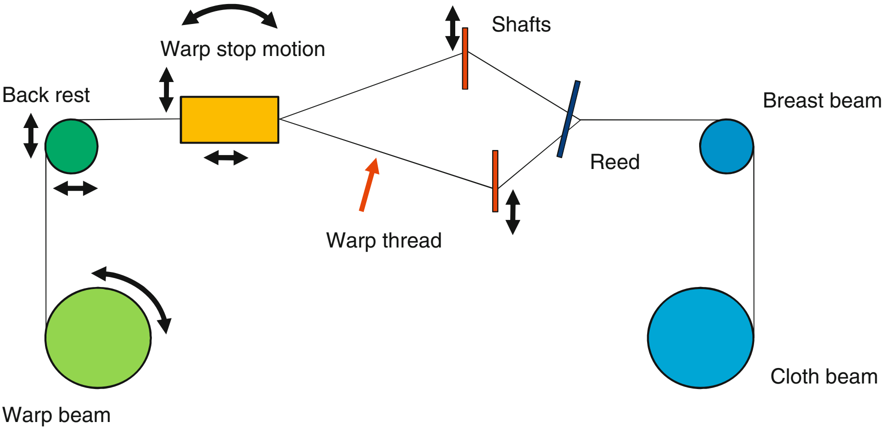


Fig. 6.73 Loom principle

The warp tension of the entire warp system, the loom's speed and the weft density are gained from the loom via a network interface using a specially created software. To capture the maximum, minimum and mean warp tension a sensor developed at ITA is used, which is shown in Fig. 6.74 (Gloy et al. 2009).

The sensor enables the acquisition of dynamic yarn tension. The data acquisition is implemented by a measurement amplifier, an analogue/digital converter and software to record the data. The values for the strength of the textile were obtained using tensile tests according to DIN EN ISO 13934-1. The mass per unit area is determined as defined by DIN EN 12127. Measuring the mass per unit area on the loom is currently not available industrially. In theory, there is an option to use radiometry sensors

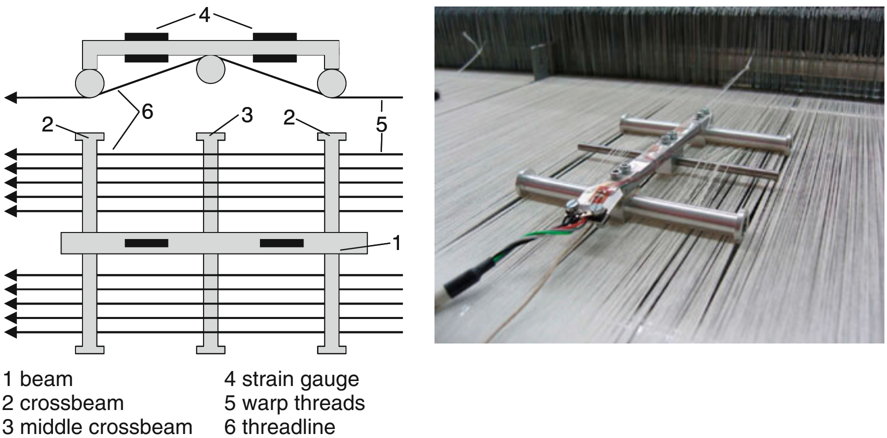


Fig. 6.74 Sensor to measure dynamic warp tension

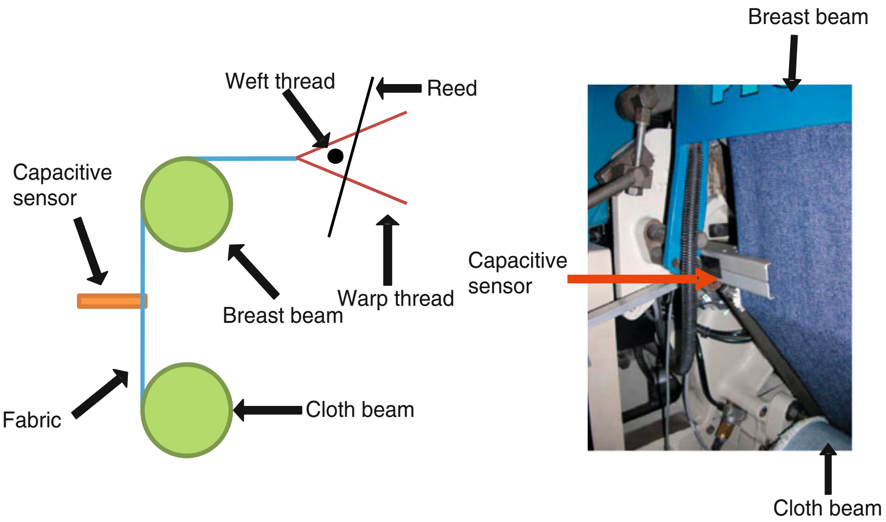


Fig. 6.75 Capacitive sensor on a loom

(Beta-back reflection method, ultrasound measurement and infrared measurement) to measure the weight of the textile on-line. But the investment required for this kind of equipment is higher than for the loom itself. Therefore, initially a capacitive sensor was inserted between the breast beam and the cloth beam on the loom (Fig. 6.75). The fabric is guided through the two capacitor plates. Any change in weight of the textile leads to a change in capacitance and so to a measurable change in the current output by the sensor.

Using the data base generated in this way, models are created to find the correlation between the setting parameters and the process variable of yarn tension. Here, it turns out that the positioning of the warp stop motion has the most impact on the course of warp tension, see Table 6.6. In contrast, the angle of the warp stop motion has only a minor impact on yarn tension and can be ignored for any further observations.

Analysing the correlation matrix of the process measurands in regard to the product quality, the mass per unit area, the minimum yarn traction has the greatest impact. Modelling of the correlation process variable to product quality regarding tensile strength is shown in Fig. 6.76. The bigger the box in an appropriate field is the

Table 6.6 Correlation matrix of settting parameters for the process variable of warp tension when weaving (from 0 minimum to 1 maximum influence)

Setting parameter	Minimum warp traction	Maximum warp traction	Mean warp traction
Revolutions	0.187	0.06	0.02
Weft density	0.195	0.129	0.133
X-position warp stop motion	0.06	0.05	0.01
Y-position warp stop motion	0.39	0.86	0.73
Angle warp stop motion	0.35	0.13	0.32

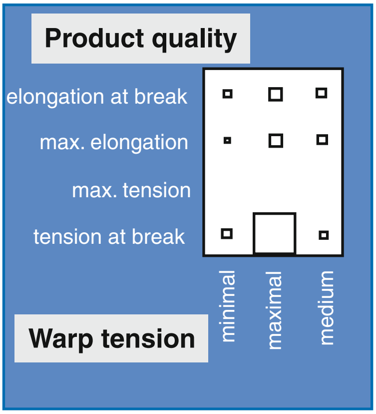


Fig. 6.76 Correlation matrix of the minimum, maximum and mean warp tension in relation to the elongation at break, maximum elongation, maximum tension and tension at break of a fabric

greater the influence of the parameter. It is notable that the maximum warp traction has the greatest impact on the strength of the textile.

Furthermore, a model was also investigated for the correlation between the setting parameters and product quality, using the "AX-Workshop" software developed as part of the project. The models used and their coefficient of determination are shown in Table 6.7. To improve the model quality, the base data needs to be expanded. Basically, it can be seen that the models are suitable for calculating product quality.

In ongoing work, inverse modelling is being researched, to enable the determination of the necessary setting parameters for a requested product quality. In collaboration with industry partner Picanol NV, Ieper, Belgium, a self-optimising set-up assistant is being developed as a next step.

6.4.5 Industrial Relevance

Increasing the competitiveness in an international scope using innovative and sustainable system solutions is a constant demand of manufacturing companies in high-wage countries. A distinct variable in meeting this general requirement is the manufacturing process itself. For companies, this creates the wanted result by the use of given social and technical production factors. The state of the art in technology in 2006 shows that manufacturing processes are not being run at their physical limits. To ensure product quality, mainly conservative manufacturing strategies are being selected, in order to avoid process instabilities from the outset. Depending on the level of complexity of the manufacturing process, and the dimensions of the determining

Table 6.7 Models for translation of SP to PQ and their absolute errors

Model for transference of SP to PQ	Coefficient of determination R^2
Regression model	0.89
Feed Forward neural networks	0.81
Exact RBF networks	0.2

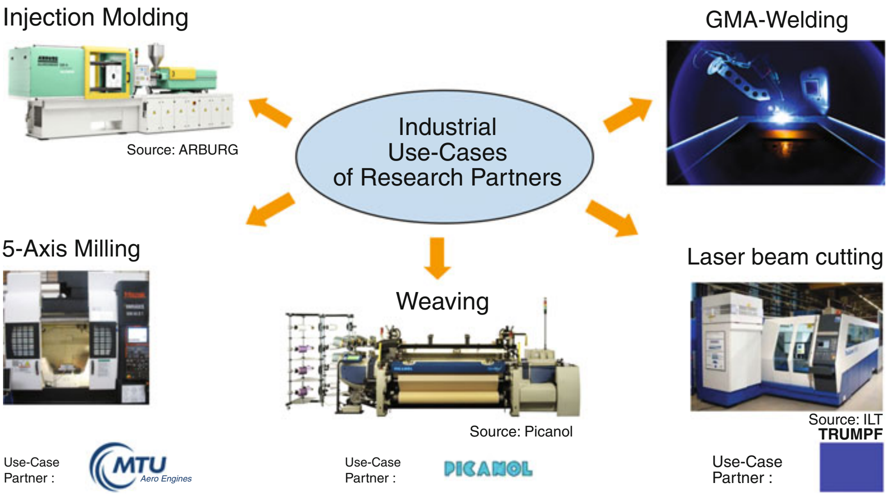


Fig. 6.77 Industrial use-cases with research partners

factor which are relevant for quality, this is achieved only with considerable set-up or inspection costs as well as re-work when deviations from predefined product quality occur. In addition, during the set-up phase after product changeovers, intensive test runs are needed in order to satisfy the required quality levels. This means, in particular for small batch sizes, a disproportionate amount of downtime in overall production. As a result, the described approach leads to high costs, in terms of both time and money.

The manufacturing industry seeks to optimise manufacturing processes in terms of processing speed and a reduction in downtime, always assuming product quality can be maintained. This goal is reflected in the basic research question, with the focus on rigging, monitoring and control of manufacturing processes. As a basis for ongoing validation of progress and growth in knowledge in close alignment with economic questions, each manufacturing process has an industrial use-case assigned to it. The industrial applications and the partner companies are shown in Fig. 6.77.

The joint work of the five research partners on the solution of the tasks ensures that solutions are generic in structure. Both the concept of Model-Based Self-Optimisation with its ISA-systems, and the process-independent methodology to generate meta-models, are tested by all five processes. This means that their application to further processes is possible, which guarantees the sustainability of the research outcomes in the conceptual domain. Sustainability in relation to gaining process-specific findings is demonstrated by the research partners' results.

6.4.5.1 Milling (WZL-TF)

The results achieved, if successfully transferred to industrial practice, will allow an improved process understanding and in a first instance enable a more efficient design and improved controllability of this manufacturing process. Integrated into the process design phase, the developed position-oriented monitoring system pinpoints unstable process behaviour in order to enable corrections in a targeted way. For industrial practice, this application would reduce the necessary number of preliminary test runs, resulting in a reduction of resource consumption and significant cost-savings for single and small lot production. Also, this optimization ensures the product's quality and therefore avoids, due to the acquired process information, instabilities by suitable parameter adjustments, In the simplest case the manufacturing process can be stopped at its current position, avoiding further damage to the component and therefore scrap costs. A reduction of both the manufacturing's scrap rate and required post-processing steps would significantly contribute to a company's competitiveness in the area of 5-axis milling.

Currently, the results achieved for machining thin, overhanging cantilevers are transferred to a demonstrator, optimizing the finishing operation of a blisk's thin-walled blades. A blisk (blade integrated disk), known as a safety-critical component in the aerospace industry, is subject to the highest quality requirements in order to avoid failures in the later use. At present, the expected product quality can only be achieved executing time consuming and expensive test runs. Integrating the results presented, a quality-oriented monitoring of the finish milling process will be established which reduces the number of necessary process design steps by far. For this purpose, the results of the process monitoring are combined with those of the virtual process design.

6.4.5.2 Model Evaluation (WZL-MQ)

The methods implemented in the AX-Workshop software for process-independent process modelling, visualisation and evaluation, enable the efficient creation of surrogate models for self-optimisation. An important part of this are the evaluation algorithms, by which the quality of the employed models can be evaluated in relation to various criteria, such as global and local representation quality, or robustness against disturbances, before they are put into use. In this way, evaluation provides the necessary criteria for the implementation of surrogate models for Model-Based Self-Optimisation in industrial practice. The AX-Workshop software and the common methodology for creating meta-models provide an important contribution to making model-generation systematic. The shift from an intuitive model-generation process, only open to implementation by experts, to a standardised procedure with software support, reduces modelling effort substantially, and allows a quality focussed generation of models.

6.4.5.3 Plastics Injection Moulding (IKV)

Injection moulding of thermoplastics is one of the most important processes in plastics processing. Thanks to the high level of flexibility displayed by mouldings' geometries and their materials, products can range from micro-components to car wings or complete telephone booths (Johannaber and Michaeli 2004; Michaeli 2006). In the automobile industry injection moulding is widespread, especially for technical parts. The injection moulding process is at the start of the process chain for creating complex assemblies. This means that fluctuations in mouldings' quality such as distortion propagate through the process chain. Therefore, it is important that the injection moulding process reacts robustly, even in the case of disturbances, as occur from the use of recycled materials in the mix. The works within the cluster of excellence focuses mainly on self-optimisation to stabilise the injection moulding process. In close cooperation with the machine manufacturer ARBURG, systems will be further developed in future to stabilise moulded part quality.

6.4.5.4 Welding (ISF)

The technological results achieved as part of the research work carried out in the area of automated gas metal arc welding, if adequately transferred to industrial manufacturing, allow to expect clear increases in efficiency in technical welding applications in the areas of cost-efficiency, quality assurance and flexibility.

By preparing and applying a global quality model, the rigging time required to prepare for manufacturing can be decisively reduced. The necessary setting parameters for the required product quality, in this case the weld seam geometry, are found almost instantaneously with the help of optimisation algorithms in the defined quality model, which means time-consuming and cost-intensive test runs are no longer needed. Depending on the application, the necessary time for setting up and running-in processes can be reduced by up to 50%. This is achieved by enabling expert staff to decide more rapidly on the adjustment of values for setting parameters, as they are supported by the technical welding expertise implicitly contained in the quality model.

The results presented in the area of on-line monitoring, particularly through reliable in situ prediction of the weld seam geometry, based on the welding current, welding voltage and sensor information such as gap width, workpiece to contact-tip distance and welding torch orientation, now for the first time enable a closed control loop to actively influence the weld seam geometry resulting from the weld pool, and to adjust the process to suit the required geometry. Thanks to this prediction of the weld seam geometry, the optimum process parameters can be determined and set in the light of the process information listed above, without extensive empirical tests, so that consistent weld seam geometry can be ensured. This leads to a sustainable reduction in rejects, and to a reduction in the necessary repair work of up to 30%.

The self-optimising gas metal arc welding manufacturing system is notable for its extensive flexibility, only restricted by the limits of the process model, and offers the possibility of reacting very swiftly, and therefore with greatly reduced economic effort, to changes in manufacturing conditions.

6.4.5.5 Laser Beam Cutting (ILT/LLT/NLD)

The process of laser beam cutting is an established industrial process. In many areas, flat-bed cutting machines are used to prepare steel sheets which are subsequently converted into switchgear boxes, façade claddings, construction components in the areas of mechanical engineering or plant engineering, or which find their use as sub-components in complex technical devices.

For the research in the area of self-optimisation, the company TRUMPF provided a flat-bed cutting machine to the cluster of excellence, equipped with a CO_2 laser beam source. This machine is sold in large quantities on the international market, and can manufacture continuously if automated accordingly. Optimisation potential remains in the area where the cutting process is not yet run at its technical limits, so where self-optimisation is capable of increasing the speed of manufacturing. This relates to both questions of process stability, as well as the achievable quality of the processing result.

The set-up assistant demonstrated as part of the project delivers a significant reduction in the rigging time of the focal position and increases the reliability of the result. The progress in modelling the cutting process already today enables the integration of information from simulation based on variable material data, which provides a way forward to its application for special alloy materials. On a second front, the prediction of the stability of the cutting process taking key parameters into account is now possible. The cognition which has been achieved in the manufacturing system, provided by sensor technology and model-based signal processing, has the potential to be extended to further parameters in the manufacturing process. This leverage has the potential for a significant increase in the processing rate, bringing it closer to the physical limits of the process. By replacing manual evaluation of a comb-cut to determine the position of the focal position with the use of focal position cognition, an important step has been taken in the direction of self-optimisation of a technical system. Parts of the solutions in this area are already finding their way into the development of a new generation of cutting machines.

It is expected that the application of self-optimisation to an extended set of parameters will enable a clear increase in the rate of manufacturing, and pave the way for expansion into other application areas, from the cutting of thicker materials to the rapid and precise processing of high-tensile steel for the construction of light-weight automobiles. Here, the research is making a contribution to current socio-political and economic success factors for high-wage countries.

6.4.5.6 Weaving (ITA)

The research in the area of weaving was carried out using looms from the loom manufacturer Picanol NV, Ypres, Belgium. Aim is the development of a set-up assistant for looms, which provides the necessary setting parameters to achieve a given quality of fabric. Picanol has provided a loom for the research work on which the sensors developed to monitor product quality on-loom are being evaluated. In addition, commands and interfaces for communication with the machine are being provided, which will allow direct transfer of results to an industrial application.

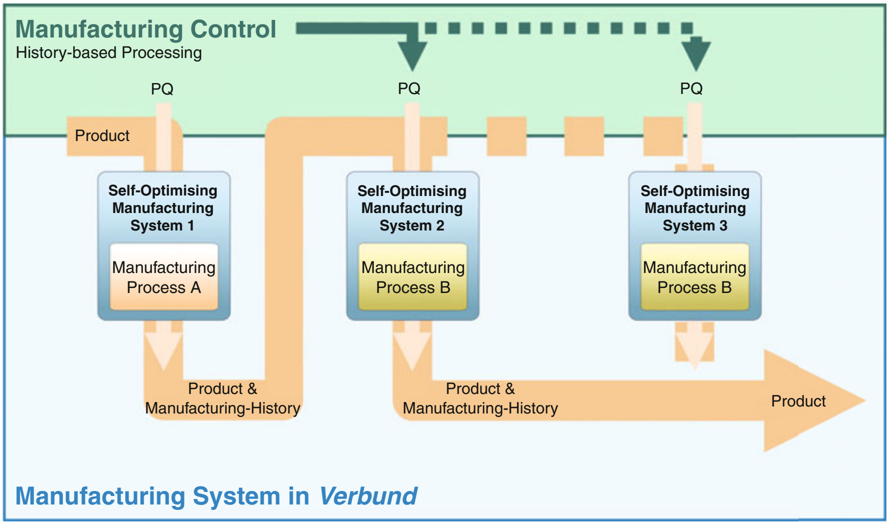
6.4.6 Future Research Topics

The research and development in the area of self-optimisation of manufacturing systems is focused on providing a generic answer to the question of how manufacturing processes can be described by models, so that they are robust against disturbances, and can themselves adjust parameters, so that the product which is produced meets its specifications. The strategy developed to implement self-optimisation, and the methods for generating meta-models, taken together with the description of processes in reduced machine-readable models, are the two main pillars of "Model-Based Self-Optimisation for manufacturing systems". The creation of set-up assistants shows the innate strength of the proposed solutions. In addition, the works in cooperation with industry partners show the possibilities for gradual implementation of self-optimisation through selective integration of ISA-systems in existing manufacturing systems. These results contribute to the reduction of rigging effort in product quality focused manufacturing.

In future studies it is necessary to further develop the concepts and methods, so that they can serve as a general basis for industrial engineering services. This is only feasible today in some selected cases. The validity of the methods for developing surrogate models has already been demonstrated, using the manufacturing processes which were studied. But to date it has not been possible to make any claims in relation to the stability arising from the interaction between meta-model, process and control system. Therefore, the expansion of self-optimisation from process monitoring and process control of isolated scalar parameters to dynamic multi-dimensional parameter spaces with the stability issues these entail, is a substantial focus for research.

The methods for the generation of models currently are targeting the manufacturing process. To successfully relocate within the parameter space, Model-Based Self-Optimisation has to take the machine behaviour into consideration. Therefore, methods need to be developed to combine algorithms for model-based optimisation with the machine-specific dynamic behaviour and performance of the machine. As it can already be predicted that, in most cases, this type of wide-ranging information is unlikely to be available in entirety and generically, methods also need to be developed to determine and describe the machine's behaviour under all relevant processing conditions.

The determination of the operating point within the manufacturing process has already been demonstrated for a reduced parameter space. Manufacturing processes which need to evaluate multiple process variables in order to reliably determine their operating point at times need to process a vast amount of data instantly. So that the



PQ: Product Quality

Fig. 6.78 Self-optimising manufacturing systems in *verbund*

analysis of this data can lead to an increase in cognition, new concepts for multi-sensor concentration are needed. Independent of its information content, acquired data needs to be used with respect to its context both for model-based optimisation systems in a superordinate scope, and also for subordinate systems within control loops.

Under the title "self-optimising manufacturing systems in *verbund*[2]" the aim of complete transparency of the manufacturing process will be pursued during the second phase of the cluster of excellence. This quality-focused, product-to-product transparency enables the realisation of quality-guided manufacturing Verbünde, where the final product defines the requirements which all manufacturing systems in the federation must comply with. Each manufacturing system also uses the manufacturing history of the product as an input variable, processes the product according to the externally defined objectives, and forwards the finished product accompanied by an extended history. Subsequent processes use this information to adjust their processing, so that they in turn achieve their externally defined objectives (Fig. 6.78).

"Self-optimising manufacturing systems in *verbund*" will demonstrate their ability to adapt to changed parameters on the input, and to changed requirements in the processing result. The gain in the area of stability, and the reduction in rigging, contribute to economic factors for the production of individual items, and address the scale-scope axis of the polylemma of production. The increase in flexibility and the

[2] Verbund in this context describes the union of cooperating manufacturing systems which themselves react autonomously to changing requirements such that all entities within the union together achieve the global objective of the union.

interaction with the planning level address the plan-value axis. By turning historical information from manufacturing steps into valuable input for the planning level, the processing parameters can be adjusted at an early stage for the subsequent manufacturing steps. This modularisation of the manufacturing site can also be applied using the model-based approach in virtual environments, as the quality-relevant behaviour exists in machine-readable form.

6.5 Integrative Product and Process Design for Self-optimising Assembly

Christian Brecher, Peter Loosen, Rainer Müller, Nicolas Pyschny, Alberto Pavim, Matthias Vette, Adam Malik, Kamil Fayzullin and Sebastian Haag

6.5.1 Abstract

In the scope of the Cluster of Excellence "Integrative Production Technology for High-Wage Countries" and more specifically within the sub-projects "Self-optimising Flexible Assembly Systems" and "Model-based Assembly Control" the application of self-optimisation principles in the automated assembly of complex, multi-variant products are under investigation. A miniaturised laser system developed in-house—the "MicroSlab" (see Sect. 7.3)—is used as demonstration scenario.

The motivation behind this research results in particular from the rapid development of optical technologies, which has caused a steady increase in the demand for products with optomechanical features. The technological development of such products is driven by a clear trend towards miniaturisation, which defines the highest demands on the precision of production and assembly technologies. Today, precision assembly is characterised by a high degree of manual or semi-automatic processes, which encourages production migration to low-wage countries. The automation of these processes is an essential prerequisite for manufacturing competitive products in high-wage countries. Due to high engineering efforts and high installation costs, a high degree of automation is currently only economically viable for the mass production of standardised products. A constant shortening of product life cycles and the demand for a bigger diversity of variants represent major impediments to extensive automation levels. Therefore, investigating the possibilities for increasing the flexibility of automation in optomechanical assembly is an essential task. Flexible automation enables the manufacturing of customised products at low production costs and contributes to the resolution of the "scale-scope" dilemma. Here, automation is studied in terms of job retention: a trained operator should be optimally supported and complemented, not replaced.

Turning automation processes more flexible induces increased organisational and planning effort. A detailed and hierarchical planning structure in form of a fixed

assembly sequence for a wide range of products could only be implemented with very high effort and considering assembly components with high tolerance (and costs). The approach used in the sub-project for resolving the dilemma between planning and value orientation in assembly planning is based on the self-optimisation of assembly processes. In this respect, some of the relevant decision making is transferred to the assembly system during the execution of the assembly process. This implies reducing the planning efforts, because not all decision trees have to be fully defined in advance, but rather they can be created during the assembly process. According to the definition of self-optimisation, the assembly system must be able to derive new goals from observations of the assembly process and adapt itself to the dynamical boundary conditions. The required level of autonomy for such purposes can be achieved by means of implementing modular substructures (for example "software agents") that can perceive their environment and react to changes based on rules or models. Hence, system operators are made responsible only for coordination tasks, such as planning the product features or evaluating time, costs or quality of the assembly end product.

The following section introduces an insight into the state of the art in the field of micro and precision assembly of optical systems, from which the research purpose is derived.

The results section firstly develops the topic of self-optimising assembly systems on a conceptual basis. Based on a basic definition of self-optimisation in technical systems, the prerequisites, or building blocks, for self-optimising assembly systems are then derived and explained (see Sect. 6.5.4.1). These include:

- the flexibility and mutability of automated systems,
- the autonomy of assembly cells,
- automated planning in flexible assembly systems and
- cognition and learning ability in automation technology.

In the sequence, some specific solution approaches for the automated assembly of the MicroSlab are explored and the solutions developed for model-based assembly control are presented (see Sect. 6.5.4.2).

In addition to addressing the topic of self-optimising assembly systems, methods for the integrative design of products, processes and equipment in assembly were developed and these are presented in Sect. 6.5.4.3. Finally, future research in the individual topics is outlined.

6.5.2 State of the Art

Technological developments in many areas of the consumer and capital goods industry are driven by dynamic changes in customer requirements and a clear trend towards miniaturisation combined with functional integration (Lotter 2006). In various areas of consumer electronics, automobile manufacture and industrial production as well as in medico-technical applications, miniaturised approaches help to find better solutions to problems or to even make it possible to solve certain problems for the very first time (Pfirrmann and Astor 2006).

Based on rapid developments in microelectronics and semiconductor technology, a multitude of different manufacturing techniques from microtechnology have been established such as coating, lithographic and etching techniques, which, through parallel production (*batch processing*), are suitable for large lot sizes and thus achieve low unit costs.

Through monolithic integration, i.e. the integration of sensors and actuators using the structuring techniques mentioned above, various and sometimes highly complex technical solutions have been developed for actuation systems, sensors and analysis systems since the beginning of the 1980s under the general term *microsystems technology* (*MST*) (Völklein and Zetterer 2006). In a microsystem, in addition to the electrical and mechanical components, there are often optical functional elements too (Hankes 1999). Although there is no sharp definition of microsystems technology, there is generally a consensus that it describes the discipline of integrating as many different functions as possible in a small space, whereby at least one component of the microsystem is produced micromechanically.

In order to manufacture microsystems, driven by applications with micromechanical, microoptical and microfluidic components, attempts are being made to open up the third spatial dimension in production technology. Despite some progress, 3D capability in production processes remains extremely limited, becomes very costly as the complexity increases and is only viable where very large unit numbers are concerned. Against this background, the type of system integration (monolithic or hybrid) is increasingly determined by the economic boundary conditions and is following the clear trend towards hybrid integration from various production environments.

The area of production technology concerned with the hybrid integration of miniaturised products is usually referred to as microassembly. According to DIN 32564-2, the term "microassembly" is defined as follows: "microassembly is the assembly of microtechnical components, the fitting of microcomponents on mounting surfaces or their installation in housings, including electrical contacting and the creation of other connections (e.g. media)" (DIN 32564 2003). A more precise classification of microassembly compared with the related discipline of precision assembly as well as conventional assembly is shown in Fig. 6.79. According to this classification, microassembly is characterised primarily by small component dimensions, whereas precision assembly is characterised mainly by the required assembly precision of less than 25 μm.

6.5.2.1 Microassembly

For Germany as a production location, microsystems technology offers long-term employment prospects. At present, there are 680,000 jobs in the direct technological environment, plus the jobs in the application sectors. Economically, the market for

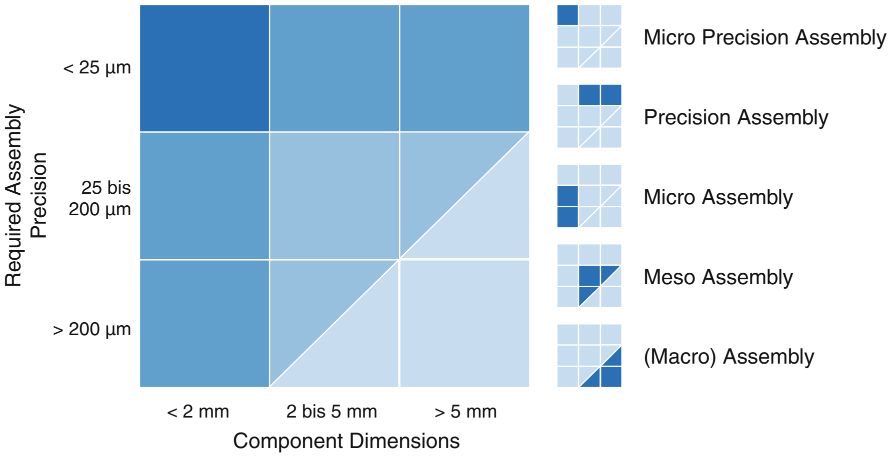


Fig. 6.79 Systemisation of assembly techniques. (Petersen 2003)

microsystems is already an above-average growth market with estimated average growth rates of 16%. These estimates do not include the value-adding portion of microsystems technology which results from integration into innovative products and of which experts expect a "leverage" factor of seven (Salomon 2006).

Due to the economies of scale through mass production, the unit costs of standard parts and components for microsystems are often very low, whereas the cost caused by microassembly is extremely high—up to 80% of the overall manufacturing costs of a microsystem. This is due to the predominantly low level of automation combined with high wage costs in Germany (Hesselbach et al. 2002). This could be interpreted as the reason why, in contrast to monolithic microsystems, a distinct increase in the unit numbers is still proving elusive in the field of series production of hybrid microsystems (Korb 2005).

In addition to high wage costs, there are other reasons for conducting intensive research into the flexible automation of hybrid microassembly. The small component dimensions cause difficulties with manual assembly because the highly precise positioning and focussing on small parts, e.g. under a microscope, are ergonomically suboptimal for human employees. If miniaturisation continues to progress, it will become more difficult for humans to carry out the required assembly operations safely and efficiently while maintaining the same quality.

6.5.2.2 Precision Assembly of Optical Systems

Optical technologies have now developed into a cross-sectional and key technology. This has resulted in a constantly growing demand for products with optical and, in particular, microoptical components in many technical fields (medical technology, information and communication technology, measuring technology, production

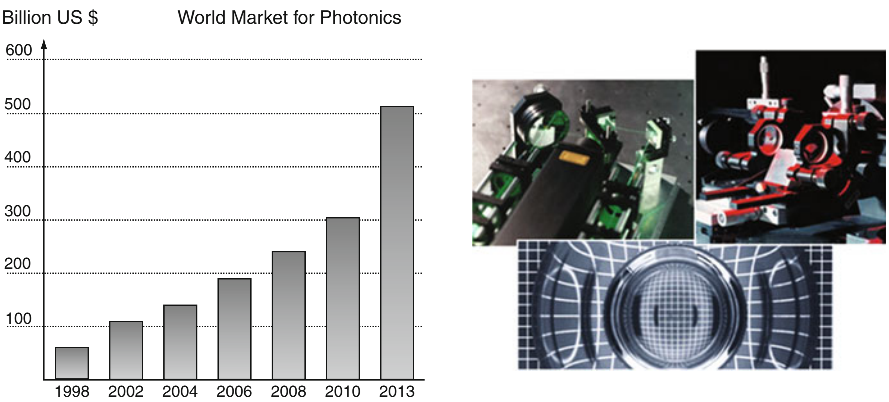


Fig. 6.80 Market forecast for optical technologies. (From von Witzleben 2006)

technology, etc.) and, therefore, these technologies are steadily gaining significance for the overall economic development of industrial countries (Fraunhofer IOF 2006) (Fig. 6.80).

Similar to the development of microsystems technology, through continuous miniaturisation and functional integration, optical systems are also opening up more and more new areas of application for non-optical components too. Essentially, these applications involve generating, manipulating or detecting light using optoelectronic components as well as components from electronics, actuator engineering and silicon micromachining (Beckert 2005).

From the perspective of the production technology polylemma, optical assemblies are characterised by evermore complex, multi-variant systems. At the same time, only small to medium unit numbers are usually required. Macrooptical systems are, for the most part, assembled and aligned in manual processes, whereby efficient execution requires many years of experience (Bauer 2008).

The classic system design is based on the so-called optical bench onto which the optical components are mounted in mechanical, manually adjustable manner. Optics, electronics and mechanics are generally treated as separate entities. These design principles cannot be transferred to miniaturised optical systems because they do not satisfy the requirements in terms of packing density, functional integration, robustness and unit costs (Beckert 2005).

As a result, there is a strong link between product design and the production process whereby the optical design of the assembly determines, in particular, the requirements on the assembly process. Miniaturisation restricts manual assembly to the extent that new solutions need to be developed especially in terms of flexible automation of the assembly of microoptical modules. In this respect, microoptical systems need integrative approaches that take an integrative view of product design, component manufacture and assembly.

To build complex hybrid microsystems with optical elements, in addition to an automation-friendly design, new and complementary technologies are required for joining and above all handling optical components (Bauer 2008). From a production engineering point of view, the joining methods used are decisive when it comes to different material pairings and, in particular, for opening up the third spatial dimension in assembly and alignment. Another key challenge is the precision required, which, depending on the application and component, is in the submicrometre to low two-digit micrometre range (or between low microrad and millirad) (Beckert 2005).

Hence, for further market penetration and to secure or recover jobs in high-wage countries, manual processes need to be replaced by flexible automation solutions for small and medium unit numbers (Photonics 21 2010). In this respect, based on their characteristics, the optical systems in question can be considered representative of a wide range of multi-variant, complex, miniaturised products.

6.5.2.3 Automation Solutions from Science and Industry

The relevant scope of automation technology for micro and precision assembly can be divided into the topics of handling and precision joining, which are explained in the following sections.

Handling The accuracy required for assembling optical systems is in many cases in the single-digit micrometre range and is often even in the submicrometre range. In the industrial world, huge efforts have been made to develop robots with similar positioning accuracy. Vertical articulated arm robots can achieve a positioning accuracy of approximately 50 µm (in some cases 20 µm) and SCARA robot kinematics, depending on frame size, can achieve down to 10 µm. The repeatability of Cartesian systems with linear axes, which are widely used in microassembly, is 1 µm for single axes with optical scales, but this is not necessarily the spatial positioning accuracy (Hesselbach et al. 2002).

Furthermore, manufacturers like PI and Micos produce high-precision single axes which are used in stacked systems for multi-axis alignment. There are also parallel kinematic concepts which allow multi-axis alignment with extremely high precision, but tend to only cover small working areas (e.g. Micos SpaceFab, PI HexAlign, PI M-810).

The high accuracy of precision robotics is associated with a high acquisition price. In many cases, for efficient assembly, large working areas are required to be able to feed the components and integrate the wide variety of sensors mentioned above. This increases the costs many times and sometimes is not even conceptually feasible.

Fully automated assembly systems are predominantly available for electronics production (pick-and-place, SMD, die bonding) or tailored to very specific applications (e.g. the Sysmelec SMS1000 for assembling microlenses and optical fibres). There are no flexible and cost-effective systems that allow integration of various modules from different manufacturers and therefore fast reconfiguration of peripheral equipment (Gramann 1999).

This fact reduces the availability of precise handling systems for laboratory automation, especially for young high-tech companies, which makes it impossible to exploit any potential for rationalisation in product development and series production.

In research, different approaches are used to achieve the required assembly precision (e.g. micromanipulators, microrobots). However, to date there is no consistent overall concept for modular systems. Prototype assembly systems developed in collaborative projects are always designed to facilitate assembly of a product.

The gripping of precision- and microcomponents has been investigated in various collaborative projects (e.g. German Collaborative Research Centre SFB 440 "Assembling hybrid microsystems") and various physical principles can be exploited to grip different components. For assembling microoptics, however, there are no solutions that address the specific requirements of these components: (1) to build a product, many different components need to be gripped and held, which means the gripping system needs to be flexible; (2) there is usually only minimal surface area available for gripping the component; (3) hybrid gripping, alignment and joining of components increase the demands on the robustness of the gripping principle and the process-safe holding of components in the aligned position.

Precision Joining of Microoptical Components To join microoptical components, a whole range of attempts have been made to make conventional joining methods from macroscopic production also suitable for microscopic applications. The feasibility of joining specific material pairings plays a central role in this respect. The German Collaborative Research Centres SFB440 and SFB356 focused in particular on the joining methods as micro-soldering, micro-electron beam welding and micro-bonding.

For optical components (non-metallic materials), soldering and bonding processes are most suitable. During assembly, it is crucial that these components can be aligned and fixed in the designated position. Using UV-cured adhesives, in principle, this kind of alignment is possible in the liquid adhesive provided a corresponding adhesive gap is created and maintained. In this regard and with respect to the state of the art, there has been no systematic investigations about the shrinkage behaviour of different adhesives and gap geometries, nor is it known for sure how alignment in the adhesive can be reliably and efficiently performed.

The company Coherent GmbH uses a patented soldering procedure (PermAlign) to assemble optics which requires relatively complex preparatory steps (Woods und Pflanz 2008). Various research projects are also striving to develop suitable soldering procedures.

Conclusion The state of the art is characterised by manual or semi-automatic solutions involving the use of highly specialised soldering or bonding technologies for individual components or assemblies. In terms of cost-efficient automation, the state of the art in micro and precision assembly should therefore be considered critical. While there are some fully automated systems already, most of them are designed for a specific purpose, i.e. product-bound, and are therefore inflexible and it would take considerable effort to convert them for use with other products. Automatic assembly units are associated with extremely high acquisition costs and the investment risk increases when sales figures are difficult to predict because this means it is not possible to guarantee that the amortisation point will be reached within the product life cycle.

Automated systems that guarantee improved cost-efficiency even with fluctuating unit numbers and a high variant-diversity are of significance, especially for small and medium-sized enterprises. Regarding the considerable growth potential of product technologies and applications, this is a promising starting point for new concepts and technologies which, based on the high demands optical assembly places on flexibility and precision, can also be transferred to other products and industries.

6.5.3 Motivation and Research Question

The Cluster of Excellence "Integrative Production Technology for High-wage Countries" addresses the challenges faced by manufacturing companies, especially those in high-wage countries. Given the complex competitive situation in the globalised market, many companies are struggling to devise a strategy for meeting the increasingly dynamic and complex demand. Unification of the strategies of differentiation and price leadership has not been achieved until today due to irreconcilable differences between economies of scope and economies of scale as well as between plan and value oriented production.

The Cluster of Excellence is developing contributions to production technology which, based on their integrativity, i.e. the integrative view and thus also the interdisciplinary solution of tasks, contribute to reducing both dichotomies mentioned above.

The key question in this chapter is how complex, multi-variant products can be efficiently manufactured in high-wage locations. The focus is set on assembly, which is increasingly becoming a key value-adding element and thus a core competence of the company. Assembly is also the last value-adding stage through which direct proximity to the market is influenced the most by a dynamic environment. In this respect, strategic positioning between the opposing corners of the production polylemma is possible. Figure 6.81 shows the corners on the relevant axes.

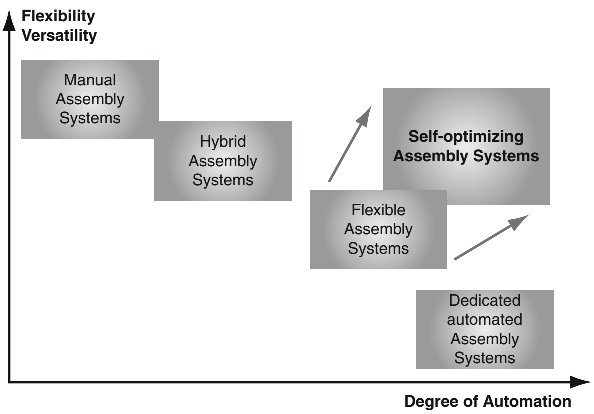


Fig. 6.81 Classification of assembly systems. According to (Weck 1998; Andreasen and Ahm 1988)

Automated assembly systems are efficient for high unit numbers and stand for economies of scale in mass production. Most processes have to be planned in advance and explicitly implemented because of the considerable adaptation work required, which means it is practically impossible to implement changes in the production process at short notice. In contrast, manual assembly is extremely flexible and copes better with variant diversity and fluctuations in demand. Value-oriented control mechanisms are increasingly being used in order to improve efficiency. The strategy of differentiation and the related economies of scope are, however, required because of the strong influence high wage costs have in international competition. As a compromise, hybrid or semi-automatic as well as flexible assembly systems are found between the extremes of the polylemma.

Concepts for self-optimisation should be implemented to create modern assembly systems that combine high productivity with high flexibility by reducing the effort required for planning, configuration and process ramp-up in flexibly automated systems. Self-optimisation offers solutions in which system adaptations are autonomously planned and implemented in order to respond to changed conditions.

The following section presents approaches for self-optimising assembly automation which help to reduce the production polylemma.

The technological focus of this chapter focuses on the assembly of complex optical systems which, as explained in the introductory sections, place the highest demands on the flexibility and accuracy of the assembly system whilst at the same time representing a wide variety of products that combine evermore and increasingly heterogeneous functional elements in the smallest of spaces in line with the trend towards miniaturisation. The basic research question is therefore:

> How can high-tech products be assembled cost-efficiently in high-wage locations?

This chapter investigates how self-optimisation can help create flexibly automated systems that enable efficient production of complex, multi-variant products. The research aims to develop concepts for automated assembly of a miniaturised laser system (see also Sect. 7.3) and, on that basis, aims to demonstrate and study cognitive and self-optimising elements in industry-related applications.

6.5.4 Results

Based on a basic definition of the term self-optimisation, the following sections will first consider the technical and conceptual bases for self-optimisation in automated assembly systems. In light of the research question, the various aspects of flexibility and adaptability, autonomy as well as cognition and learning ability are brought into the context of self-optimising assembly systems.

Then Sect. 6.5.4.2 presents concrete solutions for the automated assembly of miniaturised laser systems which were developed in the project "Self-optimising Flexible Assembly Systems". One focus of the deliberations is autonomous planning in assembly cells (Project: "Model-based Assembly Control").

The concepts and methods developed to facilitate an integrative procedure for designing product, process and production equipment are presented in Sect. 6.5.4.3 using the example of a miniaturised laser system. Detailed information about the development of the laser system (MicroSlab) is given in Sect. 7.3, “Integrative Production of Micro-lasers”.

6.5.4.1 Self-optimising Assembly Systems

With respect to the production polylemma, self-optimising assembly systems are seen as a means of alleviating the polylemma by allowing short-term adaptations to current conditions and objectives to be made with minimal planning, reconfiguration and change-over effort. In this respect, self-optimisation is to be understood as the ability of a system

- to continuously analyse of the current situation and
- to derive system objectives based on that analysis and then
- to autonomously adapt the system behaviour accordingly.

The first point ensures that the current state of the system and its environment is identified. The information obtained is then used to determine the new objectives of the system (point 2). These objectives should bring the system to an optimal operating point and they can be either selected, adapted or even newly generated. Once the new system objectives have been determined, the system behaviour is adapted to the new conditions (point 3). In layman’s terms, self-optimisation means “thinking whilst doing” instead of “thinking before doing”.

According to a definition in the German Collaborative Research Centre SFB614, self-optimising systems have the ability to respond autonomously and flexibly to changing environmental conditions, user intervention and actions of the system. Behaviour is optimised based on machine learning (Frank and Gausemeier 2004).

The basic prerequisites of an assembly system for implementing self-optimising mechanisms are therefore flexibility and mutability (as the basis for adapting system behaviour) and autonomy (for mastering complex processes robustly and without human intervention) (Fig. 6.82).

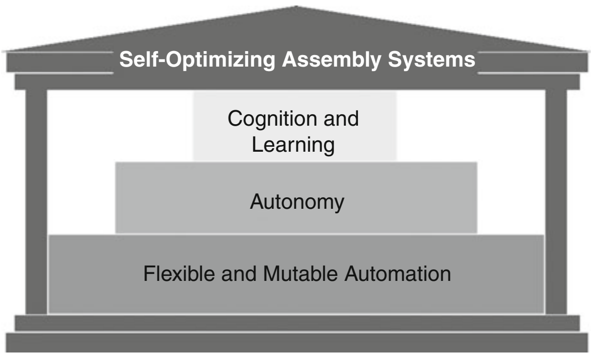


Fig. 6.82 The building blocks of self-optimising assembly systems

A self-optimising system is able to pursue new objectives based on dynamically changing boundary conditions through the use of automatic planning and optimisation functions or through the integration of cognitive structures without focussing on the planning effort of the system and without having to suffer corresponding losses in efficiency and profitability.

The next section first discusses the three main aspects of flexibility/mutability, autonomy and cognition as the fundamental basis for designing self-optimising assembly systems before presenting concrete implementation concepts for producing laser systems using an automated assembly cell.

Flexible and Mutable Automation In the context of assembly, flexibility refers to the ability to adapt an assembly system quickly and with little effort to changed influencing factors. With regard to flexibility, organisational and technical changes are limited at the time of planning by defined corridors (Abele et al. 2006). Hence a change in the number of units, for example, can be represented to a pre-defined extent within these corridors. Mutability, on the other hand, refers to the potential to implement possible changes on a reactive basis beyond the defined corridors (Reinhart et al. 2002). So, when planning mutable assembly systems, no explicit limits are set. Hence, the systems are largely solution-neutral; the necessary scope for possible changes is therefore taken into account (Nyhuis et al. 2008). Only by simultaneously increasing flexibility and mutability is it possible to meet the growing dynamics of turbulent market conditions (Spath and Scholtz 2007).

The influencing factors that lead to turbulence in the assembly system come from a number of external sources (e.g. the market, society, politics, technology, etc.) or from the production system itself. They affect assembly via a certain number of channels. The influencing factors, also known as change receptors, include:

- products and product variants,
- costs,
- time,
- number of units,
- quality and
- the elements of the assembly system (process technologies, system technology, tools, etc.).

All the requirements on flexibility and mutability can be traced back to changes of one or more of these receptors (Cisek et al. 2002). This results in various requirement categories which can be summarised as

- component and product flexibility,
- capacity and utilisation flexibility,
- process and technology flexibility and
- planning flexibility through connection to business data flows.

Component and product flexibility comprises the ability of an assembly system to adapt to frequent variant changeovers and to further/new product developments. Capacity and utilisation flexibility (or scalability) focuses on adaptability to varying unit numbers caused by fluctuations in demand or the evolvement of production volume within the product life cycle. Process and technology flexibility describes an assembly system's capability to apply different handling processes and, in particular, different joining processes depending on the component to be assembled as well as its capacity to use need-based measuring and testing technologies (Kurth 2005). A pre-requisite for planning flexibility is that the state of assembly is always transparent in the information systems, allowing the impact of short-term changes to be assessed. That way, the costs, time and quality parameters can be balanced out in direct market interaction by, e.g. by applying customer-specific pricing based on the variant and lead times (Westkämper et al. 2001).

To meet these requirements and to create a basis for reacting quickly to changing environmental conditions, an assembly system must have so-called change enablers which describe the necessary properties of the system elements. These include (Nyhuis et al. 2008):

- *Modularity*: modularity describes the ability of an assembly system to exchange standardised, functional units or elements (modules);
- *Compatibility*: compatibility between system elements allows individual elements to form networks with one another and interact in various configurations through uniform interfaces (mechanical, electronic, data-related);
- *Universality*: universality of system elements describes the dimensioning and design of these elements for various requirements and application areas;
- *Mobility*: mobility refers to the spatially variable usability of system elements;
- *Scalability*: scalability means the technical, spatial and personnel-related extensibility and reducibility of an assembly system.

Figure 6.83 summarises the information about the flexibility and mutability of assembly systems.

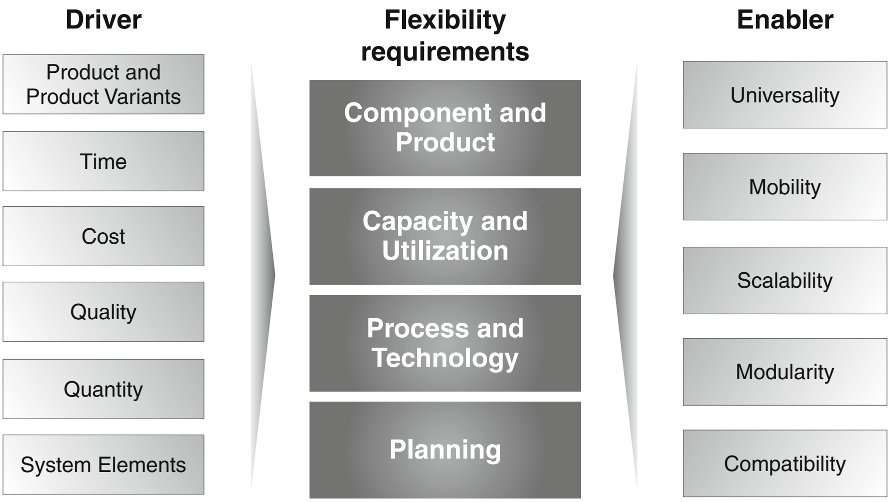


Fig. 6.83 Change drivers and change enablers

Flexibility and mutability create the degrees of freedom required for short-term re-planning and reconfiguration of assembly systems and therefore provide the basic foundations for self-optimising systems. Additional degrees of freedom lead first to increased system complexity and then to greater complexity in the planning and control of highly flexible systems. To cope with this complexity and get closer to the basic goal of self-optimisation with a view to reducing planning effort, additional components are also needed.

Autonomous Assembly Cells With respect to an assembly cell, autonomy means the capacity for reliable and error-free execution of complex assembly processes for a longer period and with a maximum degree of independence. Autonomy enables the system to respond proactively to environmental factors of influence (Pfeifer and Schmitt 2006). The need for an assembly system to offer the greatest level of autonomy possible is due to the fact that with the increasing depth of process automation, the number of process faults also tends to increase. In addition to producing a deterioration in process and product quality, this means that in order to remove these faults, time-intensive intervention by specialist staff is required and that staff must also be capable of coping with the increasing requirements caused by the growing complexity of processes and systems. For the transition from rigid automation solutions to flexible, fast and efficient assembly systems, the following three sub-aspects of autonomy can be identified:

- *Autonomy for the user through optimised support* The assembly system provides the user with a variety of functions and information and allows the user to intervene in the assembly process and in the planning of the next assembly steps;
- *Autonomy through fault tolerance* The assembly system has functionalities that enable it to react autonomously to faults caused by parameter fluctuations and other influences. Faults can be rectified autonomously if there is a high level of integration of the various system modules and fault rectification strategies can be planned;
- *Autonomy through closer integration of sub-functionalities* The assembly system benefits from an extended range of functions through the automation of all sub-operations of handling, joining, aligning, measuring and control as well as through additional services for process planning, monitoring and control.

Based on the key requirement of enhancing the responsiveness of automated assembly cells in the event of fluctuating influencing variables and making efficient use of the increased flexibility, there is a need for new functions along with new hardware and software concepts in the control system. One essential element in this respect is an information and communication technology infrastructure for connecting a variety of actuators and sensors in the form of a flexible and extendable sensor/actuator network. Furthermore, a self-optimising assembly cell must also have significantly expandable planning functions if it is to achieve the desired degree of autonomy.

Automated Planning in Flexible Assembly Systems The task of cross-process planning at the higher levels of production systems involves scheduling and coordinating the production and assembly orders that result from customer or planning orders and reacting to process-related changes caused by re-planning activities. To meet the needs of increasingly complex planning as result of dynamic influences, planning efforts have to be purposely decentralised, i.e. moving them to autonomous subsystems coordinated centrally (Pfeifer and Schmitt 2006).

Hence, flexible assembly systems are organisationally and technically divided into decentralised units (in automated systems: into autonomous assembly cells) which carry out process-related planning and control. If numerous cells have to work together to assemble a product, these cells are coordinated by a central entity. Coordination or even assembly planning includes allocating individual tasks to the subsystems. These tasks can be described in the form of complete programs that can be loaded into the cell control system for direct execution. This, however, involves major planning effort in the central unit and the result can usually only be used for a predefined cell configuration. Optimisation of process execution is usually done offline. A steady flow of new and changing tasks calls for frequent planning (diversity of variants, shortened product life cycles). Taking uncertainties into account when deliberating the plans hugely increases the workload and even hinders automation. If conditions vary, plans need to be changed and this delays execution. Program creation and optimisation (as part of planning) is time-consuming and requires a great deal of expertise, yet parts of programs can rarely be re-used.

What is considerably more flexible is an abstract, task-oriented interface between assembly planning and the executing cells, which are translated into device-related actions within the cells. Hence for an automated assembly cell, planning means determining the execution of a defined task. The result is a plan that describes a chronologically ordered sequence of actions. A complete plan contains the following aspects of the action description (Siegert and Bocionek 1996):

- Action planning (What?)
- Sequence planning (What is the temporal relationship?)
- Resource planning (Who?)
- Timing (When?)
- Execution/process planning (How? With which parameters?)

For automated planning, the cell control system requires a knowledge base with an environment model (cell model) and with rules regarding how a task is to be broken down into individual steps. At each planning level, only the section of the world relevant to the problem should be considered as this limits the number of states and therefore the planning effort.

In addition, synchronisation patterns are required for coordinating the activities and various algorithms are needed for planning specific device actions (grip planning, path planning, sensor integration). In planning, one distinguishes between the following procedures (Siegert and Bocionek 1996):

- *Goal-oriented planning*: In a perfect world, an overall plan would be created and executed without the actions being monitored.
- *Reactive planning*: A complete plan is created for an assumed world model. During execution, if action monitoring detects any deviations, a new plan is created or the goals are changed.

- *Opportunistic planning*: Plans are created for alternative scenarios and all goals are pursued simultaneously. In the next step, the plan promising the greatest success is selected based on the current state of the system.
- *Reflexive planning*: This involves an event-driven, step-by-step plan based on the actual data to hand with few forward-looking assumptions.

In conventional automated systems, planning is usually detached from execution. This basically corresponds with goal-oriented planning and results in rigid systems with a lack of responsiveness. During execution, the other types of planning mentioned above reflect the principle of self-optimisation: analysis of the current situation, determination of system goals, adaptation of system behaviour.

For self-optimising automation, it must be possible to move planning scopes, in the sense of reactive, opportunistic or reflexive planning, into the control system (depending on the planning task different alternatives should be preferenced).

Automated planning in flexible assembly systems presents a challenge because of the special conditions. As the last link in the production chain, assembly is constantly exposed to changing quantities, variants and products. Thus, assembly depends on a lot of parameters of the prior production steps. As already mentioned, the difficulty is that not only faults and problems of the prior production steps have a negative impact on assembly precision or even the feasibility of assembly, but changing production tolerances and production processes also influence assembly. This reaffirms the need for reconfigurable and adaptive assembly systems.

Automation technology has been helping to make production processes faster, more precise and reproducible (Weck and Brecher 2006). At the same time, it has become a limiting factor. This is because the increasing complexity of the processes, and therefore also the ever-higher degree of subsystem networking, is confronted with the impermeability of different control domains (e.g. NC, RC, MC, PLC) combined with "weak" programming paradigms. Today's automation technology is well suited to reliably facilitating well-specified processes with short cycle times. Varying requirements or boundary conditions, however, are difficult to depict during process time and call for time-consuming planning processes or the direct intervention of human experts. Furthermore, the traditional approach hardly allows production and assembly processes to be viewed across multiple subsystems.

Cognition and Learning Ability in the Control Technology Cognition enables logical thinking about the relevance of the information acquired, allows planning and learning processes to be devised based on individual experience and facilitates intelligent decision-making as well as reliable and robust system behaviour (Beetz et al. 2007). The processes of logical thinking, learning and decision-making can be performed both on the basis of statistical data analysis and using the techniques of artificial intelligence.

Russell defines weak artificial intelligence as machines behaving as though they were thinking. In fact, they are not thinking, they are following a set of rules that can be as complex as desired. If these rules can be extended and adapted independently from the system during run time, then this is called machine learning (Russell and Norvig 2003).

Intelligent systems with the ability to learn are used in order to make decisions about their actions independently. Here we see a direct reference to the autonomy mentioned above. The better the decisions a system makes independently, the greater the degree of autonomy.

One difficulty in implementing intelligent systems is the wide variety of actual processes and process environments. This makes finding a solution highly complex. In most cases, not only is the solution path of interest, but also the associated costs. Variables, such as time duration, energy consumption, can be interpreted as costs in this respect. For example, efficiently programming complex automated systems is a major challenge. Only automatisms can drastically reduce programming effort. Such automatisms make use of the tools of artificial intelligence to determine optimal action plans, for example, or generate specific paths for robots. Using learning techniques such as neural networks, the automatisms can be refined independently of the system. Systems that able to achieve this to the full can presently only be found in the highly specialised domains with a very limited scope. The field of artificial intelligence remains an important research topic. Practical solutions designed to fulfil the requirements described with respect to flexibility and autonomy are of particular interest.

The means of cognition that an assembly system must possess for self-optimisation may, in the simplest case, be based on experience and the process data acquired may simply be statistically evaluated. In its most advanced state, however, cognition may take forms of artificial intelligence that are able to not only evaluate experiences, but also to make judgements and draw conclusions. In this respect, heuristics constitute another basis for self-optimisation whereby alternative decisions are specified and the system can chose between them.

Self-optimisation can be applied at various hierarchical levels within an assembly system. Planning and system configuration represent the top level while the lowest comprises the individual assembly processes such as joining a component. Ideally, self-optimisation should include all levels, allowing the entire assembly process to be adapted without human intervention. Should the boundary conditions or objectives change, the system autonomously re-configures itself, uses other joining technologies, measuring technology or handling units and changes the assembly and alignment process. At lower levels, individual processes optimise themselves: initial values for alignment tasks are adapted as the system learns in order to reach the desired result faster, join parameters are optimised in order to minimise deformation and misalignment and to allow faster assembly cycles. The changes made flow back to the user as information and the knowledge gained is fed back into the development process so that the necessary changes can be planned, assessed and implemented. In this respect, cognition presupposes the existence of flexibly deployable sensors which collect the necessary information, evaluate it and store it in a knowledge base so that situation-dependent decisions can be made. These decisions are converted within the system into corresponding control commands and fed back into the process via various actuator components.

Interim Conclusion The information above leads to the following understanding of the concept of self-optimising, automated assembly cells:

- Self-optimisation must mean that commands at task level can be transferred to the assembly cell control system either by the operator or by direct coupling to central planning systems, disregarding concrete hardware and solutions;
- The cell control system takes over the planning responsibilities, i.e. it independently breaks down the tasks into individual, device-related actions and generates the associated program code;
- The execution of actions is monitored and necessary changes to the plan are automatically implemented during run time (reactive/opportunistic/reflexive planning);
- Knowledge (learning effects) is gained from execution of the actions and taken into account in new planning operations (cognitive control).

From this it is possible to derive a variety of fields of research regarding reducing the production polylemma in assembly. The concepts and solution approaches presented in the following section are based on the results of the research projects "Self-optimising Flexible Assembly Systems" and "Model-based Assembly Control" conducted by the Cluster of Excellence "Integrative Production Technology for High-wage Countries" based at RWTH Aachen University. The work focuses on the following aspects:

- Increasing the flexibility and mutability of assembly systems along the entire process chain in order to improve efficiency and profitability when producing a broad product mix with low unit numbers;
- Increasing the degree of automation in flexible assembly through innovative, versatile assembly techniques and measuring equipment, by developing set-up and change-over procedures that can be automated and by designing open, integrated equipment control systems;
- Developing concepts for mapping and using planning systems and cognitive structures in assembly systems for self-optimising automation;
- Improving system transparency throughout the product creation process through systematic analysis of the dependencies within the process chain, from product development to process design to manufacturing equipment development.

6.5.4.2 Flexible Assembly Cells for Self-optimising Production

The automated assembly of complex products or diverse product variants requires conception of a flexible assembly system capable of continuously adapting to the production parameters such as quality, time and costs (Fig. 6.84). In some cases, various components and product variants can be assembled similarly, yet still have different features and this requires a change in the assembly system behaviour. Accordingly, the assembly system must demonstrate self-optimising behaviour by continually monitoring its current state and independently adapting to the new circumstances.

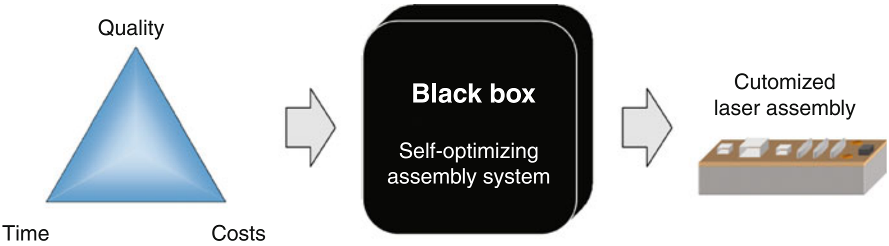


Fig. 6.84 Quality, time and costs as input parameters for self-optimising assembly

A highly flexible assembly system requires a completely modular equipment design with standardised interfaces for integrating various handling, joining and measuring technologies. Providing flexible assembly techniques in the form of interchangeable modules and developing automated change-over and calibration procedures as well as an open, integrated control system increases the degree of automation of flexible assembly systems and thus enhances the productivity and profitability of assembly. With respect to assembly control, solutions need to be developed to embed the decentralised control units of individual modules, especially of cooperating robots, in one central control system. This supports flexible order planning, thus delivering the prerequisite for self-optimising operation.

The following sections present solutions for designing a modular precision assembly system along with flexible and configurable equipment control systems as well as for automated planning of complex systems. The solution approaches focus on the self-optimising assembly of miniaturised laser systems.

Modular Multi-robot Cells for Precision Assembly As described in Sect. 6.5.2, as areas of assembly technology, micro and precision assembly are characterised by the component dimensions and precision requirements involved. Both of these factors play an important role when it comes to designing automation solutions for precision assembly. They produce the following overall set of requirements and tasks:

- Manipulation of objects in the micro[μm] and meso[mm] range;
- Setting up extremely precise spatial relationships between components (alignment);
- Sensor guidance of assembly processes;
- Monitoring and control of contact force;
- Execution of various physical and chemical processes (gluing, soldering, bonding, etc.);
- Linking/hybridisation of alignment and joining process.

It is in essence possible to derive design guidelines for precision assembly systems against the background of a high degree of product-specific operations and customer-specific adaptations to a wide variety of applications (Siciliano and Khatib 2008):

- Holism: a precision assembly system is composed of a wide variety of functional units (robots, manipulators, cameras, sensors) and their interaction must be considered integratively.
- Integrativity: although the handling technology in micro and precision assembly presents particular challenges with respect to robots and grippers, integration of the joining process is especially important for system design.
- Reconfigurability: as a basic requirement, reconfigurability can be achieved with a consistent, modularised system design that supports automated module changeover with respect to tools.
- Tolerance management: setting and complying with tolerances in the spatial relationship between components is the decisive factor for designing an automated system for precision assembly. One of the key aspects here is the purposeful shortening of tolerance chains in the system through referencing and closing local control loops. The precision that can be achieved then depends on the resolution of the sensors used and the movement resolution of the positioning units.

Fully automating the laser system assembly places highest demands on precision, but also on the flexibility and mutability of the assembly system. Hence, different component-dependent handling and joining processes, on the one hand, and need-based measuring and testing technologies, on the other, are used for building the MicroSlab. In order to implement the resulting variety of assembly tasks in one assembly cell, a modular system concept has been developed based on the use of multiple cooperating robots with standardised interfaces for interchangeable tools.

The basic idea of the system concept is to maximise the flexibility and mutability of an assembly cell by ensuring system modularity right down to tool level and thus reducing the configuration and set-up effort for the various process steps to a minimum.

During assembly, industrial robots take over positioning of the modular tools in the working area. These tools are used for gripping, fine-positioning, joining or measuring depending on the process step. For more complex assembly operations whereby components are actively aligned and a joining operation is performed at the same time or directly afterwards, several tools are used with three cooperating industrial robots (Fig. 6.85).

For joining optical components, a novel resistance soldering process is used which, based on the development of an automated joining module, can be transferred to a robot-based soldering process (see Sect. 7.3).

Since the robots do not provide satisfactory precision and step sizes for aligning certain optical components of the MicroSlab, an additional micromanipulator was developed providing the necessary precision. In line with the concept of a modular mounting head, this micromanipulator is designed for mobile use attached to conventional industrial robots to compensate for their positioning errors and to perform the subsequent high-precision alignment of components (Fig. 6.86).

This produces a seamless link between macro and micro working areas and therefore reduces the dichotomy, common in micro and precision assembly, between high precision and large working areas.

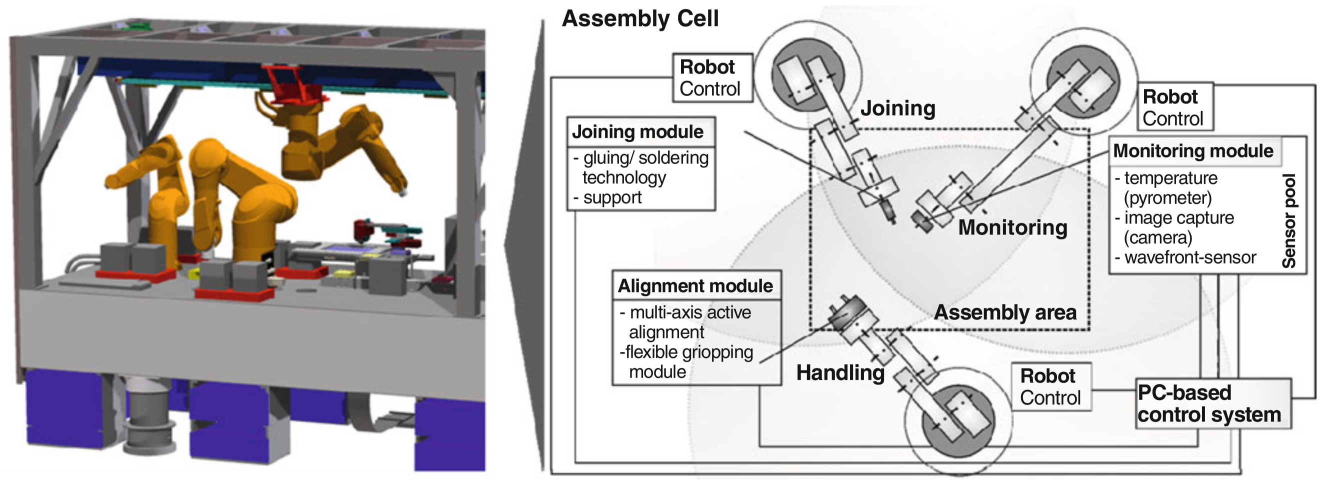


Fig. 6.85 Modular multi-robot cells

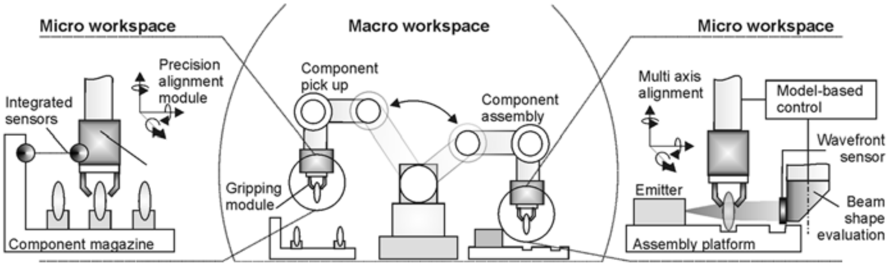


Fig. 6.86 Concept for integrating macro and micro working environments

Since the alignment of some components requires movement in all six degrees of freedom, the micromanipulator is based on a hybrid kinematic concept which meets these requirements while enabling the smallest increments in the micrometre range. The concept involves a parallel-kinematic structure whereby the movable end effector is connected to the frame via three struts of invariable length. The base points of the struts are mounted so as to allow movement in two translational degrees of freedom in the plane, resulting in six degrees of movement on the platform altogether.

All the passive joints of the kinematic structure are composed of special flexure-based joints which offer superior properties compared with conventional notch or leaf spring joints. Basically, with respect to the precision and movement resolution of a system, flexure-based joints offer advantages in that there is no friction or play and they are less susceptible to wear and tear. In theory, these advantages imply that the accuracy of the movement of the platform is ultimately determined solely by the movement behaviour of the drives, i.e. the smallest increment sizes correspond with the computational resolution of the system.

In the context of the metrological characterisation of the mounting head, it was possible to prove that the movement resolution of the drives can be transferred to the platform via the elastic structure without losses. Interferometer measurements show that the drive movements, including control oscillations in the single-digit nanometre range, are transferred to the platform and can be measured there. Unidirectional repeatability measurements with the same measuring setup yield a characteristic value of: 0.15 μm repeat accuracy (40 measurements with a standard deviation of 50 nm). The associated translational working area is approx. 4 mm $\times$ 4 mm $\times$ 4 mm (Fig. 6.87).

A vacuum gripper is used for handling components. For a highly accurate and reproducible grip, this gripper has multiple reference surfaces to which the components automatically align themselves when they are vacuum gripped.

With respect to assembly control, solutions need to be developed within the context of the system concept that will allow decentralised control units of individual modules of this type to be embedded in one central cell control system that supports hardware configurability.

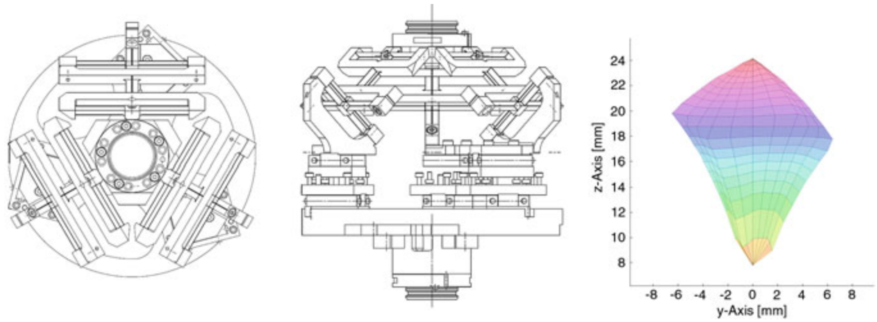


Fig. 6.87 Micromanipulator based on flexure-based joints (wireframe graphics and working area)

Agent-Based Control of the Assembly System To achieve the afore-mentioned flexibility, autonomy and cognition aspects desired in the assembly system, a concept for integrating distributed hardware and software platforms using agent-based technologies is used.

Modelling the control of an assembly system using an agent-based structure aims at the optimal utilisation of the assembly resources (Russell and Norvig 2003; Jennings et al. 1998). Although there are various definitions for the word “agent” in the literature, it is generally understood that an agent “is an entity that perceives its environment through sensors and can provide feedback to this environment by means of actuators” (Russell and Norvig 2003). The agent autonomously seeks one or more goals and usually remains in close contact with other agents in order to achieve these goals. Before the agent decides for a particular action, it takes into account its perception of the environment, its goals, its knowledge and its experience (Fig. 6.88).

The development of software agents focuses on the decomposition of a problem and the distribution of responsibilities to smaller, autonomous entities that follow

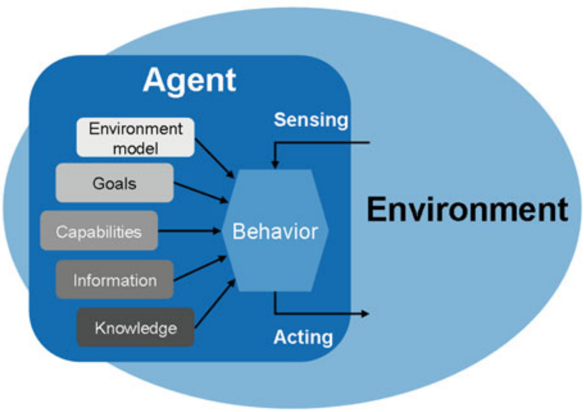


Fig. 6.88 Model of a software agent

certain principles, e.g. encapsulation, goal orientation, reactivity, autonomy, proactivity, interaction, persistence, adaptivity and intelligence or ability to learn, in order to accomplish their locally or globally distributed activities (Jennings et al. 1998; Göhner et al. 2004).

Within flexible production environments, at least three positive factors are introduced by using agent-based control structures (Brecher et al. 2009):

- The structure of an agent-based system allows an easy introduction of new agents (software or hardware) to the system without needing to re-programme the control logic;
- The application is based on a distributed system that allows the use of multiple operating system platforms and ensures communication between different hardware systems via a common medium according to a predefined communication protocol;
- Cooperation or competition between the various agents is the main cause for the desired operational autonomy of the production system.

The agents of a multi-agent system vary in terms of their abilities and responsibilities. There is, however, no specific command hierarchy between the agents from a functional point of view. In principle, all agents work on the same software level and exchange relevant information and services in order to achieve their goals. Nevertheless, from an organisational point of view, in terms of modelling and designing an agent-based assembly control structure, it is valuable and aids comprehension if the agents have a specific hierarchy. The hierarchy highlights the level of difficulty of implementation and the importance of the agent's functionality. These aspects are explained in the next section.

Agent-Based Control Structure of the Assembly System The self-optimising behaviour of the flexible assembly system can be divided into various levels or control loops. Figure 6.84 shows the highest level of self-optimisation whereby a dynamic strategy for laser assembly is derived directly from the production parameters of quality, time and costs defined by the customer.

The following extreme scenarios better illustrate this idea:

- *Assembly of a high-quality laser system:* In this case, the criterion quality is set as the highest priority while the time and costs of assembly are disregarded. Thus, the assembly sequence is carefully planned and the most suitable optical components as well as the most appropriate handling, positioning and joining processes are carefully selected. All necessary quality characteristics of the laser are metrologically assured and the optical components featuring critical positioning tolerances undergo active (in-process) alignment procedures;
- *Fast assembly of an inexpensive laser system:* By contrast, in this case the criteria time and costs are set as the priority while the quality is disregarded. Hence, the time allowed for planning the assembly, selecting the optical components and for handling, positioning and/or joining is reduced. The metrological systems are only used when necessary. To save on costs, fewer optical components are used where possible as long as the laser system remains operational.

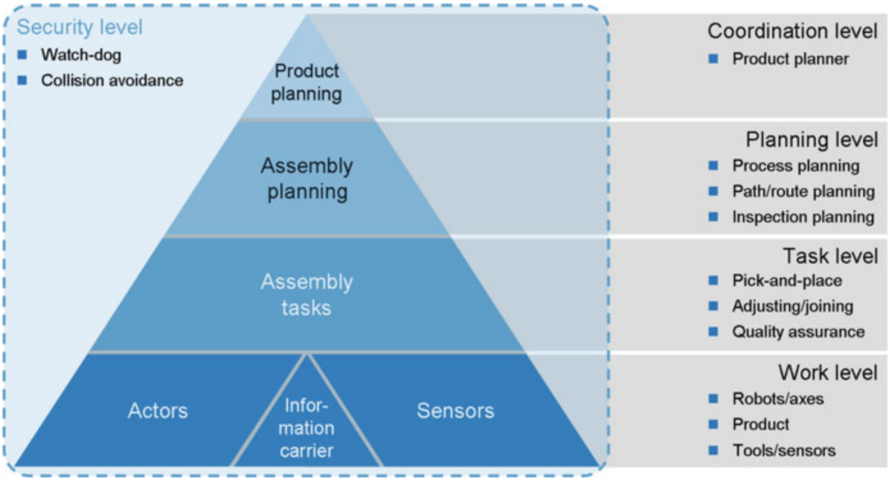


Fig. 6.89 Agent-based control structure of the self-optimising assembly system

Other assembly scenarios can also be depicted by choosing other concurrent configurations of the variables quality, time and costs. For the highest level of self-optimising assembly to take effect, the small, inner self-optimising control loops of the system must also be working. In this case, the black box element of assembly is conceived as a multi-agent system, whereby the cooperative/competitive exchange process between all agents involved creates the highest level of self-optimising system behaviour. Within the control structure, some agents are involved in reduced work groups in which smaller, self-optimising control loops operate. Figure 6.89 shows the overall agent-based assembly control structure, which is distributed across various levels (Brecher et al. 2009). Examples of tasks and agents that appear on different hierarchical structure levels of the system with different responsibilities and capabilities are as follows:

- *Coordination level:* Here, a product agent has high-level responsibility for product assembly planning. The desired production parameters quality, time and costs, as specified by an operator or user of the system, are incorporated directly here as input. Therefore, a strategy for assembling the product in accordance with the currently available resources is dynamically devised and orders are forwarded to the other lower-level agents in the control system. If no suitable strategy for assembly can be found because the combinations of production parameters are not viable, then a new, viable combination of parameters is suggested;
- *Planning level:* Here, several agents are responsible for optimised assembly planning based on the orders from the coordination level. Planner agents cooperatively define which and how assembly tasks should be carried out, in order to ensure that the expectations of the product agent are met. They translate the abstract assembly goals defined by the product agent into new, smaller and more specific orders for the task-level agents;

- *Task level:* Here, several agents are responsible for the interface between planning and working levels. Task agents translate the specific planned task orders into hardware and software commands that are to be executed by the work agents. They coordinate and assess whether the desired goals have been achieved after the task has been executed;
- *Working level:* Here, several agents are responsible for controlling the hardware and software modules (e.g. actuators and sensors). Some are also responsible for managing and providing information about the current state of assembly and of the product;
- *Safety level:* Safety agents are responsible for the integrity of the assembly work as well as of the agent-based system itself and avoid dangerous situations.

From this description of the responsibility assigned to each level of the control system, it is clear that the upper levels (task, planning and coordination levels) have a more complex level of design and implementation. These agents are characterised by intelligent capabilities (planning and coordination) and therefore need cognitive means in order to weigh up their actions and safely and robustly accomplish their tasks. The lower-level agents do not usually need these cognitive means and behave reactively (receive order, execute) to complete their activities.

The organisation of the various agents as well as their communication within the agent-based system is in accordance with FIPA[3] standards. Hence, FIPA-compatible tools are used for modelling (FIPA AUML[4]) and implementation (JADE[5]) of the system.

Top-Down and Bottom-up Approaches for Designing the Control Structure Based on these preliminary considerations, the control structure of the system for assembling the miniaturised laser was designed using a top-down approach. The behaviour of the entire system and the relevant smaller, inner assembly control loops were derived from the highest degree of self-optimisation via the coordination level. There are already important, inner, self-optimising loops in the planning level (e.g. assembly planning, path planning, test planning (Brecher et al. 2009)). Since the results generated during the assembly process do not always correspond with those that were predicted in the planning phase, these planning activities must be adapted based on the current state of assembly (Fig. 6.90).

Up to the planning level, the execution of agent activities is software-intensive and based on simulation and classification methods. From the task level down, these activities are closely related to the hardware control system, e.g. for component handling, positioning, aligning and joining.

Although the conception and rough modelling of the control architecture of the assembly system follows a top-down approach, the fine modelling and implementation of individual agents is developed using a bottom-up approach, because the lower-level agents constitute the basis of the assembly work and are essential for

[3] FIPA: Acronym for "Foundation for Intelligent Physical Agents" (http://www.fipa.org/).

[4] AUML: Acronym for "Agent UML" (http://www.auml.org/).

[5] JADE: Acronym for "Java Agent DEvelopment Framework" (http://jade.tilab.com/).

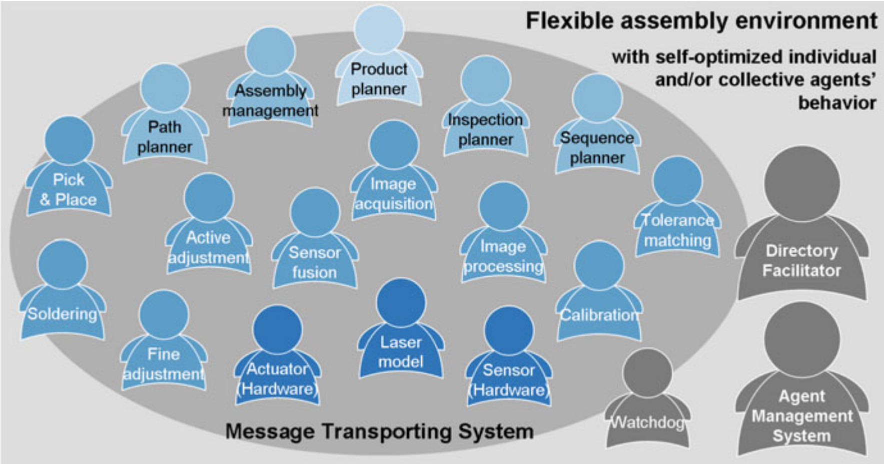


Fig. 6.90 Agent-based organisation for self-optimising assembly. (According to FIPA standards)

the minimal functionality of the system. This also includes their lower behavioural complexity in comparison with the planning agents.

The autonomy of the assembly system can only be achieved through the additional intelligence of machines (intelligent agents), which have the ability to perceive their environment using sensors and to process the obtained information in a goal-oriented way for fulfilling their task. Accordingly, sensor integration, intelligent controls and generic structures are essential for production systems of this type. On the basis of such formal descriptions, planning and optimisation components can determine action plans that are implemented by the control system or a management system (Fig. 6.91).

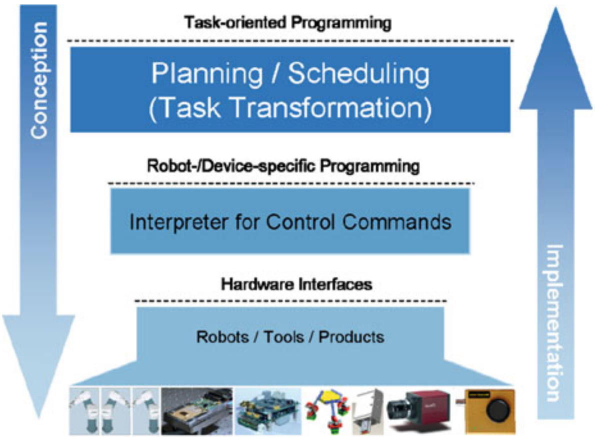


Fig. 6.91 Top-down and bottom-up approaches to the conception and implementation of the system

Agents for Dealing with Failure States The quality of the laser system depends not only on the quality of the individual optical components, but also on their positioning and correct pairing. In this respect, alignment precision has to meet the highest requirements (µm range). Deviations in component properties lead to new dynamic setpoint values for controling the assembly process. Hence, the entire assembly process must be flexible and supported by measuring technology for assessing the state of the system. Furthermore, the measuring technology should also be used for monitoring the entire assembly process, and active alignment of the optical components should be possible.

Depending on how the key performance indicator "quality" of assembly for a laser variant is defined, different laser system assembly steps must be metrologically assured. For a high-quality assembly, various laser characteristics have to be measured and tested:

- Geometric characteristics;
- Presence/absence and identification of components;
- Compliance with positioning tolerances;
- Characterisation and dynamic selection of components;
- Active (in-process) alignment of components with critical positioning tolerances;
- Characterisation of the laser beam profile and laser output power;
- Identification and resolution of failure states.

A number of measuring methods are used and integrated into the robot cell or the robots in order to ascertain the state of assembly (Fig. 6.92):

- robot-based, high-resolution image processing systems;
- a camera-based laser beam analysis system;

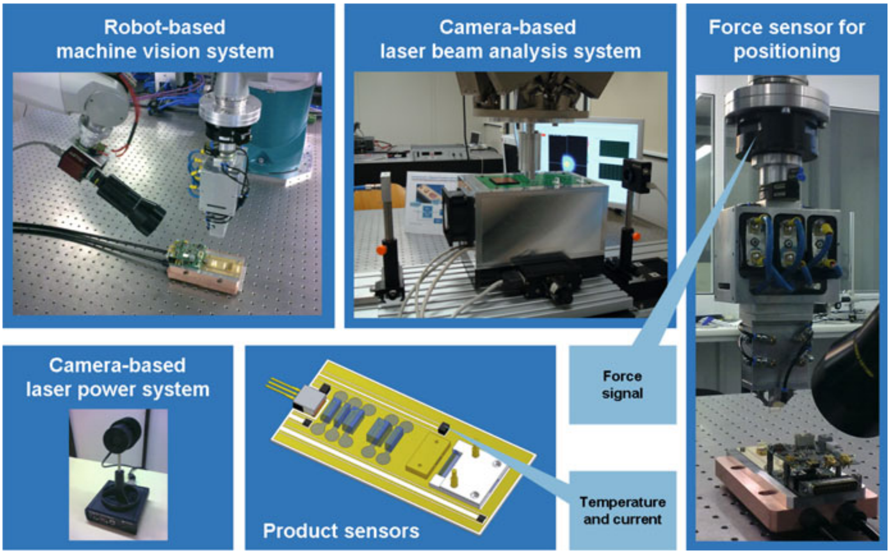


Fig. 6.92 Optical measuring methods for ascertaining the state of assembly

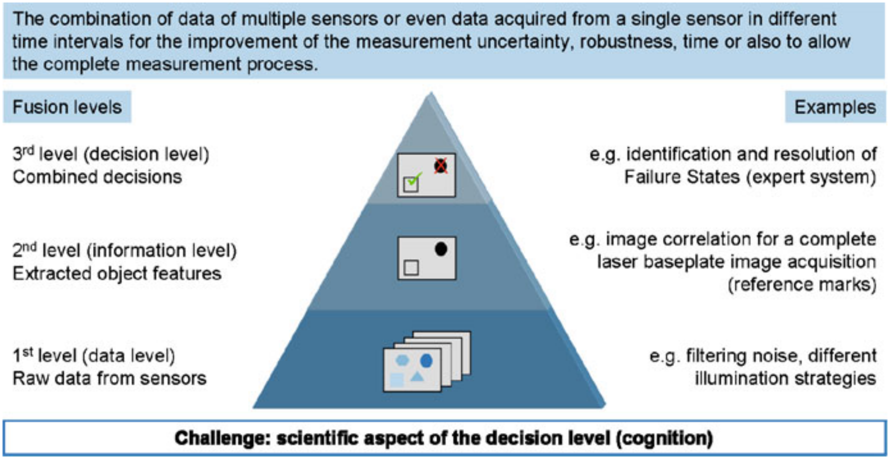


Fig. 6.93 Levels of the sensor data fusion strategy

- a camera-based laser power system; and
- electronic sensors for checking force, temperature and current.

Each sensor system is assigned a corresponding measurement and testing agent. Agents of robot-based measurement systems must constantly cooperate with robot agents to allow flexible measurement. The information gathered by the sensors is analysed by processing agents and stored as general knowledge. The complete state of the entire optical system can only be mapped by the fusion of data from all the measurement systems (Russer and León 2007). Combining the data from multiple sensors, or just the data gathered by a single sensor at different time intervals, improves measurement accuracy, robustness and time, or actually makes the measurement task possible.

The data gathered can be combined at three different levels (Brecher et al. 2009) (Fig. 6.93). At levels 1 and 2 (the data and information levels), images and features are combined to solve typical image processing problems. The capability of fault tolerance (one of the major challenges of laser assembly) is located on the 3rd level (the decision level). Fault-tolerant behaviour is coordinated and executed by an intelligent agent from the task level (failure states agent). The intelligence of this agent is based on an expert system which manages a model (part of the knowledge) of the laser assembly.

The aspects of cognition and decision-making ability of the failure states agent, which are required for the system to be fault tolerant, will now be briefly discussed.

The Intelligent Failure States Agent The use of cognitive aspects in technical systems focuses ensuring the capacity for intelligent decision-making. Modelling and implementing such cognitive abilities is supported by knowledge-based systems. These systems serve as the basis for knowledge representation as well as for inference and learning abilities that are required for dynamic and adaptive systems.

Compared with traditional programming approaches, knowledge-based systems differ by strictly separating the knowledge representation from the knowledge processing (Beierle and Kern-Isberner 2006). Hence, various knowledge-based approaches are based on a common core structure comprising a knowledge base for storing the data and an inference component for knowledge processing. The knowledge base can be created with various types of information which usually falls into two different classes (Beierle and Kern-Isberner 2006):

- *Case-based knowledge:* This is the specific kind of knowledge that relates only to the specific problem case. These are facts that exist because of observations or experimentations;
- *Rule-based knowledge:* This is the core of the knowledge base and may contain domain specific knowledge (theoretical knowledge or even practical knowledge) and general knowledge (problem-solving heuristics, optimisation rules or even knowledge about objects and relationships in the real world).

The failure states agent is supported by a special kind of knowledge-based system: an expert system. The crucial characteristic that differentiates the expert system from other knowledge-based systems is the origin of the knowledge contained in the knowledge base. In an expert system, the knowledge comes from experts who have a high degree of competence in the relevant field, both through appropriate training and through extensive practical experience (Beierle and Kern-Isberner 2006).

The failure states agent has a lot of "if-then" rules within its rule-based knowledge which depict a particular systematisation of the assembly of the laser system. Different information about the assembly is stored this way, e.g. the different components of the system, their setup and tolerances, the variety of metrological systems for checking the conformity of assembly tasks as well as the hardware resources for controlling and changing the system state. The intelligent agent requires further information from planning agents as input so that the inference machine can be initiated. The agent therefore obtains from the planning level a specific sequence or assembly plan for the components as well as the predefined quality level of the laser system. Thus, the agent can guarantee proactive and preventative intervention intermittently with the assembly tasks, in order to always assure the state of the assembly.

The agent then coordinates the measurement and test operations and delegates tasks to the metrological agents so as to support and assure the state of the assembly. This means first identifying the failure states (clear deviation between the desired and the actual assembly situation) by collecting and analysing metrological data. This process runs via the rule-based expert knowledge of the agent. The resulting state diagnosis (logic for deriving the failure state) can be presented by the agent in a comprehensible and transparent manner via an explanation component. Next, the causes of the failure need to be identified and interpreted. In some cases, additional metrological investigations can be used to initiate measures to correct the failure. The process of linking a failure with a solution usually applies the previous experiences of case-based knowledge.

Figure 6.94 shows a typical failure that occurs at the start of laser assembly. In this case, the pump beam is emitted from the diode laser, but instead of passing exactly

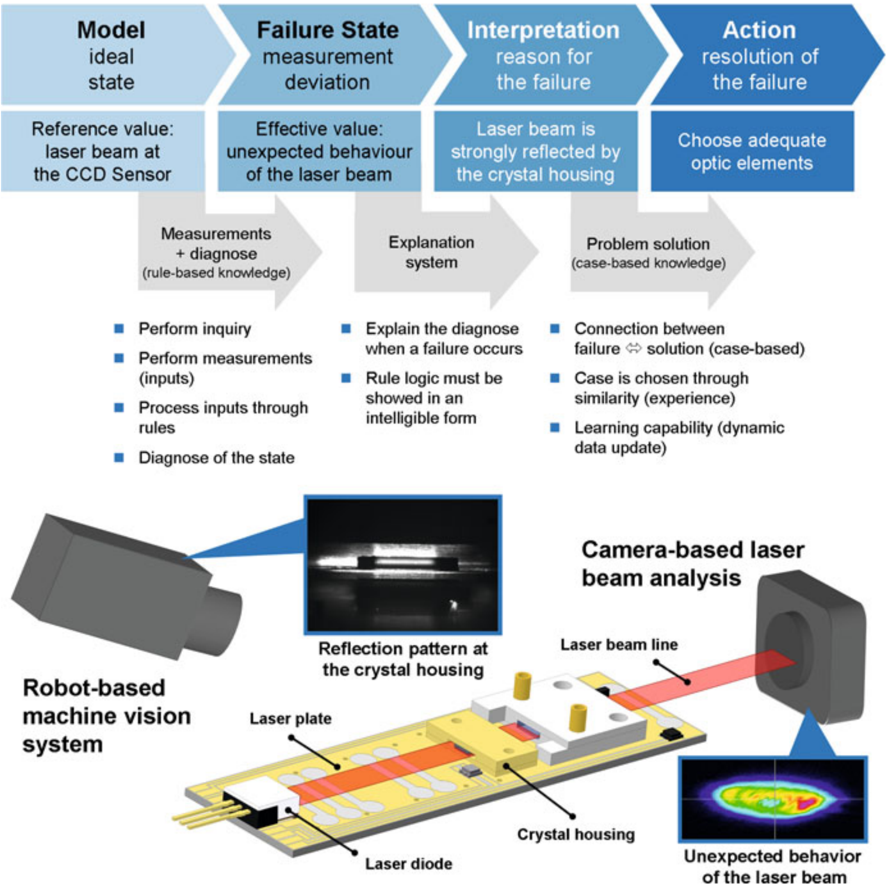


Fig. 6.94 System for identification and resolution of failure states

through the laser crystal, it hits the crystal housing. This failure can be identified by looking at a reflection pattern. At that moment, the diode laser and crystal or crystal housing are already firmly mounted or joined onto the laser mounting plate. A possible solution in this case is the dynamic selection of new optical components which are placed between the diode laser and the crystal, in order to correct the path of the pump beam. Planning agents run the dynamic selection of components as soon as the failure states agent provides the measurement data for the current state.

Model-Based Assembly Control In today's automation solutions, the knowledge about a particular task or system and the "solution competence" as expert know-how of the companies are usually inextricably intermingled. Corresponding subsystem-specific implementations are neither flexible (with respect to changes in the boundary conditions or objectives) nor transparent (with respect to global system behaviour) and thus offer no basis for self-optimising behaviour. Hence, separation into problem-specific modelling and a generic control architecture is presented as the basic scientific hypothesis.

With respect to self-optimising systems, it is still largely unheard of to have an interdisciplinary system and process modelling that covers the entire life cycle of a system ((Adelt et al. 2009) offer a concept geared mainly towards design). Hence, it was a particular scientific challenge to develop a metamodel based on general modelling languages, see for example (Brecher et al. 2007). The second scientific challenge was to develop control architectures capable of creating goal-directed behaviour based on the model information present, whereby current system states are incorporated into further planning during run time. Approaches from the fields of *artificial intelligence*, *operations research* or *cognition* that have been transferred into production engineering to date have promised little if any success (some approaches in production management engineering offer holonic production systems, cf. Schild and Bussmann 2007; Kempf 2010).

Integrative Vision The integrative vision for self-optimisation is based on continuous, model-based development, programming and implementation of self-optimising production cells. In this respect, the aspect of self-optimisation implies the assembly system's capacity for autonomous, in-process adaptation of system behaviour based on varying boundary conditions or objectives (Brecher et al. 2008b). If this takes place in the context of an automated solution, such a system can in principle be used for low-volume, multi-variant production processes as well as for producing large unit numbers (*reduction of scale/scope dilemma*). By using a consistent model description for the entire life cycle of a system, not only can engineering phases in the application be shortened but, above all, they can also be synchronised across all disciplines. In this respect, using the same model information for the runtime system too allows, in principle, global goal attainment (*reduction of plan/value-orientation dilemma*).

The potential offered by an integrative approach is obvious and will be demonstrated using a real assembly process for a complex product. Of particular interest is the possibility that the approach may be transferred to other systems and processes.

A self-optimising system that results from this will in the future be able to make essential, process-related decisions during run time both autonomously and with involvement of human experts and will be able to implement these decisions in different subsystems of the system without jeopardising the global, product-related goal variables (performance, costs). The specific goal of this subproject regarding self-optimisation was to prove the innovation potential of this integrative approach using the example of a realistic automated assembly process.

Results The problem of increasing individualisation and dynamisation of demand on the part of the customer and the resultant need for assembly systems that can be quickly and cost-effectively adapted to varying customer requirements and new products was already extensively investigated during exploratory work at the Chair of Machine Tools regarding operations control for flexible manufacturing systems with explicit focus on the shop floor level (Peters 2001; Possel-Dölken 2006; Fayzullin 2010). Although there are similarities, there are also many differences between the two domains of automated flexible assembly, on the one hand, and flexible manufacturing on the other, as per the flexible manufacturing systems usually found in industry. While, at shop floor level, the dependencies between the manufacturing system components involved are temporally and spatially encapsulated and can therefore be relatively easily organised and formalised in terms of function, at the cell and control levels the system components involved are closely interlinked. The assignment of functions to the individual components is not obvious here and can in fact change dynamically depending on the specific product and the application. A robot can, for example, be converted from a processing unit to a transport or handling unit by changing the gripper. Furthermore, the degree of abstraction of a control task varies: whilst at the shop floor level, clearly delineated work cycle elements, e.g. transport of an object from location A to location B, are referred to as elementary processes, a robot controller in an assembly cell can only run a linear interpolated movement between any space coordinates. In the latter case, there is no reference at all to the semantics of a process. When multiple robots are used in one assembly cell, the working areas of the individual robots may overlap, adding the problems of coordinate transformation and collision detection/avoidance. In terms of the desired self-optimisation, the operational planning task in the shop floor level refers to the allocation of the available degrees of freedom and the determination of an optimal sequence for the individual operations of multiple parallel-dispatched workflows. Planning tasks in cell and control levels are often based on few or even just one workflow instance which, however, has significantly more degrees of freedom. Obviously, the problem of operational control is characterised at cell and control levels by a large number of degrees of freedom, complex and not formally defined semantics as well as significantly higher demands on response time.

The concept of a possible control hierarchy for a self-optimising assembly cell includes a layer that is capable of optimally planning and controlling the assembly procedure during run time (Schmitt et al. 2008). The state of the art for such superordinate control systems is to date purely reactive, usually PLC (programmable logic controller)-based, master control systems. As demonstrated for the shop floor level, it is also assumed that the means for implementing a self-optimising system at cell level are the formal declarative model-based description of the problem domain and the development of algorithms for evaluating these models, together with implementation of the necessary optimisation runs and their execution during run time. The findings resulting from the development of a model-based control system for the automated assembly of the MicroSlab laser include four methodologies:

- Methodology for semantic modelling of flexible processes
- Methodology for model-based engineering of control systems
- Methodology for direct model execution (in contrast to imperative programming or model-to-code transformation)
- Methodology for optimising and planning the processes using models

These methodologies are described below.

Methodology for Semantic Modelling of Flexible Processes Different boundary conditions prevail at each level of the automation pyramid despite the general self-similarity of the control tasks. Quantitative characteristics vary dramatically from level to level, e.g. with regard to the guaranteed response time or the potential diversity of alternatives. Even in discrete manufacturing, processes continually running at the lowest levels, such as movement guidance, also have to be taken into account. In contrast, continuity at the higher levels of the automation pyramid is no longer a binding constraint because it is encapsulated within the limits of the respective manufacturing processes (Fayzullin 2010). The resultant diversity of variants at the individual levels calls for different approaches for implementing self-optimisation, taking into account the different boundary conditions.

The foundation for any process optimisation and planning is the formalisation of the relationships within the production engineering domain by means of adequate models. The need to create appropriate models also results from the demand for inexpensive and flexible development processes. In computer science, a new development paradigm has become established in recent years: model-driven development (Kühne 2005). In model-driven development, relationships within the specialist domain in question are mapped in appropriate models. This allows the problem specification to be separated from the implementation details. Furthermore, model-based development allows new approaches to be used with respect to validating the proposed solution (Fayzullin 2010).

A modelling approach for the semantic modelling of flexible process at shop floor level has already been developed at the Laboratory for Machines Tools and Production Engineering of RWTH Aachen University (Brecher et al. 2008a, c). Based on object-oriented modelling, an abstraction layer was designed in the form of a domain-specific language, which addresses the particularities of process descriptions in a flexible manufacturing system. However, because of the specific problems involved at cell control level, new modelling approaches and modelling input methods needed to be developed in order to engineer the model-based control systems. The significantly higher frequency of events occurring at cell control level and the associated need to map them leads to more detailed models and until now this required considerably increased effort in engineering the models. The purely graphical modelling tools developed thus far allowed the higher levels of the automation pyramid to be modelled with reasonable effort. For the cell control level, however, a new approach was required that would allow the existing meta-meta-model to be reused for faster engineering. The approach taken was to develop a modelling language that allowed graphical modelling to be supplemented with more textual input options.

Methodology for Model-Based Engineering of Control Systems As mentioned above, engineering model-based control systems is, to a considerable extent, comprised of creating the necessary semantic models. Since modelling is very time-intensive, it offers great potential for optimisation with a view to drastically shortening the overall engineering phase for model-based control systems and thus taking into operation of a control system faster. One option is to use template techniques. Templates can be used for modelling certain recurring parts of models. Modelling effort is reduced because actions are no longer elementarily modelled; instead submodels of specific, recurring processes, such as transporting a component from location A to location B, can be inserted directly into the model. To this end, a library of re-usable submodels was created. Another approach is the development of a text-based modelling language that provides a domain-specific scope. The limited scope of a domain-specific language offers the advantage of a shorter learning curve. The additional option of entering all input for the submodels via keyboard significantly reduces the modelling effort.

Methodology for Direct Model Execution The model-based approach is already being used in control technology. However, the state of the art is the transformation of the models created into rigid control logics processed by programmable logic controllers (PLCs). A much more flexible approach is the interpretation and direct execution of models because only this approach enables actions and processes to be planned online (Brecher et al. 2007). The interpretation of models offers some advantages over pure code generation. Change can be implemented faster because adaptations only affect the model and there is no need for renewed compilation and testing. The model is interpreted during run time and can in theory also be amended during run time. The interpreter can also be implemented for other platforms (cloud) and process the same models there.

Methodology for Optimising and Planning the Processes Using Models The duty cycle of the model-based control system includes perceiving information about the state of the assembly system, storing this information in the internal data model, planning actions online using the internal model and executing the plans created by triggering real actions. While these tasks are being processed, the model-based control system can have read access to the information stored in the semantic model regarding the types and instances of the system components and processes. If a state change is registered by the model-based control system or if an internal schedule changes, the control system modifies the content of the instances. The chronology of the tasks taking place in the decision cycle of the model-based control system is based on the preparatory work undertaken at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University (Brecher et al. 2008b). However, unlike event processing at shop floor level, where querying the state of the system components is time-discrete, the time-critical nature of the cell control level requires a different approach.

Validation A flexible assembly cell was developed for the purpose of validating the methodology developed (Fig. 6.95).

Since the assembly cell consists of three cooperating robots that can be equipped with various grippers, tools or measuring instruments, a great deal of flexibility is offered in terms of performing assembly operations. The challenge was to create a close link between the system components at cell and control levels. As mentioned previously, each system component can be assigned different functions which may

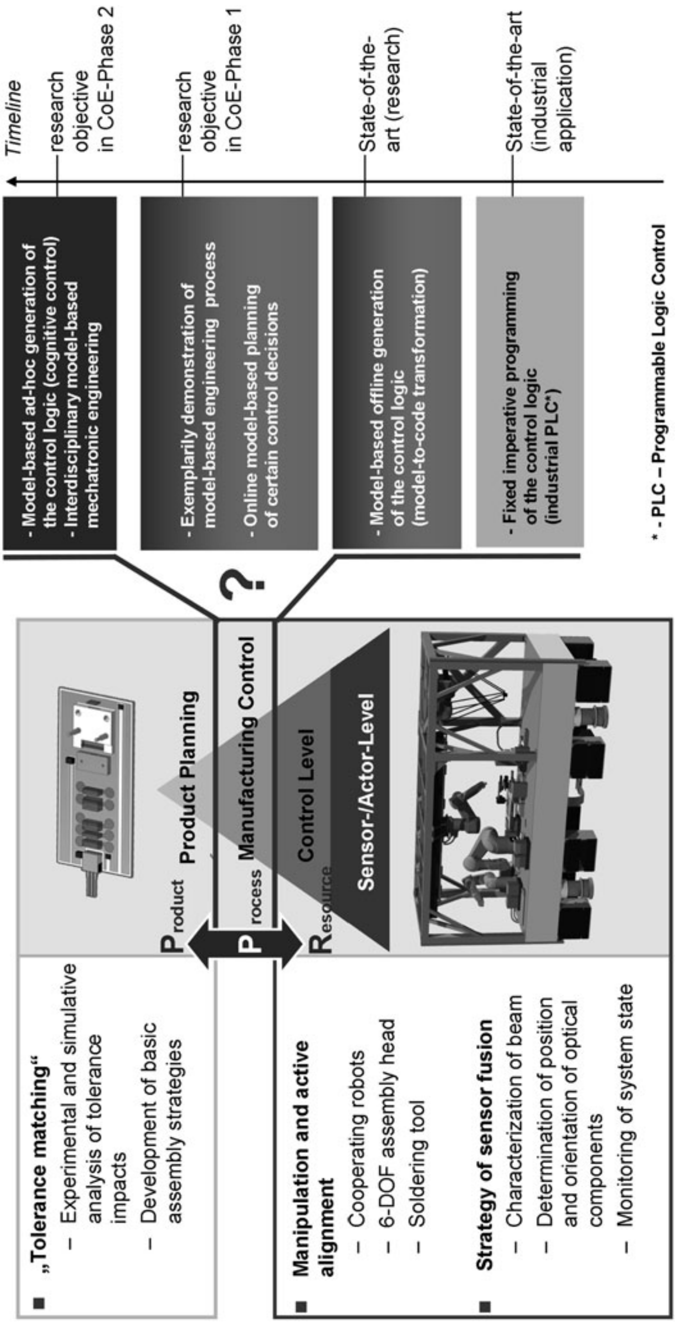


Fig. 6.95 Timeline of development of the cognitive control systems

be changed during the assembly process. Depending on the application or the product to be assembled, a robot can be reconfigured from a transport robot, for example, to a soldering robot merely by changing the gripper. This means the semantics of each element may change within the elementary processes compared with the superordinate process.

6.5.4.3 Integrative Product and Process Design

In assembly, products with defined functions are manufactured from a range of components or sub-assemblies. As part of this process, complex products have to be not only joined, but also put into operation, checked and aligned. In assembly, these steps are closely interlinked. Since assembly is usually at the end of the company's value-adding chain, it is the key to ensuring that product features are implemented at the desired quality level. It represents the last opportunity to influence features and their characteristics.

When planning an assembly system, the product to be assembled is usually taken as the starting point for deliberations. Taking this into account, first a rough process chain is created. Then the production equipment to be used in each process is selected based on that chain.

Given the objectives, some of which are fixed, with respect to quality and output quantity and in view of the goal of cost-optimal production, however, a strict unidirectional procedure would not make sense. There are mutual dependencies between product, process and production equipment that must be taken into consideration. Hence, the assemblability of a product should be taken into account during its design and development.

By integrating product and process design at an early stage, cost-intensive and time-consuming iterations can be reduced through integrative consideration of product, process and production equipment. Restrictions arising from the assembly process or the production equipment can therefore be taken into account during product design. This makes it possible to avoid costly changes to the product after ramp-up. Likewise, this avoids potential subsequent adjustments of the specified assembly process and the production equipment acquired.

With flexible assembly systems in particular, the process must be designed to be robust so that the system can be returned to operation quickly following a reconfiguration. This makes it possible to automate small series economically too.

A key factor for success in this respect is to collaborate with the company divisions involved, especially in the early planning stages. Traditional organisational structures with rigid, self-contained departments often have an entirely departmentalised way of thinking which causes significant friction losses in intersectoral cooperation. In this case, the departments usually only consider the part of the value-adding process that they are directly entrusted with. With the focus on departmental cost objectives, the individual departments tend to keep their areas of responsibility as small as possible and thus always choose the problem solution variant that is the least expensive. The aim of efficient business organisation, however, is to reach the overall optimum,

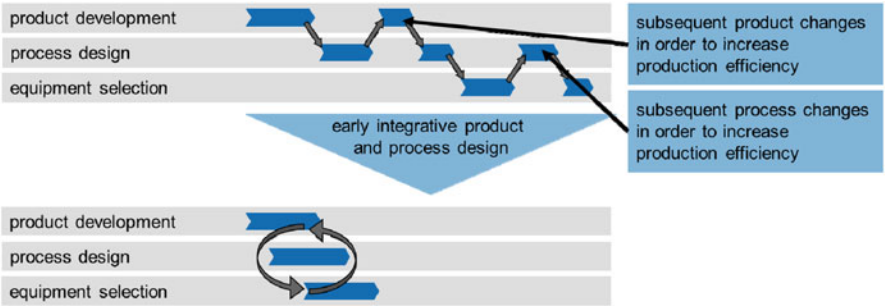


Fig. 6.96 Reduction of time-to-market and costs by early consideration of product, process and production equipment

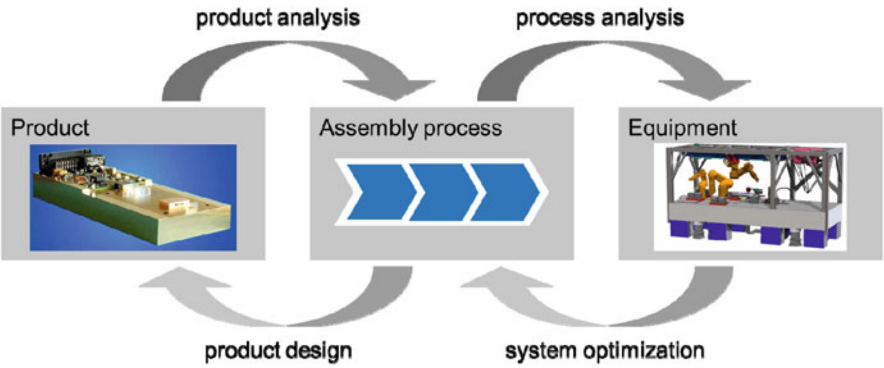


Fig. 6.97 Interactions between product, assembly process and production equipment

which in general cannot be achieved with localised optimisation measures. In fact, a procedure that is particularly inexpensive from one department's point of view can often lead to increased costs in the downstream business divisions. Accordingly, the best approach in terms of optimising production company-wide is to ensure the individual departments are closely interlinked or to even completely eliminate rigid departmental boundaries (Fig. 6.96). This facilitates, amongst other things, the integrativity of product design and production planning.

Due to the restrictions caused by the product, process and production equipment and due to economic influences, planning is becoming very complex, especially where flexible and adaptable assembly systems are concerned. In this respect, methods have been developed to assist the user in the design of flexible assembly systems. These methods can be assigned to the various stages of assembly system planning (Fig. 6.97).

Methodology for Product Analysis In assembly, to provide the required product features at the desired quality level, functional tolerances are defined during product design. Tolerance zones have to be defined because in no technical process can the target value of a property be achieved exactly and consistently (Trumpold et al. 1997). The objective of establishing tolerances is to find a compromise between the functionality of a component and the manufacturing costs. In this respect, the focus is often on the production of individual components and the features structure of the end product.

Often, products are designed which, due to their tolerances, can only be produced using the most complex assembly processes and the associated cost-intensive production equipment. Furthermore it is commonplace for assembly processes to be planned that only meet the required tolerances after the most laborious adaptation. To prevent this, it is essential that the tolerance problems are taken into account at an early stage in the product creation process. However, because of the large numbers involved, explicit consideration and control of all specified tolerances for complex products, e.g. the miniaturised solid-state laser, is unreasonable from an economic point of view and is often impossible from a technical point of view (Thornton 2004).

To facilitate early and efficient assembly planning despite the above challenge, the tolerances to be taken into account during planning are derived from the customer's requirements, which determine the product requirements. Whilst deriving this information, a fundamental understanding of the product is generated, which is essential for developing an efficient assembly process.

The tolerances to be taken into account are identified using the key characteristics method introduced in the 1980s in the USA and continuously further developed since then within the framework of scientific publications (Whitney 2004; Merget 2004; Müller et al. 2009a).

The term key characteristic (KC) refers to a quantifiable feature of a product, an assembly, an individual part or a process whose expected variation from the target value has an unacceptable impact on the costs, performance, quality (as perceived by the customer) or safety of a product (Merget 2004). Hence, to avoid such impact, KC variations must be limited by tolerances and these must be explicitly taken into account during assembly planning.

An essential element of the KC method is the key characteristic flowdown which shows the relationships between customer requirements and the KCs of individual parts (Merget 2004). In creating the KC flowdown according to the top-down approach, the customer requirements can be broken down into the KCs of the individual parts (Müller et al. 2009a). The tolerances to be taken into account can then be derived from the KCs of these parts (Merget 2004).

For a marking laser (see Sect. 7.3), beam quality, for example, is a customer requirement for marking on surfaces. The beam is influenced by the arrangement of the lenses and the crystal. For high beam quality, the beam must hit the centre point of the crystal exactly. To do this, the lenses must be aligned with the crystal (Fig. 6.98). Based on this, it is possible to define the functional tolerances that will ensure the beam offers the required quality.

Based on these functional tolerances, it is possible to derive the assembly sequence as well as the possible joining and handling technologies.

Methodology for Assembly-oriented Product Design Assembly processes and equipment are influenced significantly by the product design. Assembly-oriented

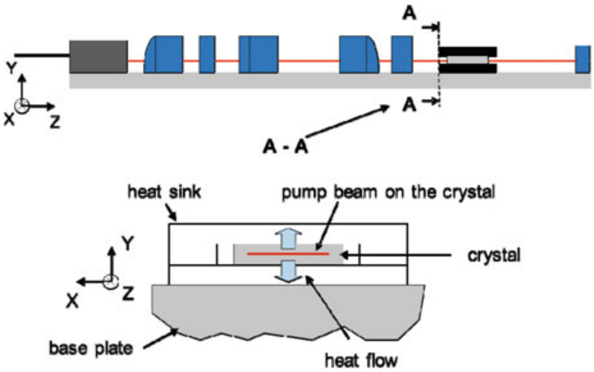


Fig. 6.98 Relationship between beam quality and arrangement of lenses

and handling-oriented product design is therefore always at the centre of the development process. The relevant literature contains a large number of guidelines to support the designer in this work. These guidelines provide a very good basis and offer valuable advice regarding assembly-oriented design. Due to the diversity and scope of the guidelines, however, it is seldom possible to fully incorporate all the recommendations.

When designing products, the restrictions caused by the process always have to be taken into account. Accordingly, it is technically impossible to, for example, assemble components with arbitrary precision. Hence, it makes sense to tailor product design to the assembly process.

Figure 6.99 shows an example of a conventional laser. Each lens is fixed in a separate carrier. The carriers have various settings as required for aligning and setting the beam. Especially for a complex product like the laser, whereby the interactions between the lenses, in particular, affects beam quality, alignment is very laborious and can only be performed by highly skilled and experienced personnel.

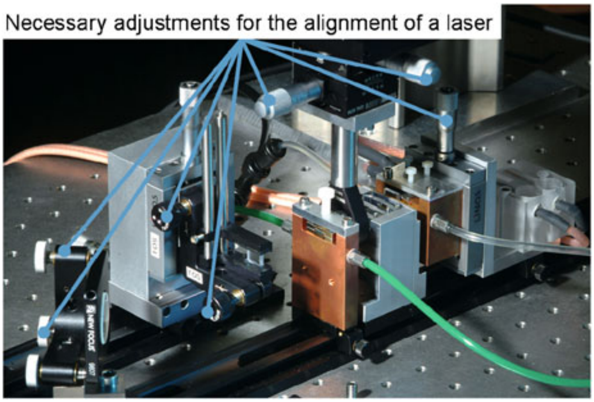


Fig. 6.99 Construction of a conventional laser

With assembly-oriented product design, the product can be modified so that it can be joined and aligned in an automated precision assembly process. In this case, the laser was designed so that the lenses could be joined from above using a robot. The new rectangular design of the lenses helps reduce the degrees of freedom of component positioning and orientation compared with the previous circular design. The lenses are joined to the motherboard by means of a soldering process. This allows the parts to be aligned during assembly whilst in the solder bath. When the solder cools, the lenses remain in the set position and orientation.

Furthermore, the laser was designed according to the so-called nest assembly, i.e. the individual parts or assemblies are arranged to a large extent side by side, as with an electronic circuit board for example (Lotter 2006). The order of installation is freely selectable. This results in a large number of acceptable process variants and optimisation approaches through variation of the assembly sequence.

Methodology for Process Analysis Process analysis considers the individual boundary conditions and objectives in the process. For the assembly process to be implemented, the functional tolerances must be maintained during assembly. The influences of the components and the production equipment on assembly tolerance must be identified in order to allow economic design of the production equipment. The aim is to compile the system so as to meet specific needs because excessive demands on the production equipment are usually associated with higher costs.

The tolerance analysis can be used in the production equipment design process. This analysis makes it possible to determine the assembly tolerance as a function of the production equipment. In this respect, possible sequencing options for the individual assembly steps are identified. In each of these steps two parts are assembled to form one assembly. These alternative sequences are then analysed to determine which of them make it possible to maintain the specified tolerances. The tolerance chain is visualised to facilitate identification and analysis of the sequencing options.

This step of the standardised approach includes, amongst other things, working out which assembly equipment and joining procedures can be used to assemble the individual parts. In addition, any component-induced or process-induced deviations arising during assembly are also revealed. To what extent these deviations need to be changed in order to efficiently meet the specified tolerances is also investigated. At the same time, implementation options for changing the deviations also have to be determined.

To help identify these options, the deviations that arise during each assembly step alternative are visualised in a graphical model (Fig. 6.100).

The nodes of the graph symbolise the characteristics of the assembly system (e.g. workpiece carrier stop). The edges of the graph describe the geometric deviations between the characteristics (e.g. the deviation in distance between the workpiece carrier stop and the edge of the laser mounting plate). This method of illustrating deviations is based on visualisation of the dimensional chains, the tolerance chains, the liaison diagram and the datum flow chain (Müller et al. 2009a). Since the graph visualises the greatest possible deviations between the actual and target variables and

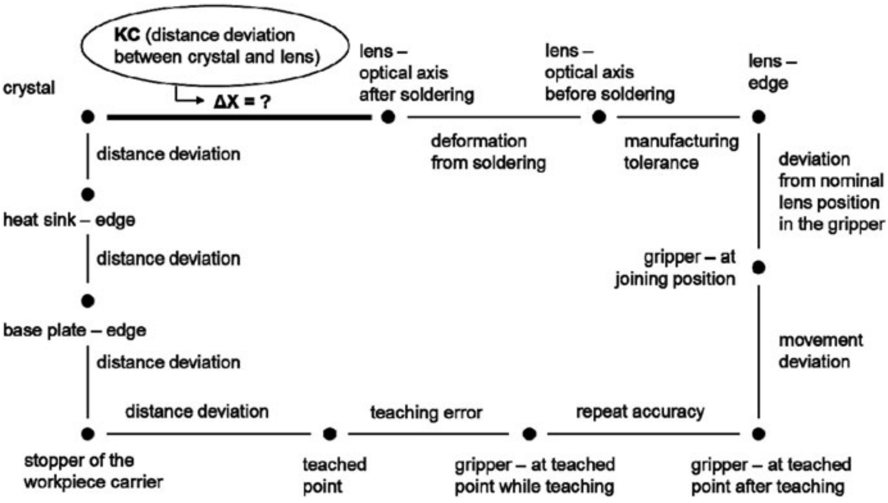


Fig. 6.100 Tolerance chain of laser assembly step alternatives

these determine the tolerances that can be satisfied, the graph is also referred to as a tolerance chain.

The deviation in distance between the crystal and the lens in the tolerance chain of the MicroSlab laser (Fig. 6.98) must be explicitly taken into account during assembly planning to ensure the laser can perform (i.e. this concerns a KC). The remaining elements of the tolerance chain symbolise the deviations that arise during the assembly step alternative in question and cause the deviation in distance between the crystal and the lens.

Methodology for System Optimisation The methods presented in Fig. 6.101 can be used in order create more efficient or shorter tolerance chains.

In the first method, the length of the tolerance chain is reduced by shortening one or more elements (Fig. 6.101, no. 1). So, in the tolerance chain from Fig. 6.100, for

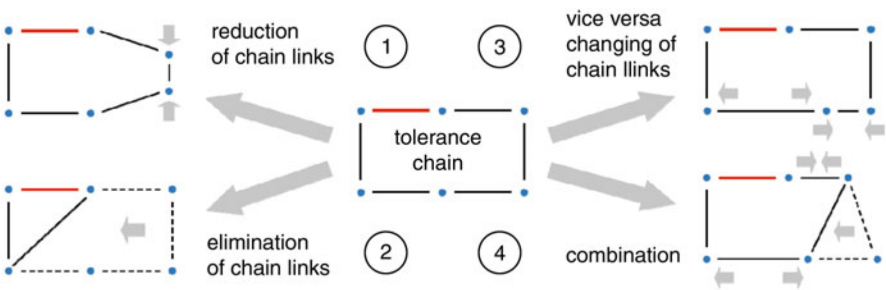


Fig. 6.101 Methods for changing the tolerance chain

example, one element can be shortened by using a robot with greater positioning accuracy. However, this is usually associated with higher costs.

Another method for shortening the tolerance chain is to reduce the number of chain links (Fig. 6.101, no. 2). Hence in the tolerance chain visualised in Fig. 6.100, the number of elements can be reduced by measuring the position of the lens with respect to the crystal because in doing so, the functionally relevant characteristic itself would be measured (distance between the lens and the crystal). Consequently, the repeat accuracy of the robot, for example, would no longer have any impact on the deviation in distance between the crystal and the lens.

To increase the efficiency of implementation of the tolerances while complying with the overall criterion of the tolerance chain (i.e. the specified tolerance), reciprocally changing individual elements can also be effective (Fig. 6.101, no. 3). This is the case, when, for example, shortening one element is associated with relatively low costs while lengthening another is associated with relatively high savings. Hence, instead of an articulated arm robot, a simple handling device could be used in conjunction with a fine-positioner. That way, the handling device puts the fine-positioner within the range required to join a lens. As already shown in Fig. 6.101, no. 2, the inaccuracies of the handling device are compensated by additional measuring technology.

Simultaneous application of the three methods specified shortens and/or increases the efficiency of the tolerance chain (Fig. 6.101, no. 4). These explanations clearly show that by examining the tolerance chain, it is possible to identify ways of shortening it and to uncover potential for making tolerance implementation more efficient. This potential is uncovered mainly through modelling and therefore focussing on the deviations and tolerances. The three methods of system optimisation are presented in detail below.

System Optimisation by Shortening Individual Elements The flexibility required for self-optimising assembly is usually achieved through robot systems. In this connection, storing a model of the robot and its environment in the control system allows robot programs to be automatically adapted when parts or processes are changed. The precision of execution of the new instructions is primarily determined by the model's consistency with the actual assembly system. However, significant deviations due to production tolerances mainly occur when nominal robot geometry is used to create the model (Elatta et al. 2004).

Within the context of self-optimisation, the assembly tolerances specified must be met without additional and complex automatable corrective measures, which is why highly accurate models are required. In this respect, through integrative system identification, it is possible to determine the relevant model parameters for the assembly system extremely accurately and thus significantly reduce the gap between simulation and reality. The following diagram (Fig. 6.102) illustrates the basic system identification process (Müller et al. 2010).

First, a nominal model is created for the assembly system. For the purpose of analysis and identification, the entire system is broken down into subsystems. The parameters of each subsystem are identified using established, state-of-the-art methods.

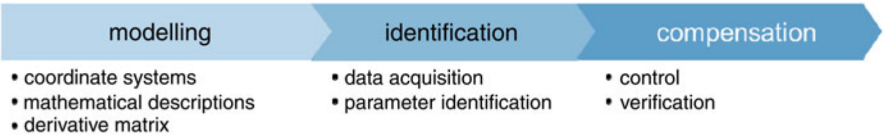


Fig. 6.102 System identification procedure

To this end, internal and external measuring systems are used to collect the data required to calculate the characteristic parameters needed to describe the model. Next, the control system and simulation environment are adapted accordingly. Finally, further measurements are taken to verify the model identified.

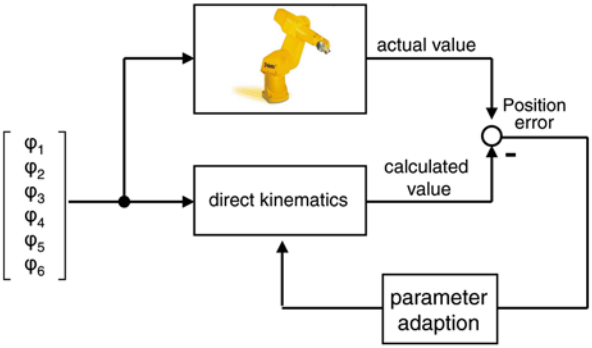
The first step of the procedure is to model the assembly cell using a parametrisable model. Based on the CAD data, the models are represented in a graphical simulation environment. For a complete model, not only the handling devices are included but also the product to be assembled, the necessary production equipment and the collision bodies (fencing, equipment housings, etc.). Next, all objects in the cell are provided with coordinate systems.

In offline programming, system behaviour must be precisely mapped in order to minimise the adaptation effort required in the actual cell. Complete and error-free mapping of the assembly system in a model, however, is impossible because some errors are of a stochastic nature and therefore cannot be explicitly modelled. On the other hand, not every single influence that can possibly be mapped has to be included in the model because identifying the model parameters based on the real system can be extremely laborious. In many cases involving robots, only the exact kinematic parameters are determined while the dynamic effects are disregarded because this simple calibration leads to a substantial and more than adequate improvement in accuracy.

The model to be identified must not contain uninfluential or linearly dependent parameters. In describing the robot using Denavit-Hartenberg parameters (DH parameters), some parameters may be interdependent. With a vertical articulated arm robot, for example, the movement parameters of the second and third axis are interdependent. Hence, for identification, one of the two movement parameters is fixed and not available for optimisation. Also, it is necessary to ensure that only those parameters that affect performance are made available for optimisation. If it is just the position that is determined with the measuring system, then parameters that only affect TCP (tool centre point) orientation cannot be identified because they cannot be monitored using the measuring system.

To allow numerical optimisation, it is important to use a reduced and non-redundant model. If this criterion of minimalism is not met, the system of equations cannot be solved explicitly. To be able to minimise the subsequent positional error between the model and the actual kinematics, therefore, a regular system of equations must be created for parameter identification. Then the parameters can be identified in the next step.

Fig. 6.103 Schema for parameter identification for industrial robots



A numerical optimisation function is used to identify the parameters (see Fig. 6.103). To this end, the robot's TCP is measured for a particular set of joint angle combinations using an external measuring system. The TCP poses corresponding to these poses are calculated using a parameterised and non-redundant model. Based on the measured and calculated poses, an error function can be set up for each joint angle combination.

Using the least error squares method, the available model parameters are purposely changed until the positional errors for all joint angle combinations reach a particular termination criterion.

In addition to determining the exact model parameters, the kinematic parameters of the production equipment must also be identified. To this end, using an external measuring system, numerous significant edges, surfaces or points of the body are sampled in a reference coordinate system. By determining these characteristics, the body's own coordinate systems can be identified.

The final steps of the process are compensation and verification. In order for the identified model to be used, both the cell control system and the simulation environment have to be adapted. Problems can arise, especially with respect to adapting the robot controls, because not every control manufacturer provides the option for the user to adapt the relevant parameters. This is due to the fact that with the reverse transformations implemented, simplifications were made. Should the model deviate from these values, the stored reverse transformation can no longer be used and another way to adapt the control system must be found. This can be achieved by means of an upstream algorithm for specifying the Cartesian coordinates. In this respect, the target coordinates from path planning are converted into joint coordinates using the calibrated reverse transformation. With the nominal model used for the robot control system, the joint coordinates can be converted into Cartesian target coordinates. The manipulated Cartesian coordinates are then converted into the target coordinates by the robot control system and the actual kinematics (see Fig. 6.104) (Wiest 2001).

Once the models have been adapted in the simulation and to the control system, the result has to be validated. This step is essential, especially for numerically generated analogous models since these models have only been optimised for a specific working area.

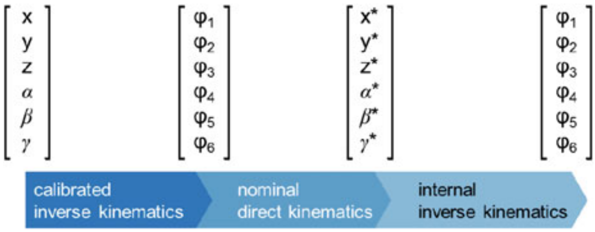


Fig. 6.104 Algorithm for specifying Cartesian coordinates

This approach to system identification makes it possible to improve the accuracy of models for robots and assembly systems and thus significantly shorten the corresponding tolerance chain.

System Optimisation by Eliminating Individual Elements The accuracy achieved by determining the exact kinematic parameters is sufficient for many applications, e.g. pick-and-place tasks. However, the accuracy usually remains insufficient for precision applications. Hence in many cases, additional measuring technology and sensors are used in order to improve the accuracy.

Figure 6.105 presents an example of camera-supported lens positioning. One robot holds a camera that is pointing towards the base plate. Another robot independently brings the lens to be mounted into the camera's image area. By measuring the base plate and the component, the camera incrementally guides the robot into the right joining position. This serves to eliminate many effects of the tolerance chain, e.g. the positioning accuracy of the robot.

6.5.5 Outlook

The theoretical foundations of self-optimising assembly systems presented in this paper are investigated and demonstrated with respect to various aspects based on the fully automated assembly of a miniaturised solid-state laser.

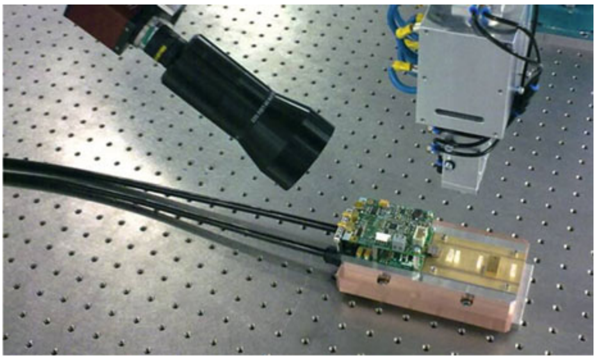


Fig. 6.105 Inline measuring technology for improving positional accuracy

Towards the vision of autonomous, self-optimising assembly systems, it is important to continue building on these results especially in the field of automated planning at cell and control levels. Due to the specific characteristics of multi-robot systems in particular, transferring the shop floor engineering solutions of flexible manufacturing systems to flexible assembly cells is only possible to a limited extent and requires new approaches. Through greater integration of system components at the lower control levels and through the configurability of components (mobile units, modular tools and sensors), functions and capabilities cannot often be assigned clearly and therefore considerably increase the degrees of freedom of planning activities.

When several robots are used in one assembly cell, the working areas of the individual robots may overlap, which is why additional aspects of coordinate transformation as well as collision detection and avoidance are added. Also, besides a large number of degrees of freedom plus complex and not formally defined semantics, the problem of assembly control is characterised at cell and control levels by significantly higher demands on response time, too. Only with appropriate real-time control solutions for modular, distributed systems along with new description models and planners for configurable and cooperative capabilities can the next step towards continuous self-optimisation in automated production be taken.

6.6 Self-optimising Assembly Systems Based on Cognitive Technologies

Marcel Ph. Mayer, Barbara Odenthal, Daniel Ewert, Tobias Kempf, Daniel Behnen, Christian Büscher, Sinem Kuz, Simon Müller, Eckart Hauck, Bernhard Kausch, Daniel Schilberg, Werner Herfs, Christopher M. Schlick, Sabina Jeschke and Christian Brecher

6.6.1 Abstract

Like business and production processes, entire production systems are often based on hypotheses that only provide a partial view of the value-adding chain or result in specific technological interactions. We rarely fully understand the way that processes, materials, production resources and humans interact or how this affects the product itself. It is impossible to predict how changes will impact on value creation as a whole. Because it is so difficult to gain an integrated view, optimisation processes often only focus on individual elements of a system. Existing ways of optimising the behaviour of a single element can become the focus of too much attention and can tie up resources even though, in certain situations, doing so can have a negative impact on behaviour in other areas of the system.

The solution to this conflict lies in designing a system that can adapt its goals according to the situation. While most system optimisation processes happen

externally—i.e. a human controls them—having the system itself perform them can be an interesting option in many cases. However, developments in automation show that, even in comparatively simple situations, we still have work to do to achieve this goal. This means that humans often still play an important role in system optimisation. Implementing self-optimising capabilities offers great potential for resolving issues associated with planning efficiency.

Using cognitive functions as a basis, it is possible to take parts of the original range of tasks—like detailed algorithmic planning of an assembly process, and rule-based re-planning and adaptation planning—and transfer these from humans to the machines themselves. The first step in achieving this involves designing and developing a software architecture that is capable of fulfilling the specific requirements of a cognitive system, such as independently planning actions and executing them. This architecture is based on a three-layer model that is expanded to include modules for knowledge representation and human-machine interaction. The deliberative part of this architecture—the planning layer—is where the actual planning and decision-making happens, through targeted linking of production rules. A hybrid approach is used to achieve a flow-oriented, error-free assembly of randomly fed-in components. The approach combines traditional planning processes (an "offline planner") with reactive cognitive control (an "online planner"). Prior to execution, the machine's offline planner calculates the complex geometric analysis steps for assembly planning. This involves an assembly-by-disassembly strategy, which identifies all the assembly steps necessary to bring together individual components into a finished product. To do so, all possible pairs of subsets of components are disassembled and checked to establish whether this can be done collision-free. This process is repeated until the product has been completely disassembled. Because an assembly action can reverse every disassembly action, read backwards the disassembly steps produce all possible assembly steps. The results of this preliminary analysis are recorded as a state graph. The edges in the graph are weighted with cost values for the different work steps and also show information on any additional components needed. Weightings for the different interim states are also assigned to the nodes in the graph.

The online planner can evaluate this information during assembly to derive a suitable assembly sequence for the current production situation and a given target system. This involves updating the state graph on the basis of the available components and the current state of the partially assembled product. Next, the A* algorithm is used to calculate the best route from the current state to the target state (the end product). Under the prevailing boundary conditions, this path represents the optimal assembly sequence. The next work step involves investigating the assembly sequence to identify possible parallelisable assembly actions. The parallelisable actions are then combined in such a way that the assembly plan is made up of a sequence of sets of assembly actions that can be executed at the same time. This plan is then passed on to the decision-making component—a cognitive control unit (CCU) based on SOAR cognitive architecture—where it serves as a basis for decisions. The CCU checks the current situation in cycles and initiates appropriate actions. Under normal circumstances, from the set of parallelisable actions to be executed next, the CCU randomly selects an action and implements it. In unforeseen situations, either re-planning is initiated or the human operator is requested to solve the problem. This process continues in cycles until the desired product has been fully assembled.

A robot-supported assembly cell is presented to validate the system. In addition to covering all key aspects of an industrial application (relevance), the cell can easily illustrate the function and flexibility of a CCU (transparency). The assembly cell has a conveyor-belt system, which is used to feed in components. There are two jointed-arm robots, of which only one is initially connected to the CCU. This robot can be moved along a linear axis and, in addition to having a flexible multi-finger gripper, is equipped with a colour camera to identify components. The second robot separately feeds in components on a conveyer belt. In the centre of the assembly cell there is a surface that can be used as a workspace and as interim storage for parts that cannot be used immediately. The human operator's workspace is currently located directly next to the surface, separated by an optical safety barrier. The workspace has a human-machine interface that can be either stationary or mobile. By displaying process information in an ergonomic way, the interface provides details on the system status and can, if necessary, help human operators to interactively identify and resolve process errors.

To enable humans to safely and effectively manage and monitor the process, machine-made decisions (e.g. on establishing the assembly sequence) must be presented quickly and in a way that is easy to understand and will ensure reliable implementation. Furthermore, the assembly sequence must be planned in a way that conforms to human expectations. With this in mind, a concept was developed that enables system behaviour to adapt to operators' expectations by using a human-centred process logic based on the MTM-1 system. This concept was validated using the assembly cell. Two series of experiments under laboratory conditions sought to establish whether humans followed certain easy-to-generalise strategies when carrying out an assembly task. The experiments identified and validated three assembly strategies as rules: (1) People begin assembling from edge positions; (2) People prefer to assemble adjacent to existing objects; (3) People prefer to assemble in layers. A specially developed simulation environment shows to what extent taking the identified assembly strategies into account as production rules within the knowledge base of the CCU makes it easier to predict the actions of the controlled assembly robot, and to what degree the rules can be generalised.

The presentation of results concludes with a lab study designed to investigate visually presenting information to humans. The starting point was the following scenario: a robot plans the workflow for a production task previously defined by a human operator and then carries it out. However, an error occurs during processing. Because the cognitively automated system is unable to identify the error itself, the human operator must be put in a position to quickly and safely identify and resolve the error. The study compared different ways of displaying the information, in order to find the fastest and most reliable one for interactive error identification. A subsequent comparison looked at the differences between a TFT screen in a workroom and a head-mounted stereoscopic display. The results showed that a display fitted over the field of view improved error detection rates.

The research results made it possible, for the first time ever, to achieve an integrated demonstration of, and carry out scientific investigation into, the design, development and application of cognitive mechanisms in automation using a robot-supported assembly cell to control them. The work also identified the need for further research into whether the results can be transferred to products widely used in industry, and to the ergonomic design of human-machine cooperation.

The cognitive automation of production systems offers a technology that, with the same or even less planning effort, can efficiently and robustly automate product families with large numbers of variants—even in cases where only small numbers of each variant are produced. This effectively makes customer-oriented mass production possible. Cognitive mechanisms in automation offer high-wage countries in particular the chance to achieve considerable competitive advantages, and thereby to directly contribute to securing and expanding their own production locations.

6.6.2 State of the Art

Autonomous production cells can be considered the predecessors of cognitively automated production systems. They are mainly characterised by physical process models integrated directly into the machine control system and by machine operation and monitoring based on these models. Safe, effective human-machine interaction means that even complex handling processes can be carried out error-free over an extended period of time, and that the work system can be operated at the optimum level in terms of performance and operational requirements (Schlick 1999; Pfeifer and Schmitt 2006). However, automated production cells only have limited self-optimising planning functions; these functions are crucial to the concept of cognitive automation. The work of Onken and Schulte (2010) built on these functions and helped to shape the concept of cognitive automation and introduce it to the scientific community. But the original concept focuses heavily on using the technology in unmanned vehicles and is therefore only partially applicable to production systems. Although the corresponding concept of the "cognitive factory" (Zaeh et al. 2009) shows that cognitive mechanisms can be successfully integrated into manufacturing systems, the overriding subject of automation using cognitive models that incorporate the "human factor" for machine operation and monitoring has remained largely unexplored.

To present the current status of research into cognitive automation in production with a view to application in assembly contexts, we must look at a number of different aspects. In view of the research results presented in Sect. 6.6, what follows will focus on four central aspects: (1) Software architectures for cognitive systems; (2) Planning assembly sequences with formal methods; (3) Industrial automation; (4) Task allocation between humans and machines.

6.6.2.1 Software Architectures for Cognitive Systems

With regard to designing and developing automated robotic systems, numerous architectures have been proposed as basic structures for simulating cognitive functions

(Karim et al. 2006; Gat 1998). These software architectures combine a deliberative part for the actual planning process (planning level) with a reactive part for direct control (action level). A widely used approach here is the three-layer model that comprises cognitive, associative and reactive control layers (Russell and Norvig 2003; Paetzold 2006). The lowest layer (reactive) contains the components that control information processing, and is designed to influence system behaviour in such a way as to ensure that the required reference variables are achieved quickly and accurately. The associative layer monitors and controls the system. The majority of rule-based auxiliary functions for automation—like process control, monitoring processes and emergency processes, and adaptation routines for improving system behaviour—are all embedded here. In this top layer, the system can apply "reflexive" methods (e.g. planning and learning processes, model-oriented optimisation processes and knowledge-based systems) to use knowledge about itself and its environment to improve its own behaviour. The focus here is on the system's cognitive ability to carry out self-optimisation.

The software architecture used in Collaborative Research Centre (CRC) 614—"Self-optimizing concepts and structures in mechanical engineering"—picks up on this model (Gausemeier et al. 2009a, b). The Cognitive Controller from the Technische Universität München is also based on a multilayer model. The signals from the production system in question are prepared and processed by a standard controller and by a cognitive safety controller. Additional general studies of cognitive systems can be found in Onken and Schulte (2010) and, regarding the production environment in particular, in Ding et al. (2008). The latter focuses on implementing cognitive capabilities in security systems for plant control. As such, the study pays particular attention to safety in human-machine interaction and to safety in the workplace.

Complementary concepts and methods can be found for automobiles and air and space travel (Kammel et al. 2008; Putzer 2004). Particularly noteworthy here are the various architectures used by the winning teams in the DARPA Grand Challenge, a competition for driverless ground vehicles. These architectures provide regular insight into the latest research (Thrun et al. 2006; Kammel et al. 2008). Many of the architectures are modelled on a "relaxed layered system". In this system, every layer can use the services of the layers below. This makes the system more flexible and efficient by ensuring that every layer is fully supplied with information (Urmson et al. 2008).

This shows that modified architecture models are used to configure and develop cognitive systems. We can see that the most-used approach involves combining the multilayer model with other models. The final architecture must therefore be adapted to the specific application to ensure the highest level of system performance.

6.6.2.2 Planning Assembly Sequences using Formal Methods

"Planners" play a major role in cognitive systems. There are numerous formal approaches to solving planning tasks in different fields of application. Hoffmann (2001) developed the Fast-Forward Planner, which is capable of deriving actions for given problems in deterministic operational areas. By contrast, other planners can handle uncertainty (Hoffmann and Brafmann 2005; Castellini et al. 2001). All the planners mentioned are based on a symbolic knowledge representation. In the case of assembly planning, which requires the geometric relationship between conditions and their transitions to be adequately represented, this kind of knowledge representation becomes extremely complex, even for simple tasks. As a result, generic planners fail when it comes to calculating assembly sequences with even a short-term planning horizon.

Other planners have been specially designed for assembly planning. They work directly with geometric information to derive action sequences. A popular approach is the Archimedes system (Kaufman et al. 1996), which uses and/or graphs for formal representation, and the assembly-by-disassembly strategy to dismantle the end product into its individual components. Thomas (2008) picks up on this concept and develops it in such a way that the system requires no additional help from the user to deduce all possible disassembly steps for the end product—as was the case with the Archimedes system. Thomas only uses geometric information about the end product and the components it contains. This approach, however, is not capable of dealing with "uncertain" planning data. We can find another approach in the assembly planner developed by Zaeh and Wiesbeck (2008), which calculates action sequences almost independently. However, it is not designed to control a technical system, but as a support system for manual assembly. This means that the decision to carry out the sequence is made by a person, not a machine.

6.6.2.3 Industrial Automation

Industrial automation comprises many different controllers which, by working together, enable production plants to function. The components of an automation system used in production are usually assigned to different levels (DIN EN 62264 2008). The lowest level houses the sensors and actuators used to detect and change the state of a production plant. Sensors and actuators are linked, either directly or via fieldbuses, to device controllers. Depending on the machine type to be controlled, in manufacturing technology either a programmable logic controller (PLC), numerical control (NC), robotic control (RC) or motion control (MC) will be used. Each machine can then be controlled by one device controller or more. Groups of machines are combined to form cells, which are coordinated using device controllers. Depending on its size, an entire production plant comprises several cells and is controlled by a manufacturing execution system (MES).

The most commonly used device controllers are designed to meet the demands of "traditional" industrial automation. They are programmed using languages that are partially reminiscent of circuit diagrams, or using machine-like programming constructs that provide maximum control over the relevant device or group of devices (DIN EN 61131-3 2003).

The models and algorithms necessary for self-optimisation are not currently part of device controllers. It is only at the control level that "intelligent" planning algorithms come into play (Brecher et al. 2008a) and the concept of multi-agent systems is used (Brennan 2003). The prevailing programming paradigms in automation result in rigidly linked processes at the cell level and the device level.

In automation the term "intelligence", applied to the lower levels, is often taken to mean adaptive control. However, intelligent decisions on targeted action, which are necessary for assembly, are taken at the cell-control level and, in automated solutions, implemented in controlled movements. As a rule, using PC-based cell controllers (Possel-Dölken 2006; Upton 1992) in industrial environments also makes it possible to use the SOAR cognitive architecture employed in this project.

6.6.2.4 Task Allocation Between Humans and Machines

Looking at the role humans play in standard automated production, we see that their main task involves managing and monitoring the manufacturing system. In the event of a malfunction, they must be able to take over manual control and return the system to a safe, productive state. This concept, termed "supervisory control" by Sheridan (2002), involves five typical, separate subtasks that exist in a cascading relationship to one another: plan, teach, monitor, intervene, learn (see Fig. 6.106).

After receiving an (assembly) order, the human operator's first task usually involves planning the assembly process. To do so, he or she must first understand the functions in the relevant machine and the physical actions involved to be able to construct a mental model of the process. Using this basic understanding, the operator then develops a concrete plan that contains all specific sub-targets and tasks necessary. "Teaching" involves translating these targets and tasks into a format that can be used for machine control—e.g. NC or RC programs that facilitate a (partially)

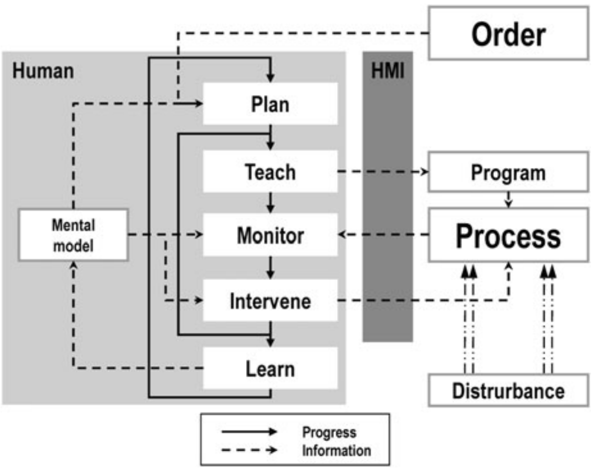


Fig. 6.106 Supervisory control. (Based on Sheridan 2002)

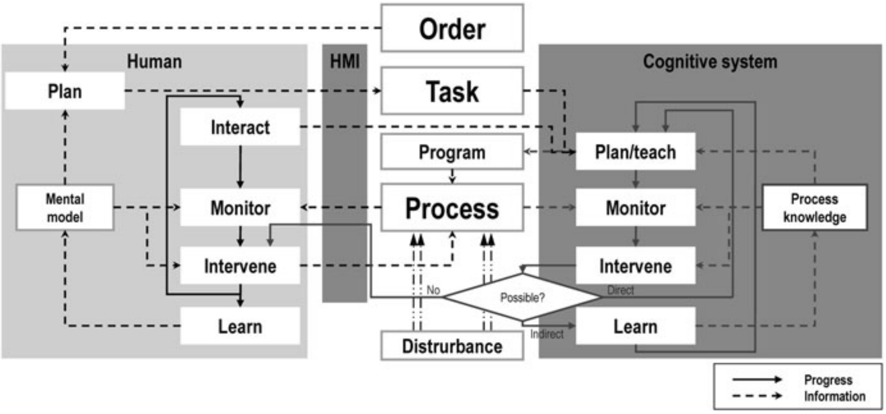


Fig. 6.107 Extended supervisory control approach for cognitive systems. (Based on Mayer et al. 2008)

automated process. This process must be monitored to ensure that it runs properly and produces products of the desired quality. The expectations for the process are drawn from the mental model the operator created at the start. In cases where reality deviates significantly from the model or where there are anomalies, the human operator can intervene by modifying the NC or RC program or by manually optimising the process parameters, for example. Ultimately, every intervention involves the human operator continually adapting his/her mental model, while existing process information, characteristic values and trend analyses help the operator better understand the process and develop a more detailed mental model.

With a cognitively automated system, the tasks change gradually, but in a conceptually relevant way (see Fig. 6.107). Because the cognitive control unit (CCU) can independently solve a certain class of rule-based production tasks, the human operator is relieved of performing repetitive, monotonous or very dangerous tasks. In a cognitive production system, the human operator defines the assembly tasks based on the status of the sub-product or end product, carries out adaptations or sets priorities as needed, compiles rough process plans, and sets initial and boundary conditions. The information-related pressure on the human operator is considerably reduced in the areas of detailed planning and teaching, because the cognitive system handles them. But shifting this load from the human to the machine can result in the human operator forming an insufficient mental model of the state variables and state transition functions in the assembly process. This is because knowledge relating to execution is already stored in the system's process knowledge. At the same time, however, humans must monitor system status and dynamics, and possibly make decisions based on this knowledge. Especially in the event of an error that the system cannot identify or solve, the human operator must receive all information relevant to the situation in an easily understandable form so that he/she can intervene correctly and enable system recovery.

6.6.3 Motivation and Research Question

In high-wage countries, automating production systems can cover over 70% of their functions. Given the law of diminishing marginal utility, raising the degree of automation even higher will not necessarily lead to a significant increase in productivity. Although automation can, as a rule, reduce the frequency of process errors, it also causes a disproportionate increase in the possible consequences of a single error (Kinkel et al. 2008). These relationships, which Lisanne Bainbridge incisively referred to as "ironies of automation" (Bainbridge 1987), are represented in Onken and Schulte (2010) as a negative feedback loop (vicious circle). To circumvent human shortcomings, a function that humans originally performed is automated. This increases the complexity of the system, which in turn places greater demands on the employee responsible for monitoring the automated function. The result is that the entire system potentially becomes less robust. The loop comes full circle when humans attempt to use automation again to compensate for these possible weaknesses. While it is not uncommon for an automated system's productivity to increase during the first iteration of the loop, humans often underestimate or even ignore the risks involved. Onken and Schulte (2010) believe that using mechanisms borrowed from human cognition presents an opportunity to "break the cycle" and design a flexible production system in which humans and machines work together safely and effectively—especially in process planning and monitoring, and in disturbance management.

A system that is capable of learning and of adapting to changing environmental conditions can increase planning efficiency by reusing acquired knowledge and transferring it to similar, new production cases. This can sometimes considerably reduce the number of iteration steps involved. Systems of this kind are known as self-optimising systems (Frank et al. 2004). Self-optimisation requires cognitive capabilities that, in current production processes, only humans possess.

In this context, the term "self-optimising production systems" describes a concept that can implement value-stream-oriented measures at the same time as increasing planning efficiency and improving process and product quality (Schmitt and Beaujean 2007). Transferring existing knowledge to similar or new production cases—the essence of self-optimisation—opens up new perspectives for production and assembly systems by enabling them to dynamically adapt system behaviour to keep pace with changing targets and situations. Incorporating humans' unique skills and experiences into the system is considered essential to self-optimisation. Innovative cognitive functions in the form of symbol-processing systems should support humans and, where necessary, relieve them of routine tasks.

Here, cognition is understood to mean processes such as perception, knowledge storage, reasoning and learning. Obviously software can only partially simulate the unique features of human cognition, but some models (Strohner 1995) can be partially transferred to technical systems and thereby provide a suitable basis for self-optimisation. Thus, cognition can largely be described as referring to the transfer and application of knowledge, and to the processing of information, either by a living being's central nervous system or in an artificial system (Strohner 1995). Within this context, the sub-project described here, which is concerned with self-optimising production systems, focuses on designing and realising a prototype of a CCU. This CCU can use symbolic knowledge representation to optimise itself according to predefined criteria. Most importantly, however, the CCU can be designed, developed, and operated safely and efficiently by highly qualified experts in a high-wage country like Germany.

Developing the design starts with the dilemma of planning orientation and value orientation in the polylemma of production (Schuh and Orilski 2007). In planning-oriented production, processes in a manufacturing system are centrally planned during operations planning and scheduling, in great detail, far in advance, and in line with the Taylorist principle of separating preparation and execution. Doing so makes it possible to closely align production steps with the overall target because all activities are analytically derived from the desired end result using a global target function. In value-oriented production, planning activities overlap with the actual value-adding process. In addition to carrying out the activity directly related to value creation, the person responsible for production defines sub-tasks, their sequence, and the use of production resources. The overall target is therefore generated in collaboration: organisational units independently define their sub-targets and sub-tasks, and these come together along the process chain. This approach has the benefit of allowing the production system to respond quickly to changing boundary conditions, which means that it can better handle the complexities and dynamics of its environment and the process itself. State-of-the-art value-adding chains are typically only partially capable of independently finding top-down solutions to specific problems on the basis of simplified models within a defined solution space. As a rule, these chains do not fully take into account interactions between processes, materials, production resources and the people working in the environment, meaning that knowledge of these interactions is usually incomplete. The same is true of how these interactions affect the design of the product structure.

Cognitively automated systems should pave the way for new concepts and technologies for production and assembly systems that should be able—through continuous data analysis, information fusion, interpretation and assessment of the actual situation—to dynamically adapt themselves to changing targets and boundary conditions. The research question that arises from this is: how can we achieve a highly dynamic system while at the same time ensuring that the targets of all activities are well-synchronised?

The solution posited in designing and using cognitive automation is as follows: a cognitively controlled production system reacts faster, more reliably and in a more resource-efficient way than a production system that uses traditional planning logic and methodology. Unlike a standard control system, a cognitively automated system can, on the basis of internal decisions, independently redefine reference variables in terms of targets, and adapt the control strategy accordingly. It would, however, be naïve to believe that such a system could function completely autonomously. Scientific research must focus on designing and producing prototypes of cognitively automated production systems that can be further developed and efficiently operated by highly qualified skilled workers in high-wage countries. These are no longer purely technical systems; they are complex human-machine systems that require an ergonomic design. Therefore, in defining the research question, we can identify two fields of activity:

> Firstly, this sub-project should address aspects of technical design and evaluating cognitive functions. Secondly, any design for this kind of future production system must focus squarely on humans and their superior cognitive skills.

The Cognitive Control Unit sub-project therefore explored the following research questions:

- Design and development of architecture for a cognitive control system: The requirements placed on a cognitive machine control system are reflected in the unique requirements placed on the software architecture of the system. The sub-project therefore investigates how this kind of architecture should be structured and designed. This involves taking into account the different time requirements for machine-oriented control systems (e.g. a robot cell) and planning for various data abstractions. The data flows and information shared between planning, control and human-machine interface must also be defined. A cognitive system works using a knowledge-based approach. This means that the software architecture must include components that allow the system to save and modify knowledge that machines can process and humans can understand.
- Design and development of a planning methodology for cognitive automation: Assembly planning presents a highly complex problem, even for current planning systems. If it also has to be flexible enough to adapt to unexpected events and ad-hoc changes in the process, we need to find special techniques for assembly planning and control that incorporate functions borrowed from human cognition.
- Usefulness of technical cognition in industrial automation: The numerous functions that have to be controlled in a production plant are predominantly designed to effectively and efficiently carry out a pre-defined process. Alongside the primary task (e.g. assembling a workpiece), supporting tasks also have to be controlled, such as transporting and handling the workpieces. To achieve end-to-end self-optimisation and control, all levels of industrial automation require cognitive functions. The sub-project therefore seeks to find out how to integrate cognitive functions into industrial automation, both conceptually and in terms of technical implementation.
- Human-centred design of a cognitive control system's knowledge base: A control system that has a knowledge representation which allows it to plan almost independently and according to the situation will have a considerable impact on the spectrum of tasks performed by human operators. Starting with the human role in this kind of cognitively automated production system, the sub-project uses robot-supported assembly to show how a human-centred CCU design can make system behaviour more compatible with the human operator's expectations. In doing so, the sub-project aims to develop a safe, productive and disturbance-free design for work processes in cognitively automated work systems.

- Ergonomic interface design: Another question relates to the way information is presented to human operators. One area that this research focuses on is ergonomically designing head-mounted displays and displays in workspaces using new methods of visualisation and interaction. In contrast to standard input and output media like keyboards and TFT screens, new technologies for user interfaces based on head-mounted, semi-transparent LCoS displays are being developed for use in production environments. They are also being ergonomically designed and evaluated with regard to their potential under operating conditions.

6.6.4 Results

6.6.4.1 Software Architecture

Russell and Norwig's three-layer model (2003), widely used in robotics, was chosen as the basic framework for the architecture. Compared with other architectures commonly used in this field, such as the blackboard model (Hayes-Roth 1985), a three-layer architecture has the advantage of a clear demarcation between abstraction levels and temporal demands. The planning layer operates on a high abstraction level with symbolic problem definitions and must satisfy only soft real-time demands. The reactive layer, on the other hand, has to monitor machine-related control loops in "hard real-time". The coordination layer mediates between these two layers. This is where abstract instructions from the planning layer are transformed into concrete machine control commands. In the reverse direction, the information from the various sensors is aggregated to form an overall picture of the situation and is transmitted to the planning layer as a basis for decisions.

In order to satisfy the demands of a holistic consideration of the human-machine system, the classic three-layer architecture was expanded to include further layers and modules (see Fig. 6.108)

The presentation layer forms the interface to the human operator. The interactive goal definition and description of the task, as well as the presentation of the current internal state of the CCU takes place by means of the human-machine interface. In addition, external data formats such as CAD data of the product to be assembled, are transmitted in internal representation forms and made available to the other components. The knowledge module contains the knowledge base of the CCU in the form of an ontology (Gruber 1993). The planning layer can send enquiries to the knowledge module by transmitting the system state received from the coordination layer. The knowledge module then analyses the objects contained in this state and can derive further information by means of reasoning via the ontology and transmit this to the planning layer. A further logging module forms the database for all the other components. All the data generated during system operation are persistently stored here so that in the event of an error, the cause can be reconstructed on the basis of the data if necessary. To support system transparency, all the data stored in the logging module during operation are accessible to the user via the presentation level.

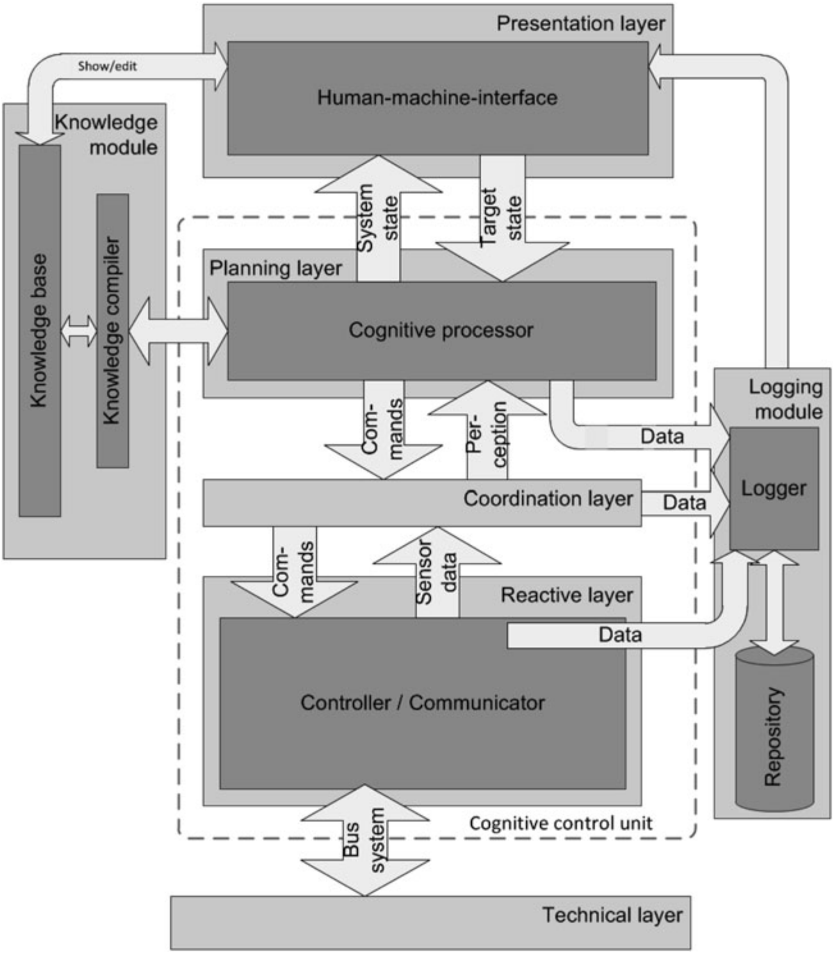


Fig. 6.108 Architecture for a cognitive planning and control unit

Furthermore, the data thus gathered can also be used for training measures that are integrated directly into work processes (embedded training, Odenthal et al. 2007).

The cognitive architecture SOAR, whose internal knowledge base is structured in the form of production rules (if-then rules), was chosen to simulate cognitive functions in the planning layer of the cognitive controller (Leiden et al. 2001; Langley et al. 2004). Compared with emergent systems such as artificial neuronal networks, a rule-based approach has the advantage of not needing time-consuming and potentially unreliable preconditioning. To a certain extent, SOAR is able to simulate rule-based human decisions and to take over repetitive and monotonous process steps (Hauck et al. 2008). It cannot, however, simulate genuine knowledge-based behaviour in the sense of reflecting on goals and their prioritisation (sensu Rasmussen 1986).

As SOAR was not designed for automation applications and has no interface to industrial control systems, a framework was designed and developed which makes it possible to model the knowledge and the corresponding algorithms necessary for controlling the assembly and the logistics (Kempf 2010). Furthermore, it provides the architecture required for system control and the necessary interfaces. The planning layer, coordination layer, reactive layer and technical layer were developed for this purpose (see Fig. 6.108). The technical layer essentially consists of the kind of control units used in today's manufacturing cells. The character of the control interfaces is thereby more or less predefined: a robot controller, for example, processes travel commands with the pattern "Move linearly from A to B at speed C"; a PLC expects switching commands with the pattern "If input A is true, then set output B"; a special image-processing software that is located in the reactive layer due to its complexity and the real-time demands imposed by the application expects an order such as "Start object detection and feed back result". The function of the technical layer is therefore to control the individual devices via a defined interface protocol with defined semantics. The main functions of the planning and coordination layers involve dynamically generating the individual device-specific control commands from a global and relatively abstract description of the task, and coordinating their execution in accordance with feedback from the sensors. One example of a component in the reactive layer is the image-processing software mentioned, which, due to its frame rates, has to be located in this layer.

A SOAR agent that takes control decisions independently consists of the model of a given problem space and the mechanisms for its processing provided by SOAR. Each agent has its own long-term and short-term memory, and its own input and output areas, where it operates cyclically, in a manner comparable with the function of a PLC. The planning layer contains SOAR agents which are responsible for selectively assigning control instructions for the production system. One such "dispatcher" thus represents an order. The coordination layer contains agents which execute the control instructions, whereby each operation is represented by a dedicated agent, the "executor". This split between several agents takes into consideration the aspects of hierarchisation and parallelisation (Kempf 2010, see Fig. 6.109).

However, a control architecture is characterised not only by layers and interfaces, but also by the way in which the internal temporal sequence is organised. The cognitive control architecture is generally based on a fixed processing cycle, comparable with the well-known behaviour of a cell controller in the form of a PLC. Similarly

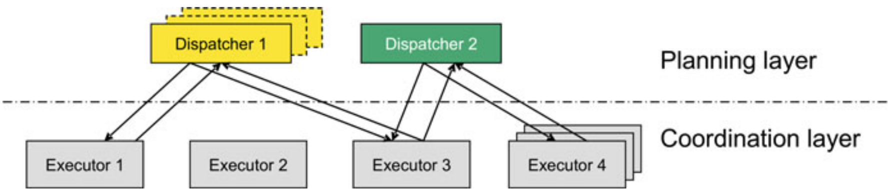


Fig. 6.109 Arrangement of the SOAR dispatcher and executor agents in the layer model

to that, the length of a processing cycle in this case also depends on the number and complexity of the active agents. Since there is no real-time operating system, no hard real-time can be achieved here. But this is generally not necessary for the execution and decision layers. The real-time-critical processing takes place in the reactive and technical layers. The SOAR agents are instantiated as independent objects of the agent class. First the active dispatchers are executed sequentially, then the executors. This means that decision and execution phases are always synchronous with one another. The same applies to the input and output or simulation. As a result, all the agents operate with the same image of the environment during a control cycle, and are therefore quasi-parallel; this also applies to the exchange of messages between the agents. A relatively simple but efficient scheduling policy is employed for the individual agents: an agent continues to receive computing time until it generates an output to the controller. This is frequently the case after one SOAR decision cycle, but in some situations the agent requires several cycles before it reaches a decision. All the agents run in one thread with the SOAR kernel and are processed in a fixed sequence. This process is flanked by two additional measures:

1. To avoid endangering the control functions at execution level, measures must be taken to ensure that an agent does not take too long (e.g. during virtual planning). This is achieved by ensuring that it is always the next agent's turn after a specified maximum number of decision cycles have elapsed. This also means that decision processes of individual agents can extend over several control cycles.
2. As only a limited number of resources are available at the decision level, a dispatcher at the front of the execution sequence tends to have an advantage. Users can prioritise production rules relative to one another. However, if they do not do this, the process follows the principle of "fairness" and the execution sequence in each SOAR cycle is defined at random.

Asynchronously to the actual processing, user inputs can be made for process control and visualisation, with semaphores ensuring data consistency.

6.6.4.2 Hybrid Method for Assembly Planning and Control

Starting from the general function of the CCU, which involves carrying out the planning and control of a robot-supported assembly process on the basis of production rules in conjunction with a formal-mathematical product model, this chapter initially deals with only the cognitive processor. As already mentioned, this is located in the planning layer. In the validation study presented here, the only input the CCU receives is a description of the finished product to be manufactured. "Action primitives" stored in the knowledge base of the CCU in the form of production rules serve as control commands for the industrial robot for component assembly, and are also linked in the processing cycles by the processor, depending on their state, to create a complete and efficient assembly process.

Simulation experiments investigated the influence of various factors on the dependent variables "CCU processor time" and "number of assembly cycles" in the whole process of assembling the target product (MTM-1 cycles, see Sect. 6.6.4.4). The target product consisted of identical parts. The independent variables of the simulation experiment are: (1) The size of the target product (six levels: 4, 8, 12, 16, 20 or 24 parts); (2) The number of parts fed in on the queue (seven levels: 1, 4, 8, 12, 16, 20 or 24 parts); (3) The type of feed-in (two steps: deterministic feed-in of the parts required or random feed-in, including parts not required). For the feed-in, a simple buffer in the form of a queuing model was used that is operated similarly to the FIFO principle. One hundred simulation runs were calculated for each of the $6 \times 7 \times 2$ combinations of factor steps. Given the high amount of computing time required, the simulations were performed on the RZ cluster in the computer centre at RWTH Aachen.

The simulation results show that the target product was correctly assembled in all 8,400 simulation runs. No assembly errors or blocking ("deadlocks") of the cognitive processor occurred.

The CCU processing time and the number of assembly cycles for the random feed-in of the parts are shown in Fig. 6.110. If we first consider only the time required, we can see that the processing time increases when more parts are used in the target object (Fig. 6.110, left). However, looking at the number of parts on the queue at the same time, it is surprising to note that processing time decreases as the number of parts increases. If we consider the number of simulated assembly cycles for a target product of a given size and for feed-in with given parts, we see that the anticipated number of assembly cycles is obtained (Fig. 6.110, right).

The corresponding results for the deterministic feed-in of parts are shown in Fig. 6.111. The simulation results clearly show a disproportionate increase in processing time as the size of the target product and the number of parts on the queue

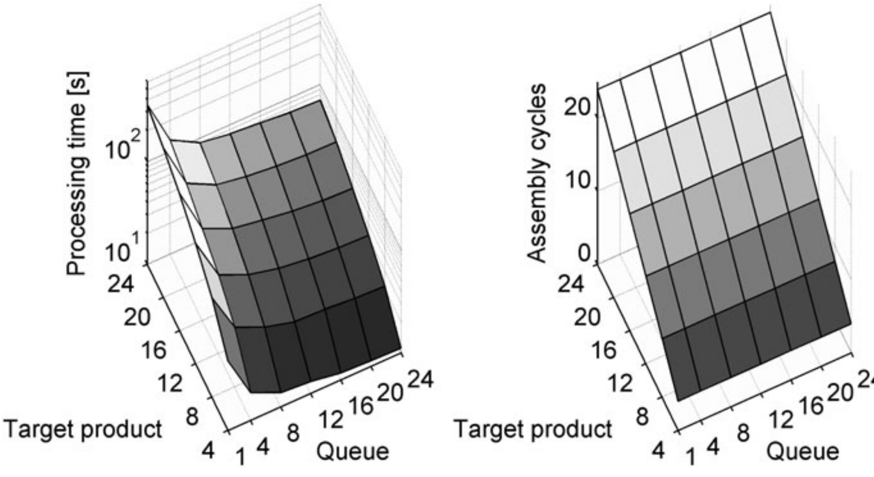


Fig. 6.110 CCU processing time (*left*) and number of assembly cycles required (*right*) as a function of the size of the target product and the number of parts fed in on the queue in random feed-in

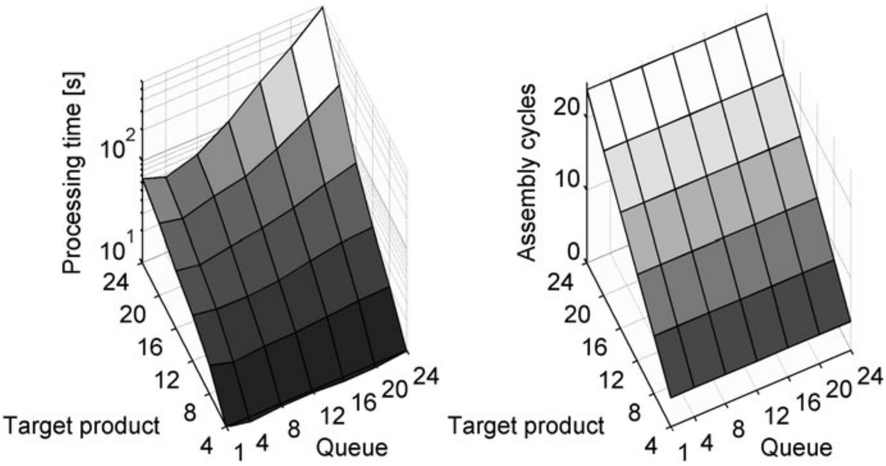


Fig. 6.111 CCU processing time (*left*) and number of assembly cycles required (*right*) as a function of the size of the target product and the number of parts fed in on the queue in deterministic feed-in

increase (Fig. 6.111, left). The number of assembly cycles, on the other hand, behaves as expected in relation to the size of the target product (Fig. 6.111, right).

The results of the simulation study show that a CCU based on SOAR cognitive architecture is able to reliably perform assembly planning. In the given application, SOAR proves to be particularly suitable for reactive planning under the boundary condition of random feed-in of parts. In the case of deterministic feed-in, however, the study showed exponential run-time behaviour. Although difficult to understand intuitively, this simulation result can be explained by the way SOAR functions during the decision-making process. It compares each required part on the queue with every possible position within the target product. These comparisons result in "proposals" that provide the basis for decisions on executing action primitives. Due to the combinatorics, this can lead to a non-polynomial increase in run-time behaviour (see, e.g. Barachini 1990).

If we want to achieve dynamic system adaptation to changing boundary conditions using—instead of SOAR—generic planning algorithms such as the Fast-Forward Planner (Hoffmann 2001), the only way to do so is by performing continuous re-planning or by completely forward planning all possible sequences of fed-in parts. Complete forward planning for products comprising just 15 parts would require drawing up over 3.6 billion plans. Forward planning can therefore no longer be controlled in normal cases by combinatorics. Continuous re-planning during assembly is also not possible, since combinatorics cannot satisfy the real-time demands for comparable reasons. By contrast, the concept presented in this paper of a CCU based on SOAR meets the demand for dynamic system adaptation very well, as the simulation results for random feed-in clearly show. However, the results also show that, unlike generic planning algorithms, the CCU has only limited suitability for classic deterministic planning tasks. The study observed, for example, exponential run-time behaviour with a strictly deterministic feed-in of required parts.

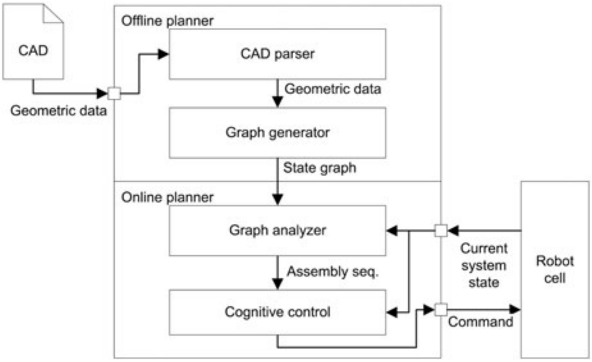


Fig. 6.112 Hybrid plan of the assembly planner

For this reason, a more advanced hybrid solution was developed (see Fig. 6.112). It involves making a distinction between an “offline planner”, which computes complex planning steps based on geometric analyses prior to assembly, and an “online planner”, which uses the results of the offline planner to generate assembly plans during the assembly process, depending on the current situation.

The offline planner consists of two components: the CAD parser and the graph generator. The parser receives the CAD description of the product to be manufactured and extracts from this the geometric information on the whole product and the individual components. This information is then forwarded to the graph generator. The generator follows an “assembly-by-disassembly” strategy (Lin and Chang 1993) to generate all the expedient assembly sequences for the given product. This involves starting from the fully assembled state and generating all possible disassembly steps recursively, until there is nothing but individual components. Read backwards, these disassembly steps show all the possible assembly sequences. The graph generator performs this process using the method presented by Thomas (2008). The system examines pairs of all subsets of components to establish whether these can be separated from each other, collision-free. Invalid disassembly steps are rejected, while valid steps are added to an and/or graph (Homem de Mello and Sanderson 1986) and further examined and evaluated on the basis of various criteria. An and/or graph for a simple tower made of four bricks is shown in Fig. 6.113.

Each hyperedge of the and/or graph represents one disassembly action, or, in the opposite direction, one assembly action. As some assembly actions have to be performed by different tools or by a person, different costs depending on the type of action are assigned as vectors to each hyperedge. The finished and/or graph is then transformed into a state graph (see Fig. 6.114) because the online planner uses a process—similar to those applied in assistant systems (Zaeh and Wiesebeck 2008)—that cannot be used for hyperedges.

During the transformation, each hyperedge is converted into a simple edge. The corresponding nodes, which represent the resulting disassembly steps, are merged into one node: if both nodes consist of more than one individual component, a new node is generated that contains both sub-products. If a node contains only one component, this node is removed and the respective component is added to the new

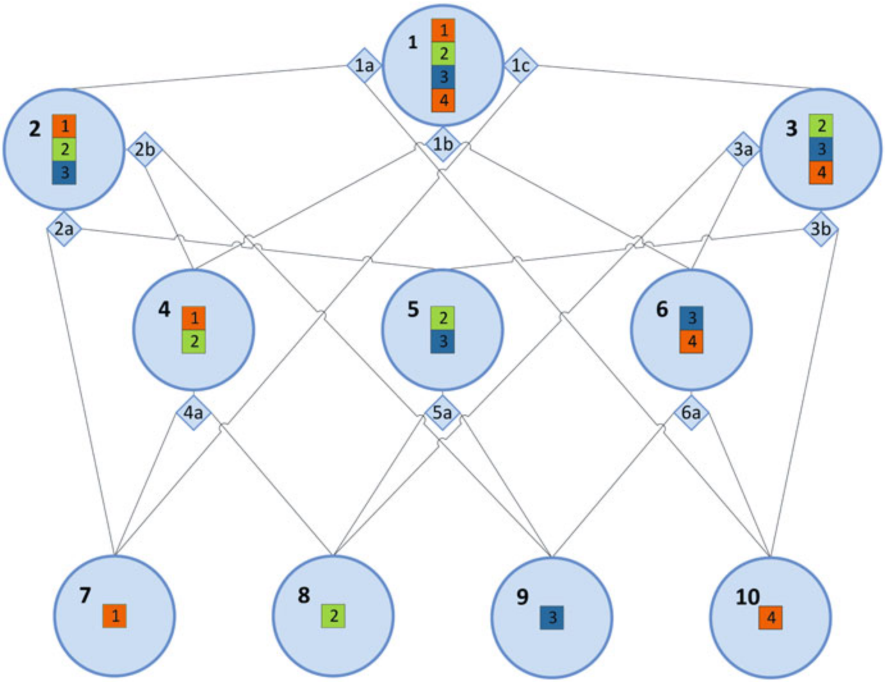


Fig. 6.113 And/or graph for a tower of bricks

edge as a requirement. The resulting state graph is then transferred to the online planner so that assembly can begin.

Within each assembly cycle, the current system state is transmitted to the online planner during assembly, the next action to be performed is calculated, and the robot cell is activated accordingly. Decision-making is carried out by three components: the graph analyzer, the paralleliser and the cognitive control (CC). The graph analyzer contains the system state and updates the state graph. First, the current state of the assembly cell is located in the diagram. Then all the following edges are updated. Edges whose required components are not contained in the current state receive additional "penalty costs". Edges further from the node describing the current state are assigned lower costs. The online planner thus gains the ability to speculate. Actions that are not possible at present, because the corresponding components are not available, are more likely to be integrated into a plan the further they lie in the future. The algorithm "hopes" that the required component will be delivered before it is required.

After updating the state graph, the least expensive path—according to the given target system—from the current state to the target node is calculated using the A* algorithm (Hart et al. 1968). When passing through the graph, A* selects as the next node to be examined the node x, for which the function $f(x) = h(x) + g(x)$ is minimal. The function $g(x)$ designates the costs of the path to be taken to arrive at node x.

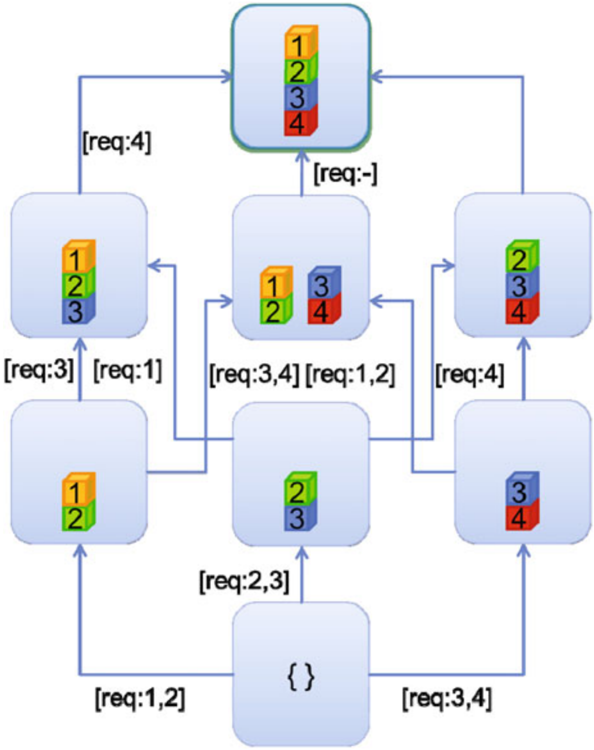


Fig. 6.114 State graph produced from the and/or graph

In addition to the costs of the individual edges of the graph, the cost calculation also takes into account any necessary tool changes. In view of prevailing safety regulations, higher costs are allocated to changes between robots and humans than to changes between tools. Paths with fewer changes therefore have lower costs. The function $h(x)$ is a heuristic which estimates the proximity of a node to the target node. The heuristic employed here uses the number of components already correctly assembled and the valuation assigned by the offline planner with respect to machine transparency. The path calculated in this way is passed on to the next component as the plan to be followed. The process described ensures that this is an optimal plan with respect to several criteria:

- High penalty costs for actions that cannot be carried out ensure that the online planner selects the most realistically probable assembly sequence.
- Additional costs for tool changes create a preference for choosing sequences where the same tool is used for longer periods or which can be carried out by a single human operator.
- Reductions in penalty costs for assembly steps that cannot be carried out, depending on their distance from the current state, create a preference for assembly sequences that can be started immediately, because actions that cannot be carried out are shifted to the end.

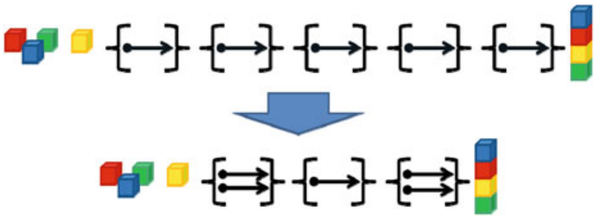


Fig. 6.115 Assembly plan as a sequence of sets of assembly actions—before and after parallelisation

A further optimisation step to accelerate assembly is performed in the paralleliser component. The fixed sequence of assembly actions received is examined for actions that can be carried out in parallel. The sequence is thereby seen as a sequence of sets of assembly actions. Each of the sets is examined as to whether it can be carried out in parallel to the following set, i.e. whether all elements of both sets can be carried out in pairs at the same time. This parallelisation can be derived from the and/or graph: two actions, x and y, can be carried out exactly in parallel when their corresponding hyperedges—e_x and e_y—have a common predecessor hyperedge and when neither e_x nor e_y is a predecessor of the other. If two sets can be carried out in parallel, they are joined and examined to see whether they can be carried out in parallel with the next set. The result of this process is a sequence of sets of assembly actions (Fig. 6.115) which is then transferred to the decision-making component, the CC (Fig. 6.112).

Transferring a plan as sets of actions that can be performed in parallel has a number of advantages. On the one hand, assembly can be accelerated because several robots or even the human operator can work on the assembly at the same time. On the other hand, the CC has greater freedom for decisions because it is free to select the sequence in which the actions within a set are to be processed (if they cannot be performed in parallel).

As already mentioned, the CCU is based on SOAR, a cognitive architecture which attempts to simulate the human decision-taking process. The schematic sequence within this component is shown in Fig. 6.116. During assembly, the CC receives the current system state, transmits it to the online planner and receives back the assembly plan described above. On the basis of this plan and the current state, a decision on the next action is taken using the stored rule base. The CC can decide to follow the given assembly plan and to execute a suitable action from the first set of actions at its own preference. Alternatively, it can decide to ask the human operator for assistance, or to wait until the situation changes, for example through a new component being fed in. After executing the respective action, the resulting system state is queried and checked. If it corresponds to the target state, i.e. it contains the finished product, the assembly process ends. Otherwise the system examines whether the state has changed purely as expected, i.e. as a result of the action performed. If this is the case, it continues processing the assembly plan already received. If the situation has changed unexpectedly, the CC initiates re-planning by the online planner.

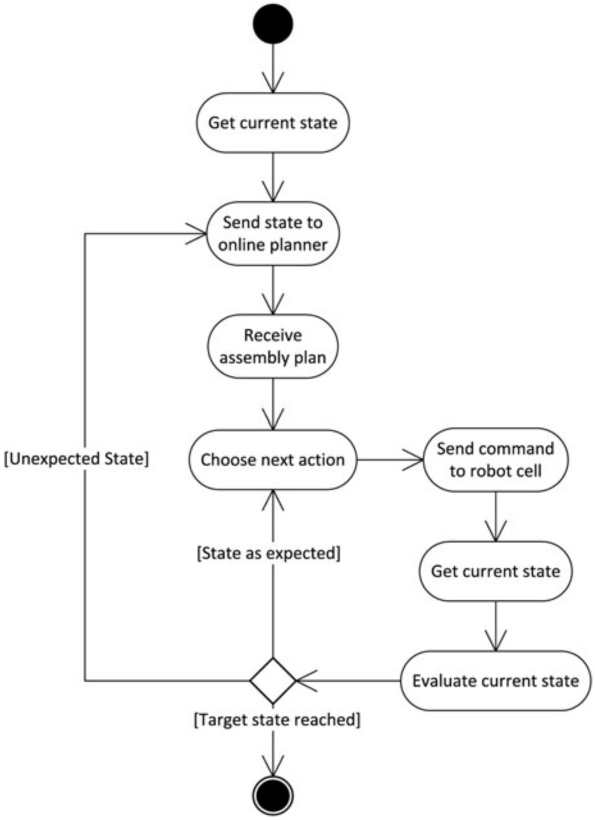


Fig. 6.116 Sequence of information processing in the CCU

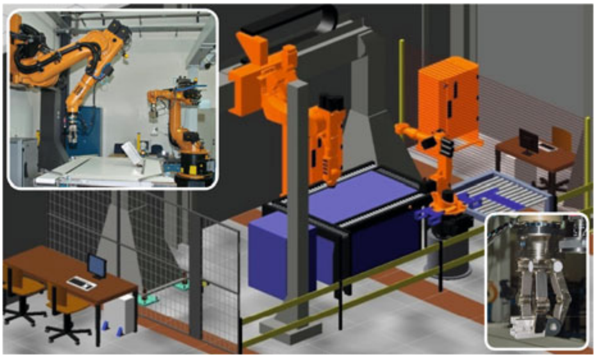
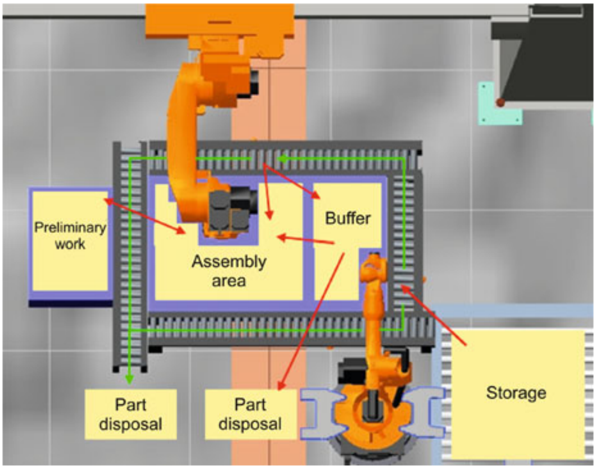


Fig. 6.117 Layout of the assembly cell

6.6.4.3 Prototype Realisation of a Cognitively Automated Assembly Cell

To test and develop a CCU in a near-reality production environment in a variety of different assembly operations, a robotic assembly cell was set up (see Kempf et al. 2008). The layout of this cell is shown in Fig. 6.117. The scenario was selected to

Fig. 6.118 Assembly and storage areas in the assembly cell



comprise major aspects of an industrial application (relevance), and at the same time to easily illustrate the potential of a cognitive control system (transparency).

The main function of the demonstrator cell is the assembly of known objects. Part of the cell is made up of a circulating conveyor system comprising six individually controllable linear belt sections. Several photoelectric sensors are arranged along the conveyor route for detection of components. Furthermore, two switches allow components to be diverted onto and from the conveyor route. Two robots are provided for handling the components, with one robot travelling on a linear axis and carrying a tool (a flexible multi-finger gripper) and a colour camera. Several areas were provided alongside the conveyor for demand-driven storage of components and as a defined location for the assembly (see Fig. 6.118). One area is provided for possible preliminary work by a human operator. This is currently separated from the working area by an optical safety barrier. The workstation has a multimodal human-machine interface that displays process information ergonomically, allowing it to provide information on the system state and to help solve problems, if necessary. Detailed information on the configuration of the multimodal interface can be found in, for example, Odenthal et al. (2008, 2009), and in Schlick et al. (2009). To simultaneously achieve a high level of transparency, variability and scalability in an (approximate) abstraction of the actual assembly process, building an assembly of LEGO Duplo bricks was selected as the assembly task. To take into account the criterion of flexibility for changing boundary conditions, the bricks are delivered at random (see Sect. 6.6.4.2). In terms of automation components, the system consists of two robot controllers, a motion controller and a higher-ranking sequencer. The latter takes the form of a CCU.

The initial state provides for a random delivery of required and non-required components on a pallet. One of the robots successively places the components onto the conveyor. The automatic-control task now consists in coordinating and executing the material flow, using all the technical components, in such a way that only the assembled product is on the assembly table at the end.

As has already been explained, the automatic generation of valid assembly sequences, e.g. directly from a CAD model, is an extremely complex problem for which a universal solution has still to be found. In the present case, however, it is possible to find valid sequences with reasonable computing time that can be generated using the hybrid planning process from Sect. 6.6.4.2. The result of this kind of upstream planning process is a set of sequences extracted from an assembly priority graph that serve the dispatcher as inputs. The assembly planner described in Sect. 6.6.4.2 is also involved in solving the assembly problem. Even for an assembly problem of average complexity, a number of additional boundary conditions have to be observed when planning the assembly graph in order to arrive at a valid assembly sequence. Such limitations exist particularly in the following basic operators, which are based on the MTM-1 taxonomy (Drumwright et al. 2006; see Sect. 6.6.4.4):

- REACH: Does the assembly situation, in combination with the gripper geometry, permit a valid approach trajectory?
- GRASP: Does the assembly situation, in combination with the gripper geometry, make it possible to grip the component during positioning (joining) (see Fig. 6.119, left)?
- POSITION: Does the assembly situation permit movement in joining direction (see Fig. 6.119, centre)? Does the assembly situation permit a stable force couple between the component to be positioned/joined and the assembly (see Fig. 6.119, right)?

The necessary call parameters have to be provided for commanding the functional units via the device interface. Important poses within the cell are preconfigured and are processed within the coordination layer. However, the exact target pose might only emerge at run-time, when a position is located on the pallet, for example. Operation always takes place in a discrete problem space. A component can be deposited on the pallet in, e.g. a given pattern that is stored as part of the plant model. The executor

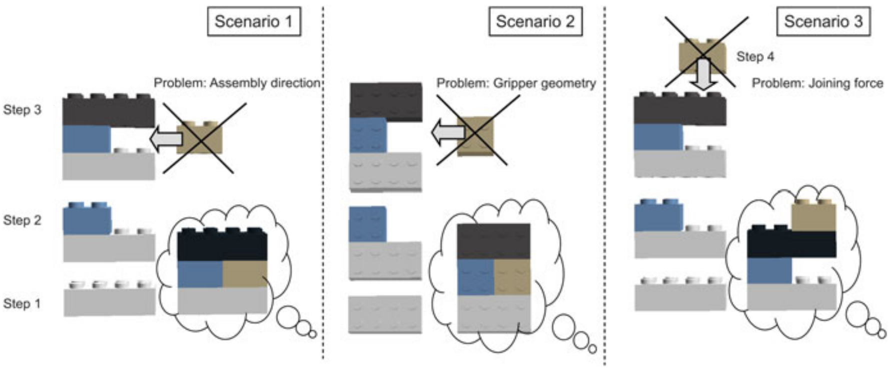


Fig. 6.119 Boundary conditions in positioning the demonstrator components

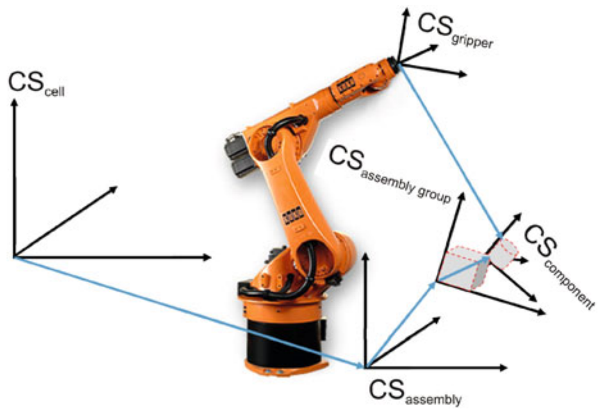


Fig. 6.120 Coordinate systems and transformations

now supplies a discretised frame that corresponds to an actual (relative) component pose and has to be scaled to the modelled grid and offset against the reference pose. All the executors use this principle. Corresponding discretised poses are:

- Depositing poses of components or assemblies (in the cell coordinate system)
- Joining poses of components in the assembly (in the assembly coordinate system)
- Gripping poses on components (in the component coordinate system)

Frame and transformation operations are used to calculate the resulting target pose. This makes it possible to easily calculate random links between different coordinate systems (see Fig. 6.120).

Atomic commands are used in the technical layer to control the robots, grippers, conveyor belts and switches. The challenge here lies in linking the PC-based cognitive controller to the industrial device controllers.

Before a plant can begin production, it must be sufficiently tested. The advantages of using virtual commissioning for this purpose are increasingly being recognised. A cognitive control system probably has even greater need of simulative testing than a classic system does. This is because system behaviour in this case is not even known "on paper"—according to the requirement, it is not generated until run-time. First the question has to be answered as to how a control framework can make it possible to test the resulting processes in advance and—if possible—to visualise them. One possibility is by using a simulation tool of the kind used in connection with classic commissioning. These kinds of simulation tools generally have an OPC interface that allows them to connect to an external controller (generally in the form of a soft PLC). To avoid having to create several interfaces in the cognitive controller to device controllers and simulation systems, an abstracted interface was used that allows the different types of device controller and simulation software to be linked transparently. This abstracted interface communicates using the real-time-capable CORBA middleware (see Fig. 6.121).

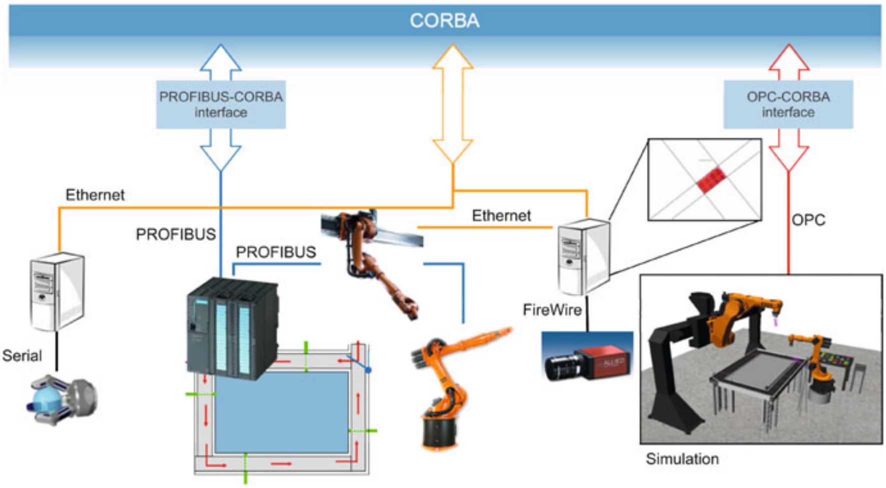


Fig. 6.121 Communication structure in the technical layer

In the technical layer, the function calls of the actuators and sensors with their parameters as control commands are either transmitted directly via the CORBA interface from the cognitive controller to the corresponding PC-based controller, or they are transferred to the fieldbus by a PC-based interface with a corresponding CORBA call. In the demonstrator cell, the following technical sub-systems are controlled via a PROFIBUS interface:

- Two KR C2 robot controllers made by KUKA
- One SIMOTION motion-control system made by Siemens
- Further actuators (two-jaw gripper, pneumatic switches) and sensors (photoelectric sensors) can also be reached as I/O systems via the PROFIBUS. These systems are linked to the cognitive controller via a CORBA PROFIBUS interface.

The following are linked directly via CORBA:

- An image processing system for identifying the individual components and detecting their position
- A gripper system for a flexible gripper, made by Schunk, with seven degrees of freedom

As the components have to be detected and gripped even when the conveyor belt is running, the robot must track them at a speed synchronous with that of the conveyor during the detection and gripping process. To achieve this, a direct control loop between robot controller and image processing system was created. At the technical level, the loop uses the KUKA RSI real-time interface to manipulate robot movements in the interpolation cycle. In the layer model, image processing in the control loop is located in the reactive layer due to its real-time character.

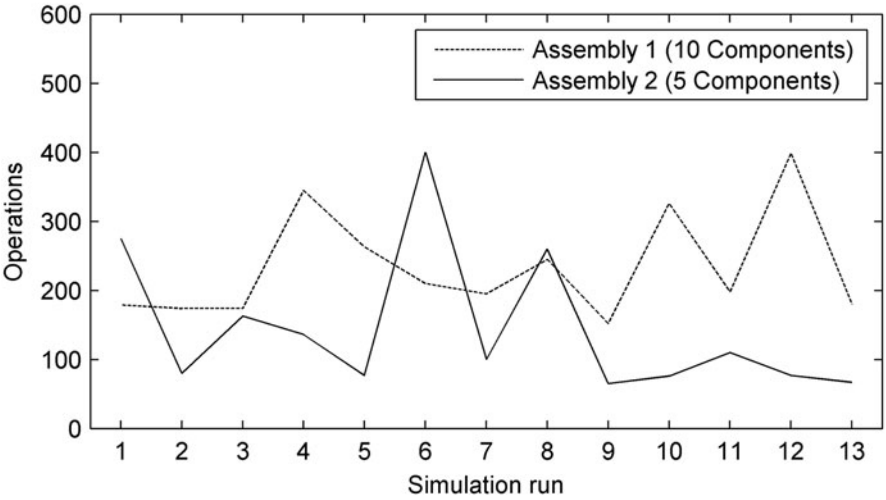


Fig. 6.122 Number of operations necessary for two parallel assembly operations

A truism for every production system is that a plant behaves correctly when it is appropriately designed, and using a cognitive control system does nothing to change this simple fact. However, the example of the demonstrator cell shows that the behaviour achieved with a correct model and self-optimisation completely met the expectations in every case (Kempf 2010). The duration of the process and the number of operations necessary depends to a large degree on the modelling and can be greatly influenced with just a few rules. Further factors are random influences—in this case the random feed-in of components, and non-deterministic decision-making within the SOAR agents. Figure 6.122 shows how the duration of an assembly process is affected by random influences alone.

Due to the semi-decidability of some planning tasks, no hard real-time behaviour can be expected from a cognitive controller at the planning level. Nevertheless, the run-time behaviour should still satisfy certain boundary conditions, particularly in the reactive layer.

As a rule, both the temporal behaviour and the necessary memory requirement are dominated almost exclusively by SOAR (planning processes are one exception; the planning module is responsible for these). The following times were measured on a normal desktop PC with a 2.5 GHz dual-core processor and 2 GB RAM. The control architecture means that there is only one control cycle that covers the planning, coordination and reactive layers. For the measurement of the cycle time it should be noted that the time resolution of the standard operating system (Windows XP) on the platform used was approx. 15 ms. To obtain a more precise statistical mean value, the total run-time was therefore divided by the number of necessary control cycles. It emerged, however, that the actual cycle time varied only very slightly from this mean value. This is also to be expected, because each SOAR agent usually performs exactly one decision cycle per control cycle.

If a normal industrial reaction time of 50 ms is assumed (see Fig. 6.123), the cognitive control system implemented always reacts within this period of time. The

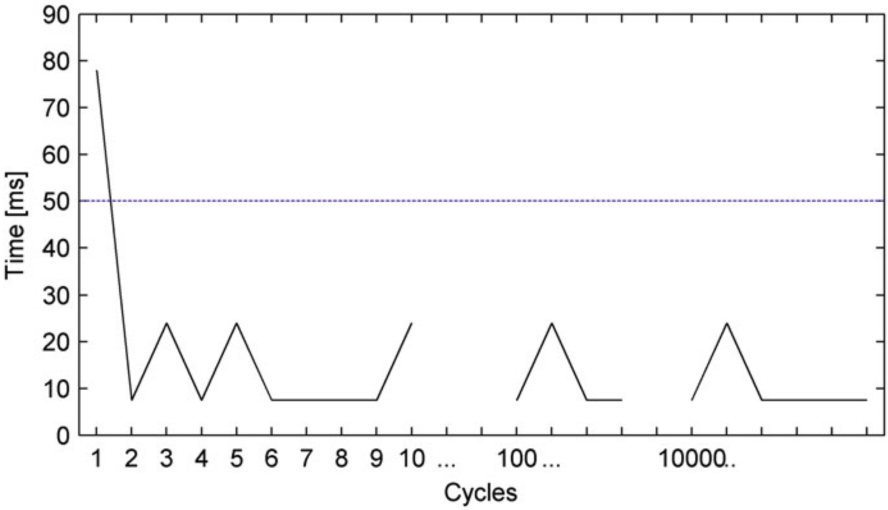


Fig. 6.123 Controller stability and cycle duration compared with normal industrial reaction time (50 ms)

coordination and reactive layers thus initially fulfil the speed demands of a suitable controller for the demonstration cell.

The current results are aimed solely at designing the technical “components” in the complex human-machine system of a cognitively automated assembly cell. From a purely technical point of view, the results demonstrated that robot-supported assembly processes can be cognitively automated, even though, with regard to planning, a hybrid approach had to be developed that combined classic planning elements with reactive, cognitive planning to meet the demands placed on reactive and adaptive planning.

Building on this, the following sections will focus on the human being in the overall system, and will examine how humans and the technical sub-systems interact.

6.6.4.4 Process Logic for a Cognitively Automated Assembly

As explained in Sect. 6.6.2.4, cognitive automation shifts the spectrum of tasks for the human being. Parts of the planning that originally lay in the human sphere of responsibility, because they involve the step-by-step transformation of a manufacturing strategy into an RC program, can be taken over by the cognitive control unit. This can lead to the abovementioned incompatibility between the human mental model and the process knowledge stored in the technical system. With the human and the machine, the working system has two totally different information-processing “systems”, which either encode and process sub-symbolically as with the human, or merely have symbolically encoded information on the production process as with the machine (Fig. 6.107).

The work presented below aims to avoid such incompatibilities by developing a cognitive-ergonomic model of the process knowledge. The idea is to adapt the process knowledge stored in the cognitive control unit to human thought patterns, and to influence the behaviour of the cognitive system in such a way that humans can easily understand and reliably anticipate it.

Elementary components of the MTM-1 taxonomy were used for the cognitive-ergonomic design of the process logic for controlling the assembly robots. The hypothesis is that a sequence consisting of empirically validated basic elements/movements that conform to expectations can be quickly learned and, if necessary, optimised by humans, even if the executing instance is a robot gripper arm (see also Gazzola et al. 2007; cf. Tai et al. 2004). The MTM components transformed into production rules, or SOAR operators, are equal and therefore not defined in a sequence. They correspond to the MTM-1 basic movements REACH, GRASP, MOVE (with integrated TURN), POSITION and RELEASE (see Sect. 6.6.4.3) that are used to control the robots in the cell. In addition, further rules are stored which, depending on the basic elements used, contain the physical boundary conditions (e.g. joining direction or conditions for positioning an element) and assess whether a fed-in element can be directly fitted or has to be stored in a buffer until a later assembly step.

As the research work focuses on evaluating the concept—not on optimising the process purely in terms of time—no tabular time information is stored initially. This simplification is acceptable if all the bricks necessary for the assembly are available or if the component is made of identical bricks, since with fixed starting and finishing positions of the end effector, the sum of the paths and hence the overall time does not change, despite different assembly sequences.

A prototype implementation of the CCU described is carried out in a self-developed simulation environment, very similar to that of the assembly cell. For simplification, the conveyor belt was replaced by a panel similar to a chess board. The fed-in bricks are laid out on the panel's fields. The supply process can be varied from randomly feeding in a single brick through to delivering all the necessary bricks. The simulation included the workstation and buffer areas as independent areas. Regarding the simulation of the gripper, it is assumed that the geometry of the gripper means there are no restrictions in approaching or joining the elements. The prototype simulation environment is shown in Fig. 6.124.

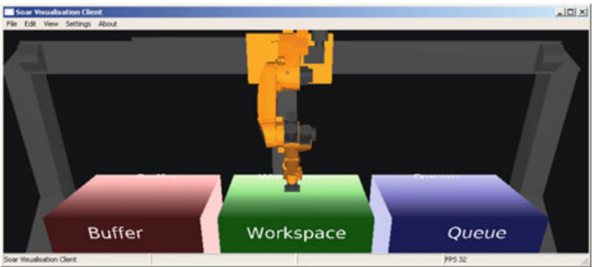


Fig. 6.124 The simulation environment

Table 6.8 Expected process sequence based on the MTM-1 taxonomy

	Step 1	Step 2	Step 3	Step 4	Step 5
Brick 1	Reach Start → Panel	Grip Brick 1	Move Box → Pos. 1	Position on Pos. 1	Release
Brick 2	Reach Pos. 1 → Panel	Grip Brick 2	Move Box → Pos. 2	Position on Pos. 2	Release
Brick 3	Reach Pos. 2 → Panel	Grip Brick 3	Move Box → Pos. 3	Position on Pos. 3	Release
...					

The prototype implementation of the CCU in the simulation environment should be considered as a reference model. The reference model was validated using several easy-to-assemble geometric objects made up of LEGO bricks. The objects differed in the number of bricks used, in colour and in shape (e.g. pyramid, cube, flat surface). The assembly results of a pyramid consisting of 30 identical bricks is explained here as an example. Identical bricks were chosen because this meant that each component could be installed in any position within the pyramid, thus resulting in a large number of possibilities ($\sim 10^{25}$) for the assembly sequence.

Without further defining the spatial dimensions within the workspace of the assembly cell, an assembly sequence like the one shown as an excerpt in Table 6.8 is expected. If we understand one complete cycle as running from the beginning of a REACH operator until the end of a RELEASE operator, we expect the number of cycles to be 30.

Repeated simulation runs (n = 1,000) with pre-picked sets of parts (all the necessary bricks were available) and with a random supply of parts (including bricks not required) show that the desired target object is always built, error-free and in the expected number of complete cycles. Given the simplified treatment of the gripper, no deadlocks occurred. This therefore replicated the simulation results presented in Sect. 6.6.4.2. It should be pointed out here, however, that the variance in the observable assembly sequences is immense, meaning that despite the use of an anthropocentric taxonomy in the form of MTM-1, the question arises as to whether the described approach is sufficient to ensure that system behaviour conforms to the expectations of the assembly cell user.

This question must be considered from different perspectives. If we look at the sequence of operators (from REACH to RELEASE) within a single cycle, the sequence conforms to expectations because it corresponds to the normal cycle of movements of the human hand/arm system. This does not apply, however, to the sequence of cycles, or assembly sequences. With SOAR these prove to be a statistical succession of possible sub-steps that is impossible for humans to understand because it happens with no recognisable heuristics.

From a technocentric point of view, the reference model provides valid results in the sense of a complete, error-free and expedient structure. From an anthropocentric perspective, however, the results appear to be inadequate since the high procedural variance means that they cannot be compatible with human expectations. Marshall's Schema Model (Marshall 2008) divides the types of knowledge for the human decision-taking process into four categories: (1) Identification knowledge (*What is happening at the moment?*); (2) Elaboration knowledge (*What has high priority and why?*); (3) Planning knowledge (*What has to be done and when?*); (4) Execution knowledge (*Who should do what?*). Regarding the discrepancy between the technocentric and anthropocentric approach, Mayer et al. (2009) conclude that the reference model significantly under-represents elaboration knowledge.

To examine the hypothesis that a stronger focus on elaboration knowledge will have a positive impact on the expectation conformity of the system behaviour, a series of experiments investigating assembly strategies was held under laboratory conditions and with 16 participants (13 male, 3 female). The participants had to assemble a complete assembly on the basis of a CAD drawing. To keep the results comparable with the assembly cell, limitations were imposed on the execution of the task. Participants were only permitted to use one hand for the assembly, and were not allowed to build sub-groups or pick up several bricks at once. The target object was a single-coloured pyramid of 30 LEGO bricks.

The analysis of the assembly strategies in the experiments produced three general rules:

- Rule 1: From the viewpoint of the participant, the first brick to be assembled is located in one of the left corner positions (87.5% of the sample cases).
- Rule 2: Preference is given to selecting bricks that can be positioned directly next to an adjacent brick during assembly (81%). This will be referred to as compliance with the adjacency relationship.
- Rule 3: The target object is built up in layers that lie parallel to the assembly surface (81%).

Validity of the Collected Data with Respect to the Defined Rules

A further series of experiments (ES2) was carried out to check the validity of the previously identified assembly rules. The series involved 25 people (14 male, 11 female) who are not involved in manual assembly during their day-to-day work. The subjects are therefore not classified as experienced workers in the sense of the MTM taxonomy. None of the participants had taken part in the first series (ES1). The average age was 26.9 years (SD = 3.4). For ES2 the task from ES1 was expanded to require the subjects to assemble ten identical pyramids in succession in a timely manner. This meant that, in spite of the laboratory conditions, the subjects gained a degree of experience that is quite comparable with that of small-series production. The reason for the expansion of the task was to avoid "unusual" methods of working due to the simple task. Subjects had to signal the start and end of each assembly sequence by double-clicking a pushbutton switch installed in the assembly area. The assembly process was subject to the same limitations as in ES1—one-handed assembly and no building sub-groups or gripping several bricks at once.

If the rules identified from ES1 are also applicable to ES2, at least equal relative frequency f_i for each individual rule should be recognisable in the empirical data.

Table 6.9 Results of the χ^2 goodness-of-fit test

	MV ES2 (EV ES1)	*df*	χ^2	*p*
ES2, Rule 1	80.4% (87.5%)	1	11.52	0.00
ES2, Rule 2	91.2% (81%)	1	16.25	0.00
ES2, Rule 3	97.2% (81%)	1	41.75	0.00

$\alpha = 0.05$; *MV* Mean value; *EV* Expected value; *ES* Experiment series

On the basis of this assumption, the following hypotheses can be formulated for the statistical test:

- H_1: The relative frequency of the position of the first brick in ES2 (f_{h2_Rule1}) is higher than the frequency in ES1 (f_{h1_Rule1}), or: H_{01}: $f_{h2\_Rule1} = f_{h1\_Rule1}$.
- H_2: In ES2, assembly that conforms to the adjacency relationship (f_{h2_Rule2}) occurs with a higher relative frequency than in ES1 (f_{h1_Rule2}), or: H_{02}: $f_{h2\_Rule2} = f_{h1\_Rule2}$.
- H_3: Assembly in layers in ES2 (f_{h2_Rule3}) occurs with a higher relative frequency than in ES1 (f_{h1_Rule3}), or: H_{03}: $f_{h2\_Rule3} = f_{h1\_Rule3}$.

To verify the null hypotheses, the χ^2 goodness-of-fit test is applied with a significance level of $\alpha = 0.05$.

The results of the χ^2 test for H_{01} (position of the first brick in left-hand corners) are shown in Table 6.9. This clearly shows that H_{01} has to be rejected. With a relative frequency of 80.4%, the observed distribution deviates significantly from that observed in ES1.

The results of the χ^2 test for H_{02} (taking account of adjacency relationships during assembly) are also shown in Table 6.9. The null hypothesis must also be rejected in this case. However, because the relative frequency of 91.2% is higher than the expected value from ES1 (81%), it can be said that the rule is followed more strictly than expected.

Finally, the results of the χ^2 test for H_{03} (assembly in layers) are also shown in Table 6.9. This null hypothesis must also be rejected. As the observed relative frequency is 97.2% (expected value from ES1: 81%), it is clear that this rule is also followed more strictly than expected.

Influence of the Rules on Prediction Quality

As already shown, the null hypotheses regarding the rules identified in ES1 had to be rejected. The results of the χ^2 goodness-of-fit test for the rules relating to compliance with adjacency relationships and to assembly in layers show that these are followed more strictly than expected. The only rule that could not be applied to the results of ES2 was the one relating to the position of the first brick. Nevertheless, in view of the relatively high satisfaction of the rule (80.4%), it will continue to be considered in the further course of the study.

Table 6.10 Overview of the simulation models compared in the study

	MTM-1 rules	Rule 1	Rule 2	Rule 3
Model 1	X			
Model 2	X	X		
Model 3	X		X	
Model 4	X			X
Model 5	X	X	X	
Model 6	X	X		X
Model 7	X		X	X
Model 8	X	X	X	X

Finally, we also wanted to investigate how the identified rules—individually and in combination—influence the CCU's prediction accuracy regarding assembly steps preferred by humans, and the generalisability of assembly steps carried out by humans. To do so, independent sets of rules, which consist of the rules of the reference model and the respective auxiliary rule, are transferred to the CCU and simulated several times in a simulation environment developed especially for the experiments. The resulting data give an indication of how each set of rules influences the prediction accuracy of the simulation, i.e. its ability to predict the next step in the human assembly activity, and provide information on the generalisability of the set of rules with respect to ES2.

The additional cognitive simulation models were systematically expanded to include the rules corresponding to the heuristics. An overview of the simulation models is shown in Table 6.10.

For space reasons, in what follows only one dependent variable is used for evaluating the simulation models. The variable is derived from the criteria provided in Langley et al. (2009) for evaluating cognitive architectures. This criterion is based on the "optimality" criterion and represents the simulation model's prediction quality regarding the assembly activity. The prediction quality of an observed model is defined as the probability of the simulation model positioning a given brick in conformity with the human action during simulated assembly. In other words, assuming a given state x_{i-1} and the stored process knowledge for reaching the next state X_i, the probability $p(x_i|x_{i-1})$ that this particular state will be reached is evaluated. This observation is followed step-by-step for the assembly sequence until the target object is fully assembled.

As the simulated assembly is a Markov process, it is admissible to factor both the overall probability into the transition probabilities $p(x_i|x_{i-1})(2 \leq i \leq 30)$ described above, and the initial probability $p(x_i)$, where x_1 represents the initial state. The conditional probability of a sequence P_s is calculated as follows:

$$P_s = \prod_{i=2}^{30} p(x_i|x_{i-1}) \cdot p(x_1) \tag{6.16}$$

The logarithmic probability was calculated to simplify the interpretation of the resulting data. The prediction quality is therefore operationalised using "logarithmic conditional probability” (LCP):

$$LCP = \sum_{i=2}^{30} \log_{10} p(x_i | x_{i-1}) + \log_{10} p(x_1) \tag{6.17}$$

In Eq. 6.17, $p(x_i | x_{i-1})$ is the conditional probability that the observed simulation model assigns to a brick that was originally positioned by the human. Thus, the LCP values vary between 0 (perfect prediction) and $-\infty$ (absolutely wrong prediction or behaviour that the simulation model cannot validly reproduce).

To examine the prediction quality of the different cognitive simulation models (see Table 6.10), the index of the simulation models is regarded as an independent variable. Expanding the reference model of the cognitive simulation (Model 1 in Table 6.10) is expected to increase the prediction quality when additional empirically identified rules are added to the knowledge base. The following null hypothesis is formulated on the basis of this expectation:

- H_{04}: The prediction quality of the cognitive simulation models exhibits no significant differences.

The Kruskal-Wallis test was used with a significance level of $\alpha = 0.05$ to test for differences in the prediction quality. This test can be regarded as a non-parametric form of the one-factor analysis of variance (ANOVA). This is necessary because the Lilliefors test for normal distribution rejects the normal distribution for all LCP values ($p < 0.01$ in each case). A post-hoc test with adaptations for multiple comparisons according to Bonferroni was also carried out.

Results

Table 6.11 shows the mean values for the dependent variable LCPs from Eq. 6. 17, which describes the prediction quality for the 250 assembly sequences from the second empirical study.

On the basis of the simulation data underlying Table 6.11, it is clear that there is a significant difference ($p = 0.00$) in the *LCP* values. H_{04} must therefore be rejected. Figure 6.125 shows the simulation data as box plots of the observed cognitive simulation models. A higher LCP value means higher prediction quality with respect to human behaviour.

Table 6.11 Mean values for the logarithmic conditional probability (LCP) for the simulation models examined

Simulation model	LCP
Model 1	– 24.595
Model 2	– 24.422
Model 3	– 20.579
Model 4	– 20.279
Model 5	– 20.430
Model 6	– 20.192
Model 7	– 15.848
Model 8	– 15.671

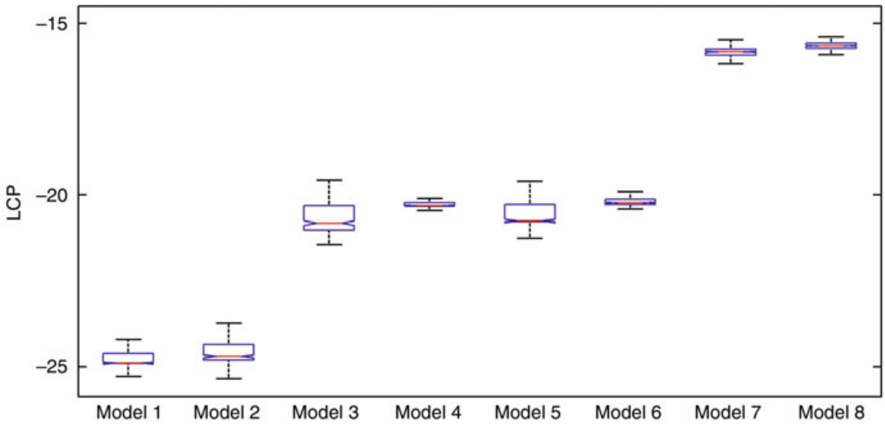


Fig. 6.125 Box plots of the logarithmic conditional probability (LCP) calculated for the simulation models examined

Table 6.12 Results of the multiple comparisons of the LCP values for the cognitive simulation models observed

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Model 1	-/-		X	X	X	X	X	X
Model 2		-/-	X	X	X	X	X	X
Model 3	X	X	-/-	X		X	X	X
Model 4	X	X	X	-/-	X		X	X
Model 5	X	X		X	-/-	X	X	X
Model 6	X	X	X		X	-/-	X	X
Model 7	X	X	X	X	X	X	-/-	
Model 8	X	X	X	X	X	X		-/-

Post-hoc pair comparisons were carried out to determine the differences between the cognitive simulation models. Table 6.12 shows the significant differences ($\alpha = 0.05$) as a cross-reference table, where X represents a significant difference ($\alpha = 0.05$).

If we compare all the cognitive simulation models, we can divide the models into three groups that are significantly different in terms of prediction quality. The first group has the poorest prediction quality and comprises Model 1 (the reference model for the cognitive simulation) and Model 2. Models 3, 4, 5 and 6 form the group with average prediction quality. Models 7 and 8 have the highest prediction quality.

The following conclusions can be drawn from the test results:

- Rule 1 has no significant effect when it is added to a cognitive simulation model.
- Rules 2 and 3 significantly influence the prediction quality of the cognitive simulation models, but they show significant differences in a direct comparison. Adding Rule 2 can improve the LCP value by between 16.32 and 22.39%. Rule 3 results in improvements ranging from 17.32 to 23.29%.

- The highest prediction quality is obtained when Rules 2 and 3 are combined. This increases quality by between 35.56 and 35.83%.

6.6.4.5 Human-Machine Interaction

As explained in Sect. 6.6.3, it is important to pay particular attention to the ergonomic design of the human-machine interface. Because the CCU can independently solve a certain class of rule-based production tasks, the human operator's duties mainly lie in cognitively demanding tasks. These include defining the production task, drafting rough process plans, defining the initial and boundary conditions and, in particular, monitoring the system state during the process and taking expedient action if disturbances or production errors arise. If the robot makes an assembly error, for example, the human operator must be able to quickly and efficiently intervene to enable system recovery. To achieve this, the sub-project designed, developed and evaluated a visualisation system that displays assembly information directly in the operator's field of view ("augmented reality", Ong et al. 2008).

A laboratory study was carried out to examine this visualisation system from the point of view of software ergonomics and to analyse how the relevant information should be displayed for the human operator. The task used in the experiment assumes that a robot has implemented the scheduling of an assembly task previously defined by the operator. During processing, however, a fault has occurred that the cognitively automated system cannot identify and remedy independently. Possible causes of assembly errors can be "noisy" sensor data regarding component detection, and flaws in the fed-in components themselves. Two separate series of experiments were performed to investigate the "optimum" ergonomic display of the assembly information with respect to the time necessary for fault detection and the corresponding accuracy. Two different displays (a head-mounted display—HMD—for visualisation directly in the field of view, and a table-mounted display—TMD—for visualisation in the workspace) and various types of visualisation were compared.

Experiment Scenario and Implementation

Different modes of visual representation (and interaction) were drafted and developed for the ergonomic system design and evaluation. These modes are based on known methods for designing written manufacturing instructions and are derived from guidelines for technical writing (see Alred et al. 2003). Exploded views and step-by-step instructions were used for ergonomically visualising assembly information in the field of view (Odenthal et al. 2008). To display assembly information in high quality and to permit quick and precise error detection with minimal mental effort, the assembly objects were accurately modelled as 3D objects and their colour and contrast were adapted to provide an ergonomic display on the HMD. LEGO bricks, which are easy to describe, were selected as assembly objects (see Sect. 6.6.1).

Table 6.13 Technical data of the HMD and TMD

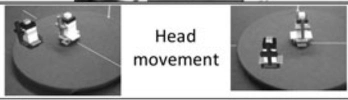
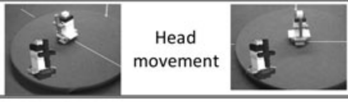
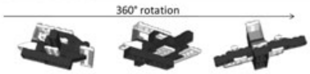
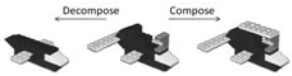
	HMD	TMD
Resolution	1,280 × 1,024	1,280 × 1,024
Image refresh rate (Hz)	60	75
Monitor size	–	17″
See-through transmission	40%	–
Weight (kg)	1.3	7
Brightness	max. 102.79 Cd/m^2	Typical 230 Cd/m^2; min. 197 Cd/m^2
Monocular FoV	60° diagonal/100%	–
Technology	LCOS	TFT
Manufacturer	NVIS	ELO

Experiment Design

The experiment design distinguishes between four factors. The first is “display type” (DT). The three other factors represent the different modes of presentation of the synthetic assembly information. These factors are: augmented vision mode (AVM); a priori presentation of the target state of the fully assembled assembly (APP); the mode for interactive decomposition and composition of the assembly during the error detection phase (DCM). These factors with the corresponding factor levels and the associated experimental conditions are explained below. Table 6.14 provides an overview of the factors and factor levels.

- DT: The visualisation systems were designed on the basis of two different display technologies: (1) A high-resolution HMD based on liquid-crystal-on-silicon technology (LCoS) with two half-silvered mirrors in front of the user’s left and right eye for stereoscopic display; (2) A high-resolution TMD using TFT technology, as is common in German industry. The TMD was selected and adjusted so that it corresponded as closely as possible to the specifications for the HMD. The light intensity was set to 500 lx at the point of the real assembly object in the working area. The technical data of the two displays are given in Table 6.13.
- AVM: The position and orientation of the HMD can be measured in real-time using an optical infrared tracking system. This makes it possible to calculate the user’s viewing direction and adapt the presentation of the virtual information in the field of view. Furthermore, the system can determine the position and orientation of the real assembly object in front of the user. This means that virtual information can be precisely superimposed onto the real object when the HMD is in use. No comparable visual superimposition is possible when using the TMD since it does not measure the position and orientation of the user’s head. Only the orientation (rotation of the real model around the vertical axis) was measured in the real model and adapted in the virtual model (rotation of the virtual image by 20° around the vertical axis). Because generating a “perspective view” is technically complex due to the tracking system, a company can only justify acquiring this if this presentation mode significantly improves performance or reliability. For this reason, a simple static information display was also investigated. The two factors are therefore “perspective view” and “static view”. In the perspective view, the

Table 6.14 The different factors and factor levels

Factors	Characteristic	Examples
DT for augmented display	HMD	
	TMD	
AVM	Perspective view	
	Static view	
APP	Rotation	
	Assembly	
DCM	Step-by-step	
	Exploded view	

display of the virtual assembly is adapted to the position and orientation of the real object and, when using the HMD, appears on the left at a distance of 4 cm from the real object. The position of the TMD was selected such that, as with the HMD, the virtual image appears to the left of the real model. The monitor was behind the turntable. This was intended to create as similar angle conditions as possible. In the static view, the virtual assembly is displayed in a fixed position relative to the monitor coordinates of the HMD or TMD. In the experiment the virtual object was displayed in the HMD, tilted at an angle of 20° to the vertical axis and in the third quadrant of the binocular display. In the TMD, the virtual object was displayed tilted at an angle of 20° to the vertical axis and in the centre of the monitor.

- APP: Before the actual error detection task, the subject was shown an a priori display of the target state (assembly). In this case a distinction was made between two factor levels: (1) Rotation: the complete assembly rotates 360° once, tilted at an angle of 20° to the vertical axis and at an angular velocity of 3.2 s.; (2)

Fig. 6.126 Experiment configuration with turntable and HMD (*left*), experiment configuration with monitor (*right*), keypad for manipulating the virtual object (*top centre*), participant wearing the HMD (*bottom centre*)

Assembly: the virtual model is assembled virtually in steps with a cycle time of 1.2 s. per sub-element.

- DCM: Two factors were analysed with respect to the mode for interactive manipulation of the LEGO model during the error detection process: (1) Step-by-step: the subject can independently compose and decompose the virtual model in steps, using two buttons on a keypad (see Fig. 6.126); (2) Exploded view: the subject can interactively explode or implode the virtual model using another button on the keypad.

Table 6.15 also shows the types of errors that are generally relevant to error detection. Because detecting a colour error is simpler than detecting a position error or an error in shape or number, the experiment does not investigate colour errors. Eight different tasks with similar degrees of difficulty were developed to analyse the remaining error types. Two tasks contained the error "Position", three tasks contained the error "Type/shape", and three tasks contained the error "Number".

Table 6.15 Possible error types

	Colour (not investigated)	Position (investigated)	Shape (investigated)	Number (investigated)	
Target state					
Actual state					

Experiment Configuration

Figure 6.126 shows the main components of the visualisation systems that were used for the experiment. The LEGO model with the assembly error that had to be detected and identified was placed on the turntable, which was located on a conventional assembly bench in the participants' central field of view. Each person sat on a chair in a comfortable upright position during the experiment. The position and orientation of the HMD and the turntable were recorded with the smARTtrack real-time tracking system (made by ART GmbH). The participants used a coloured pen, which was located behind a transparent screen, to mark the incorrectly assembled LEGO brick. As soon as the person detected the error, he/she picked up the pen, which triggered a switch linked to the screen and recorded the time of the action. This point in time was taken as the moment of error detection for all participants. The keys for manipulating the virtual image were attached to a freely configurable keypad with the following key assignment: exploded view (left), decomposition/composition (top and bottom centre), repetition of the a priori presentation (right). The key at the top right has no function within this series of experiments.

Experiment Procedure

The laboratory studies for the HMD and TMD were performed separately, each with two groups of 24 participants. The same assemblies with identical errors were used in both studies. A full-factorial experiment design with measurement repetitions using three within-subject factors (AVM, APP, DCM) and one between-subject factor (DT) was selected. The task for the participants consisted of comparing the state of a real assembly object with the virtual representation on the screen (HMD, TMD) for possible errors. The experiment procedure was split into two main phases:

1. Pre-tests and training under experiment conditions

At the start of the study, general user data (age, profession, previous experience, etc.) were gathered using a questionnaire developed specifically for the study. Then the participants' visual acuity was recorded (tested in accordance with DIN 58220), as was their stereopsis and their colour vision (using the Ishiara colour test). Since wearing a HMD can quickly lead to visual fatigue and thus significantly influence factors like human performance and stress (Pfendler and Schlick 2007), participants' visual fatigue was recorded, based on Bangor (2000), using a questionnaire before and after performance of the task. Next, participants spent ten minutes practicing with the augmented vision system under the conditions that would be used in the experiment.

2. Data collection

Each participant performed the following sequence eight times:

- Starting a run using APP of the target state of the LEGO model. The virtual sequence was presented in the participant's field of view, without him/her being able to see the real object.
- At the end of the initial presentation, the test supervisor fastened the real assembly object to a turntable. Participants could use the keys on the keypad to manipulate the virtual object. They could call up the a priori presentation at any time, but could not cancel the displayed assembly sequence. The participants' task was to compare the real object (actual state) with the virtual object (target state) for differences, without knowing whether and/or how many possible assembly errors the component contained. If they identified differences, they had to mark them with a coloured pen. Each component had one assembly error.
- Each participant then completed the questionnaire on visual fatigue.

The total experiment, including the pre-tests, lasted roughly two hours for each participant.

Participants

A total of 48 people (16 female and 32 male) took part in the laboratory study. All participants satisfied the criteria of normal vision or corrected vision (visual acuity 0.8), stereopsis and colour vision. Apart from these physiological requirements, the groups were formed according to the following criteria: homogeneous age, comparable spatial perception (cube test, Liepmann et al. 2007), comparable experience with augmented or virtual reality (AR/VR) and comparable experience of assembly.

- Group 1—HMD. The participants were between 19 and 36-years-old (MV: 26.8 years; SD: 4.4 years). 95.8% used a computer every day. 58% stated that they had little or no experience with VR. 37.5% had experience with 3D computer games (on average 4 hours per week playing time). The average experience in LEGO assembly had a value of 3.0– on a scale of 0 (low) to 5 (high).
- Group 2—TMD. The participants were between 20 and 40-years-old (MV: 26.0 years; SD: 4.5). All participants used a computer every day. 64.6% stated that they had little or no experience with virtual or augmented reality systems. 50% had experience with 3D computer games (on average 4.7 h per week playing time). The average experience in LEGO assembly had a value of 2.9—on a scale of 0 (low) to 5 (high).

Dependent and Independent Variables

In accordance with the experiment plan, a distinction was made between four independent variables (see also overview in Table 6.14):

- Display type (head-mounted or table-mounted)
- Augmented vision mode (perspective or static view)

- A priori presentation of the target state of the complete assembly model (rotation or assembly)
- Decomposition/composition mode (step-by-step or exploded view)

The experiment measured the following dependent variables:

- Detection time: This was the time between the appearance of the real LEGO model and the detection of the difference by the participant (max. 15 min.). The experiment configuration measured the start and end point of this period.
- Error detection: Different cases could occur which could be represented by separate variables: (a) The participant detects the difference (error correctly detected); (b) The participant picks out a brick which is no different from the virtual (target) state (error wrongly detected); (c) The participant does not find the difference (error not detected).
- Visual fatigue: The study assumed that this sets in quite rapidly, particularly when using a HMD. This subjective variable was therefore also included in the study.

Null Hypotheses and Statistical Analysis

The following null hypotheses were formulated:

- The display type has no significant influence on detection time (H_{01}) or error detection (H_{02}).
- The augmented vision mode (H_{03}), the a priori presentation of the synthetic assembly information (H_{04}), and the decomposition/composition mode of the virtual model (H_{05}) have no significant influence on error detection time based on display type.
- The augmented vision mode (H_{06}), the a priori presentation of the synthetic assembly information (H_{07}), and the decomposition/composition mode of the virtual model (H_{08}) have no significant influence on the given cases of error detection based on display type.

On the basis of the data collected, inferential-statistical analyses were carried out using the program "Statistical Package for Social Science" (SPSS Version 17). First, the two data sets (Group 1—HMD; Group 2—TMD) were compared with respect to the dependent variables. A four-factorial repeated measures ANOVA was calculated with for the test of H_{01} (three within-subject factors: AVM, APP, DCM; one between-subject factor: DT). The groups were then analysed separately to identify any differences with respect to the different visualisation and interaction modes. A three-factorial repeated measures ANOVA was carried out for the test of hypotheses H_{03}, H_{04} and H_{05}. The detection time data were log-transformed before the ANOVAs were carried out (see Field 2005) to satisfy the qualitative requirements. The significance level was set at $\alpha = 0.05$. A Kolmogorov-Smirnov test was performed to examine the log-transformed data for normal distribution. Chi-square tests were carried out to verify the hypotheses H_{02}, H_{06}, H_{07} and H_{08} using the nominally scaled data from error detection. The significance level was again set at $\alpha = 0.05$.

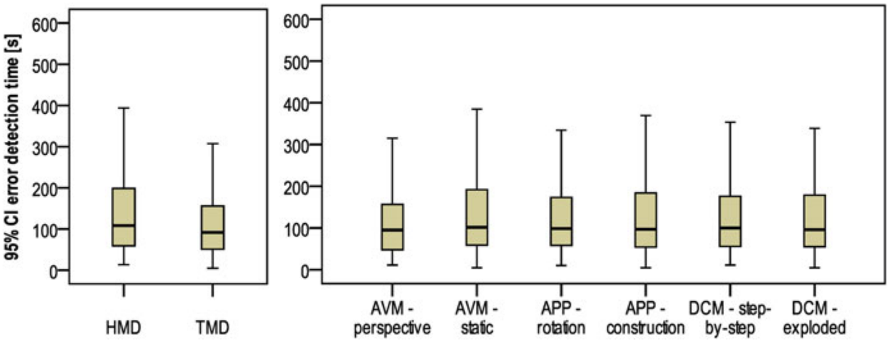


Fig. 6.127 *Left:* Error detection time for all participants, depending on the between-subject factor DT; *Right:* Error detection time for all participants, depending on the within-subject factors AVM, APP and DCM

Results and Interpretation

The log-transformed time data showed no significant deviation from the normal distribution. The mean values of the error detection time under the different experimental conditions are shown in Figs. 6.127 and 6.128 in the form of box plots.

Error-Detection-Time Results for All Participants: On average, the error detection time using the TMD was 27.64% shorter than with the HMD. However, this difference is not statistically significant ($F_{(1.45)} = 3.001$, $p = 0.090$).

A comparison of the perspective and static views in AVM showed that error detection times in the perspective view were on average 23.45% shorter than with the static view. This difference is statistically significant ($F_{(1.45)} = 8.854$, $p = 0.005$). In the case of the a priori presentation of the assembly information, the rotation condition led to an average error detection time that was 9.4% shorter than under the assembly condition. However, this difference is not significant ($F_{(1,45)} = 0.043$, $p = 0.837$).

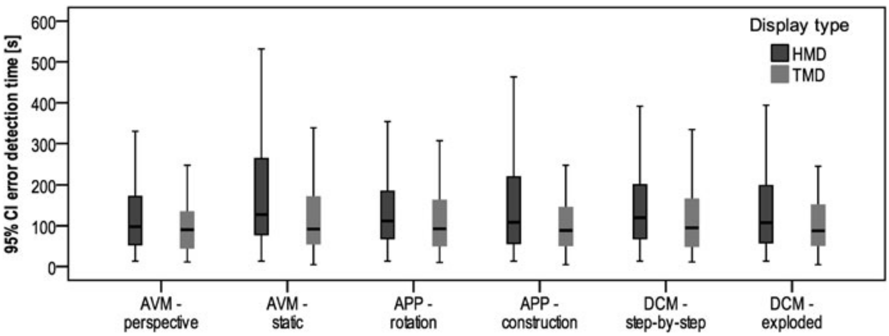


Fig. 6.128 Error detection time divided between the groups in the experiment (HMD, TMD) for the different conditions (AVM, APP, DCM)

When the participants worked with the exploded view, the average error detection time was only 2.2% shorter than with step-by-step decomposition/composition. Again, this is not a significant difference ($F_{(1.45)} = 0.103, p = 0.749$). No statistically significant interactions were identified.

Error-Detection-Time Results for Group 1 (HMD): For the perspective view in AVM, the study showed that error detection times were 29.61% shorter than with the static view. This difference is statistically significant ($F_{(1.22)} = 8.088, p = 0.009$). The rotation condition of the APP mode resulted in error detection times that were on average 18.19% shorter than the alternative assembly condition, but this difference is not significant ($F_{(1.22)} = 0.004, p = 0.929$). The average error detection times for both conditions in the decomposition/composition mode barely differed. With the exploded view, error detection times were an average of 2.78% shorter than with step-by-step decomposition/composition. This is not a significant difference ($F_{(1.22)} = 0.001, p = 0.971$). No statistically significant interactions were identified.

Error-Detection-Time Results for Group 2 (TMD): For the perspective view in AVM, error detection times were 13.07% shorter than with the static view. This difference is not statistically significant ($F_{(1.23)} = 1.526$, $p = 0.229$). The assembly condition in APP resulted in a 2.56% reduction in the average error detection time compared with the rotation condition. This difference is also not statistically significant ($F_{(1.23)} = 0.335, p = 0.568$). The average error detection times for both conditions in DCM barely differed. With the exploded view, error detection times were an average of 2.94% lower than with step-by-step decomposition/composition. This is not a significant difference ($F_{(1.23)} = 0.148, p = 0.704$). No statistically significant interactions were identified.

Error Detection Results for All Participants: The number of correctly detected errors was 36% higher with the HMD than with the TMD. The display type significantly influences all categories of error detection (correctly detected: $p = 0.019$; wrongly detected: $p = 0.023$; not detected: $p = 0.044$). With regard to AVM, the perspective view increased the number of correctly detected errors by an average of 18.5% compared to the static view. Compared with rotation, step-by-step assembly of the LEGO model with the APP mode enabled the user to correctly detect 14.5% more errors. Step-by-step decomposition/composition resulted in 10.7% more errors being correctly detected than with the exploded view. However, these differences are not significant. Error detection in the other two categories (error not detected or wrongly detected) is not significantly influenced by the levels of the independent variables.

Error Detection Results for Group 1 (HMD): A comparison of the two AVMs shows that, on average, the perspective view resulted in a 19% higher error detection probability (error correctly detected) than the static vision mode. With APP, a rotating LEGO model enabled the user to correctly detect 16% more errors than with step-by-step assembly. With regard to DCM, the difference in correct error detection between step-by-step decomposition/composition and the exploded view was far smaller (6%). The differences are not significant. Error detection in the other two categories (error not detected or wrongly detected) is not significantly influenced by the levels of the independent variables.

Error Detection Results for Group 2 (TMD): A comparison of the two AVMs shows that, on average, the perspective view resulted in a 17.4% higher error detection probability (error correctly detected) than the static view. With APP, step-by-step assembly of the LEGO model enabled the user to correctly detect 12.8% more errors than with a rotating model. With regard to DCM, correct error detection in step-by-step decomposition/composition was 17.4% higher than with the exploded view. The differences are not significant. Error detection in the other two categories (error not detected or wrongly detected) is not significantly influenced by the levels of the independent variables.

Visual Fatigue Questionnaire Results: As already mentioned, the participants' visual fatigue was recorded before the first and after the last cycle. A difference of 1 represented a subjectively perceived difference of 10% in visual fatigue. The largest average increase in visual fatigue for Group 1 (HMD) was 1.1 for the item "headache", followed by 0.65 for the item "mental fatigue". The largest average increase in Group 2 (TMD) was 0.37 for the item "mental fatigue". The other differences are smaller than 5% (0.5) and can therefore be ignored. The visual fatigue recorded was lower than expected, which indicates a good ergonomic design of the augmented vision systems. It should be pointed out, however, that the HMD was only worn for between 0.5 and 13.6 min. at a time during the experiment. The error detection phase was always followed by a 2–3 min. recovery phase, during which participants filled out the questionnaires (not wearing the HMD). It is expected that a longer cycle time would significantly increase visual fatigue (Pfendler and Schlick 2007).

Discussion

Display Type The results of the series of experiments carried out here confirm the research results of Tang et al. (2004) and Meyer et al. (2005). Using a HMD led to a significantly higher degree of precision in error detection, but not to a significantly shorter detection time. The lower error detection rate when using the TMD can be attributed to participants having to frequently shift their attention between the component and corresponding virtual model, and to the high mental strain this brings about. Furthermore, participants had to adapt to different light intensities between the real model and the representation on the screen. The participants consequently tended to overlook the error and thus arrived more quickly at a decision as to whether or not an error existed.

Augmented Vision Mode Compared to the static view, AVM with perspective view resulted in a shorter average error detection time for both groups. The perspective view did, however, have less influence on the error detection time when using the TMD. The difference between the two modes was not significant here. The perspective view resulted on average in more correctly and less wrongly detected errors.

Nevertheless, these differences were not significant in either group, which means the data analysis provided no statistically clear proof of a speed/accuracy trade-off.

The difference between the perspective and the static views—irrespective of the display used—is attributable to higher mental strain in the static view. In this view, the user frequently has to mentally rotate the virtual model during the error detection phase, which is a strenuous and time-consuming central process. Since Shepard and Metzler (1971) published their classic work, we know that people's reaction times in comparing and deciding whether or not two items are identical are proportional to the rotation angle between the two representations. The study confirmed this relationship. Reaction times increase as the objects presented become more complex (Funke and Frensch 2006). The perspective view allows users to make a direct perceptive comparison between the target and the actual assembly state, without drawing on important mental resources for rotation or translation.

As already mentioned, in the perspectively modified mode when using the HMD, the presentation of the virtual object is tracked if the turntable is rotated (vertical axis) or if the user moves his/her head (other axes). By contrast, the TMD only rotates the virtual object (vertical axis) if the turntable is rotated. These different degrees of perspective adaptation are probably the reason for the differences in performance and reliability between the perspective view and the static view in Group 2 (TMD). However, the differences are not significant.

A priori Presentation With APP, a shorter error detection time in rotation mode was observed for both groups. Using the HMD resulted in a shorter average error detection time, but correct error detection was lower than with step-by-step assembly. In other words, a tendency towards a speed/accuracy trade-off was observed here. The average error detection time when using the TMD was practically identical ($\sim 2\%$) in both rotation and step-by-step assembly. However, step-by-step assembly resulted in fewer correctly detected errors. Unlike the rotation sequence, the assembly sequence of the assembly object allowed participants to create a precise mental model of the product structure and the assembly procedure. Rotating the virtual model, which avoided the strain of mental rotation as already mentioned, resulted in faster but less reliable error detection.

Decomposition/Composition Mode In DCM, the exploded view allowed the user to switch quickly between a fully assembled product structure and an exploded view of a product structure, and to create a precise representation in their visual-spatial memory. However, in the step-by-step mode—particularly the assembly (bottom-up)—it took a certain time before participants detected the error. Nevertheless, and contrary to the expectation, only a very small time advantage of $\sim 3\%$ for the exploded view was observed in both groups. This appears to be due to the effects of overlapping, which make the exact localisation of the error more difficult. Correct error detection in the exploded view was lower than with step-by-step decomposition/composition. Consequently, the rate of wrongly detected errors and errors not detected was higher, but these differences are not significant.

Based on the results of this study, the head-mounted stereoscopic vision system was further developed to support the user in remedying the assembly error once it has been identified. Interactive graphic decomposition is possible here (Odenthal et al. 2011). A further study is in preparation (parts have already been conducted) that will compare different modes of visual assistance in the development of a cooperative human-robot decomposition strategy with regard to performance, reliability and mental strain.

6.6.5 Industrial Relevance

The concept of generic strategies first introduced by Porter in 1980 (Porter 2004) provides three strategies that a company can use to achieve a competitive advantage over its rivals. The cost-leadership strategy aims to gain a competitive advantage by keeping costs down. To do so, companies can adopt a number of different approaches (see Mintzberg et al. 2004), which may have opposing characteristics—e.g. economies of scale and scope. The differentiation strategy is about developing a unique selling point using factors like image, service or design. With the focus strategy, a company focuses on a very specific customer group or market segment, although it might also focus on differentiation or cost-leadership. A company in a high-wage country produces new products domestically and pursues a strategy of differentiation. Later in the product life cycle, however, it is very likely to move production to a low-wage country to pursue the cost-leadership strategy because other providers are flooding the market to benefit from the economic success of the product. The same applies to the focus strategy, which must undergo the same change.

The concepts and technologies for cognitive automation described in this paper make it possible—by taking targeted action on the polylemma of production—to adopt a position between Porter's strategies. A Business and Technology Case provided the opportunity to put the results into practice for the first time, as a prototype application. For this purpose concepts already tested with the demonstrator cell were simulated, in collaboration with connection technology manufacturer Phoenix Contact, in a real scenario involving switch-cabinet production.

In addition to final assembly in its mass-production-style manufacturing system, Phoenix Contact operates a customer-driven switch-cabinet assembly system. This involves assembling switch-cabinet components (control units, terminals, etc.) in configurations pre-planned by the customer, mounting them onto top-hat sections and passing them on as completed modules to switch-cabinet production (module assembly). Figure 6.129 shows an example of this system.

Fig. 6.129 Assembling switch-cabinet components on a top-hat section

Individual components must be mounted with additional plug combinations. The different mounting configurations planned are available as CAD data, as required by the CCU. The relevant mounting requirements are usually printed out, and employees then manually assemble the components on the top-hat sections. Next, even today, an image-processing-supported comparison is made between the mounting requirements and what has actually been mounted on the top-hat section.

Because the target is already available in electronic format, cognitively automating the process could bring considerable economic benefits in the future. The main challenges this involves are as follows:

- Building a continuous chain of information from the CAD system to the assembly system
- Making available robust and productive system components and joining processes for module assembly (handling devices, joining processes and aids for plugging, clamping and, if necessary, screwing)
- Applying logistical components and concepts for fitting the switch-cabinet components at the right time and in the right quantity
- Developing control concepts for establishing the ideal assembly sequence and for carrying out the assembly task itself
- Ensuring that sensors monitor the actual mounting situation and that human operators take corrective action in the case of errors

The switch-cabinet scenario bears strong similarities to the demonstrator-cell scenario. The following correlations are relevant to the transferability of the developed concepts:

- Defined workpieces: The LEGO scenario involves fitting a large number of building bricks from the whole range. The product specification means that all characteristics relevant to assembly—such as dimensions, tolerances, and colour—are known. Unfamiliar bricks or bricks that do not fit the boundary conditions for assembly are not fed in. The modules in the switch-cabinet assembly scenario correspond to the LEGO characteristics described. The modules are also precisely specified and are only fed in after quality control. Just like any LEGO bricks fed in by accident, any modules erroneously fed in can be separated out or, if they are to be assembled later or at another stage, they can be stored or fitted.
- Flexible handling technology: The flexible gripper hand installed in the demonstrator cell was chosen for two reasons: it could grasp a LEGO brick using a variety of grips depending on the assembly situation, and if the assembly scenario changed in any way, it could grip and join other workpieces with unfamiliar shapes. The gripper hand's seven axes and its push-button sensors mean that it could also assemble the switch-cabinet modules.
- Production logistics: As in the previously described scenario, the components in switch-cabinet production must be fed in individually. The demonstrator cell uses a robot to do this, along with a group of conveyor belts with photoelectric sensors and switches. A similar logistics system would also be suitable for feeding in modules for switch-cabinet assembly.

- Individual configuration according to objective: A particular challenge for the demonstrator cell involves planning and executing assembly processes that are unknown at the time of development. The definition of the target—i.e. the description of the assembly to be assembled—is provided for each assembly in the form of CAD data, which are used to establish the necessary assembly processes. Phoenix Contact produces a description of the components to be assembled for each customer-specific switch-cabinet. In the automated switch-cabinet assembly, the assembly system is responsible for feeding in the required components and for the assembly sequence. The case under discussion also showed that the system can be controlled using a cognitive controller.

6.6.6 Future Research Topics

Future research work should aim to achieve higher degrees and levels of cognition in production systems. In doing so, it should pay particular attention to topics relating to cognitive functions and technologies for fast set-up, fast start-up, flexible scalability of throughput, and to the possibility of cooperative cellular systems. Researchers must also address the question of how to construct cognitively automated production so that it is as versatile as possible and can be operated efficiently, safely and sustainably, even in the rapidly changing environment of a high-wage country like Germany. Within this context, it is especially important to integrate humans and their superior cognitive, perceptive and sensorimotor skills.

Industrial production often uses production lines with sequentially linked stations. Each station carries out the production-process steps assigned to it and then passes the (sub)product on to the next station. The only way to compensate for failures and delays at a station is by using buffers. In extreme cases, failures and delays can bring the entire system to a standstill. Using cooperative, cognitively automated cells and control structures can make the production system more error-tolerant and flexible. A shortage of materials, a broken tool, or delays would only minimally affect overall production, because the system would be able to dynamically adapt the production flow and have other CCUs take over certain production steps. Figure 6.130 shows a diagram of how this kind of process chain made up of cognitively automated cells might look.

The configuration of the process chain pictured here is based on the chaku-chaku principle (literally load-load in Japanese), also known as the “one-piece-flow system”. In its original form, this system involves all stations producing more or less autonomously, and a human operator simply transports the parts from station to station. The principle can achieve a high level of flexibility when it comes to handling variants and fluctuations in production, and at the same time it can reduce processing times and space requirements. A classic chaku-chaku line is generally based on a high degree of automation and a linear-cyclical linkage, in which humans simply perform “residual tasks” that offer almost no scope for making decisions or taking

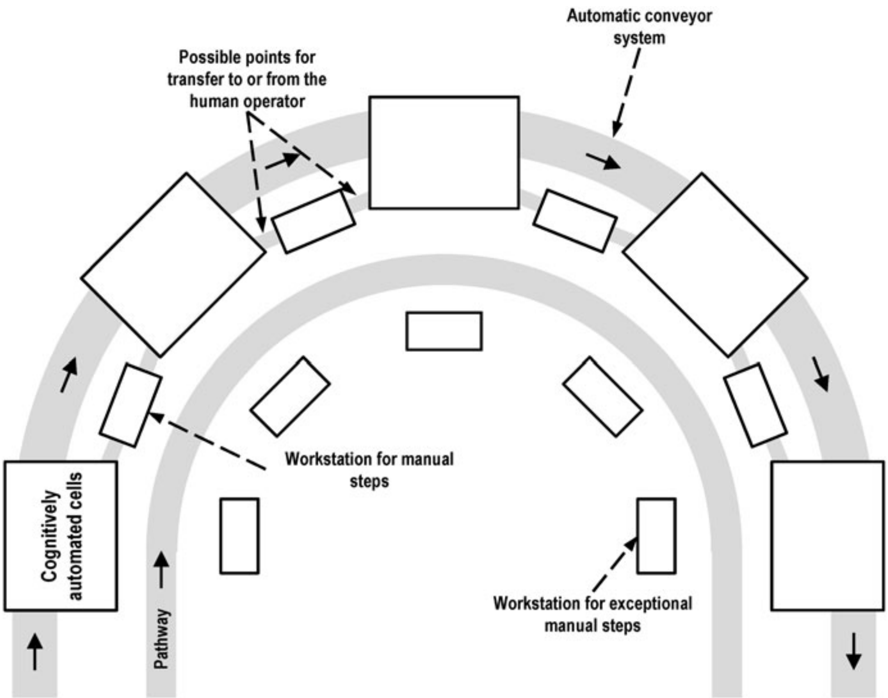


Fig. 6.130 Schematic construction of a cognitively automated process chain

action. Frieling and Sträter (2009) have investigated the productivity of such assembly systems and the health risks they pose to employees. Initial research findings from the samples they examined indicate that this kind of activity can have harmful effects on employees' mental workload (monotony and fatigue). This is presumably the result of highly repetitive assembly tasks and the lack of job rotation in the assembly concept (Enríquez Díaz et al. 2010)

To counteract the disadvantages of the chaku-chaku principle, future proposals should focus on parallelising and integrating tasks within the cells and on allowing humans to interact with cognitively automated systems in an ergonomic way. Looking at a single cognitive cell within the process chain, and based on Fig. 6.131, we can identify the following possibilities for processing a blank, a semi-finished product or a finished product:

- Fully-automated process: Feed-in, processing and feed-out happen automatically within the cell. Humans monitor and optimise the production process.
- Cooperative, partially automated process without removal: Feed-in and partial processing happen automatically. A human is required to manually perform some of the processing. To do so, the part in question is transported to the human operator's workspace, without leaving the secure, controlled area of the cell. While this is happening the cell can do nothing else. Once the human operator

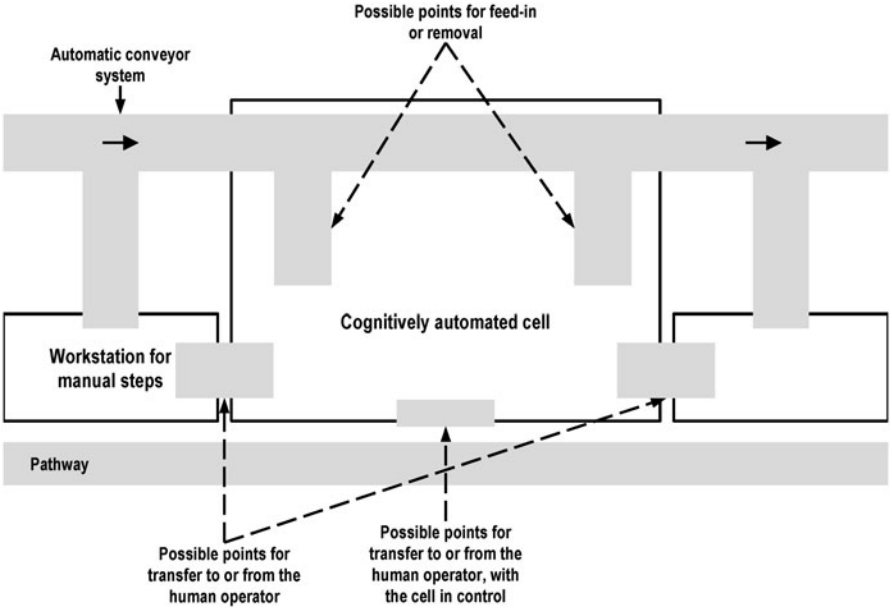


Fig. 6.131 Schematic construction of a cell in the cognitive process chain

has completed the manual work, the conveyor system takes the part through the remaining automated processing steps, the measuring and checking tasks, and on to feed-out.

- Cooperative, partially automated process with removal: Feed-in and part of the processing happen automatically. A human is required to manually perform some of the processing. To do so, the part in question is transported out of the cell to a manual workspace. The part is now outside the secure, controlled area of the cell, so the cell is free to carry out other tasks while the part is being manually processed. Once the human operator has completed the manual work, he/she feeds the part back into the cell, and the conveyor system takes it through the remaining automated processing steps and on to feed-out.

If the blank, semi-finished product or finished product does not require processing in a particular cell, the integrated automated conveyor system can make it skip that cell.

This concept requires research into the following:

- New concepts and technologies that allow cooperative CCUs to coordinate and communicate material requirements, process statuses and capacities.
- Cooperative planning algorithms that can dynamically respond to changes and adapt the behaviour of individual CCUs to keep the performance, reliability and safety of the overall system at an optimal level. Several concepts can achieve this: CCUs can coordinate directly with one another, or hierarchically structured

"meta CCUs" can perform higher-level planning tasks. The algorithms must be able to take into account limited availability of common resources, availability of materials, and the flow of products.
- This kind of system also makes it possible to integrate the specific skills of a human operator into the production process. Thus, using higher-level planning, work tasks that due to their complexity require a minimum level of experience can be assigned to the person best suited to the job. Furthermore, training and qualification processes can be carried out during periods when regular operations are slow (embedded training). This means that individualised production does not just apply to individualised products; it also refers to adapting production to the individual employees involved. Particularly in light of today's changing demographics, these kinds of systems could take physical impairments into account in production planning and execution and thereby offer staff support that is specifically targeted to their needs.

Given the many transfer points within the system, another field of research activities involves developing ergonomic designs for human-machine interaction. Because production is so varied and normally controlled by demand, one cannot assume that one fixed process model will emerge that humans can use for orientation. This underscores the necessity that human-machine interaction is safe and that it conforms to human expectations. Future research work should therefore investigate the following questions:

- To what extent can a cognitively ergonomic display of system status—e.g. a dynamic, multi-level flow diagram rather than the abstract, schematic displays that are common in today's control stations—increase system transparency, particularly given the fact that a human operator has to manage several cells and segments at once? Because the research focus in this case is not limited to intervening when errors arise, scientists need to develop additional concepts and methods for intuitively monitoring normal operation.
- How can targeted variations in a robot's motion sequences, in the sense of anthropomorphic kinematics and dynamics in the cell, inform the human operator—quickly, reliably and in a way that optimally responds to operational demands—that a transfer is coming up?
- Can anthropometric variables be integrated into overall system planning in such a way as to allow individual, ergonomic adaptation of things like transfer or processing points? How much flexibility do these points offer in terms of adapting to something like an unexpected staff absence?
- Can occupational safety aspects be taken into account during planning—e.g. by integrating an individual biomechanical model of the human operator—so that things like loads and torque are not set as standard, but are adapted to the specific human operator?

In conclusion, it is clear that a cognitive planning and control unit in the form of the detailed CCU concept described in this paper only provides a small example of what cognitively automated systems will be able to achieve in the field of production.

Looking at the process chain as a whole, it is already possible to use cognitive mechanisms to comprehensively optimise a system beyond the boundaries of existing tolerance areas. Using a targeted combination of components—known as cognitive tolerance matching (Schmitt et al. 2009b, see Sect. 6.2.4.3)—it is also possible to improve the quality of product. Taken together, the methods and systems developed so far offer companies in high wage countries in particular the chance to achieve considerable competitive advantages. By using cognitive mechanisms in automation, these companies can directly contribute to securing and expanding production locations in those countries.

6.7 Reconfigurable Assembly Systems for Handling Large Components

Burkhard Corves, Rainer Müller, Martin Esser, Mathias Hüsing, Markus Janßen, Martin Riedel and Matthias Vette

This chapter presents the results of the subproject "Reconfigurable self-optimising component handling", part of the Cluster of Excellence in the integrative cluster domain of "Self-optimising production systems". The Chair of Assembly Systems, Laboratory for Machine Tools and Production Engineering (WZL) is involved in the project as well as the Department of Mechanism Theory and Dynamics of Machines (IGM) of RWTH Aachen University.

6.7.1 Challenge

The project is motivated by fundamental changes in the operating framework for manufacturing companies in recent years. Reasons for increasing complexity and dynamics within these companies and in the industrial environment are not only increasing globalisation, but also the fast pace at which technology is developing and the altered resources situation (Möller 2008; Müller et al. 2009b). The consequences of this include the even greater shortening of product lifecycles, the persistent increase in the number of product variants and the constant pressure to reduce manufacturing costs (Lotter 2006). Assembly systems that are reconfigurable for the specific purpose are a highly promising approach in the conflict between achieving individualised production on the one hand and cost-reducing automation on the other.

Given this scenario, a current focus is the handling of large components that sometimes possess only little inherent rigidity. Such components are used in aerospace engineering for example, in shipbuilding and wind power installations. In aircraft construction in particular, we see increasingly large shell elements of CFRP materials being used for fuselages (Fig. 6.132) (Licha 2003).

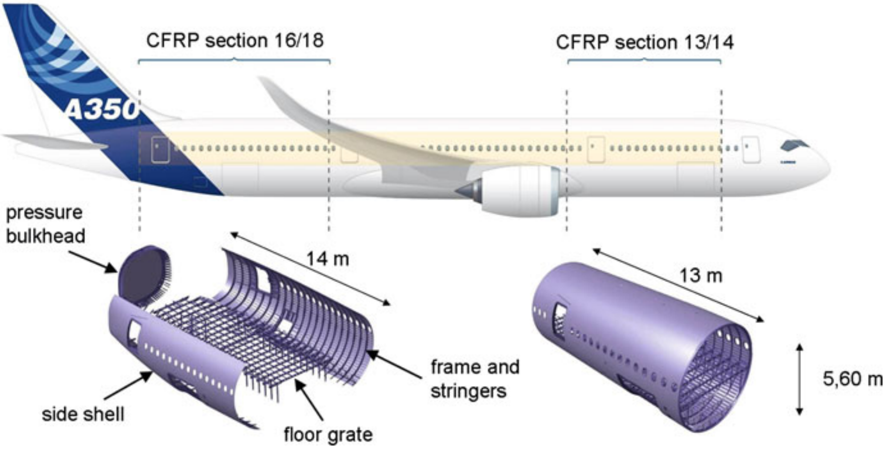


Fig. 6.132 Airbus A350 XWB sections 16/18 and 13/14. (Source: Premium AEROTEC GmbH 2011)

Especially for these components you have to consider that only limited forces may be applied to the component when it is handled. This also calls for high accuracy and an adequately large workspace. To satisfy these requirements, large jigs and fixtures are used for handling, specially matched to the particular component, and making a system highly inflexible and costly. A separate jig or fixture has to be constructed and provided in sufficient numbers for each component. The situation is complicated even more by the fact that the individual structural elements of an aircraft differ very much in shape, size and the position in which they are fitted, which will in most cases prohibit the use of a general-purpose jig or fixture.

The large variety of components and their application consequently necessitates an assembly system that can be expanded and reconfigured for the purpose. The assembly system must be scalable and universally applicable to respond to a changing economic and technical situation. To make automation economically attractive for assembling smaller production series too, it must be possible to ramp up a system again quickly after its reconfiguration to avoid the cost of long downtimes.

The approach taken here consists in composing an assembly platform (Fig. 6.133) to match the illustrated requirements. Different modules can be combined to implement the required assembly tasks and functions.

6.7.2 State of the Art

Industrial production shops are in many cases characterised by automated assembly processes. The degree of automation is relatively high in particular for small to medium-sized products in series production. However, this degree of automation becomes less as the size and complexity of assemblies and the scale of joining

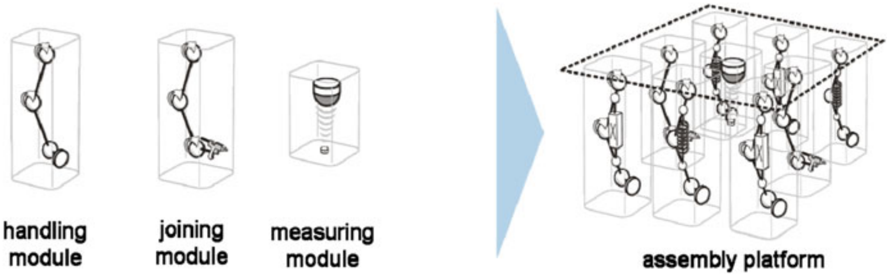


Fig. 6.133 Assembly platform

operations and processes increase. Large components in the aerospace industry for example, in shipbuilding and the construction of rail vehicles are for the most part assembled manually, supported by jigs and fixtures. The reasons for what seems to be a technical lag are to be found in the nature of the product to be assembled and high technical demands when it comes to accuracy and accessibility of the joints (Stepanek 2007).

Large components in particular frequently have little rigidity before they are assembled, so the force of gravity alone can cause inadmissible deformation, meaning that components must be picked up and supported at a number of points when handled. The task of the system includes joining components in addition to transporting and feeding them on a line. The joint tolerances may be just a few tenths of a millimetre while the components themselves are anything between ten and thirty meters (Wollnack and Stepanek 2004).

To complicate things more, the components are very different and are only manufactured in relatively small numbers, resulting in a whole number of high-investment stations and large space requirement when it comes to rigidly automated systems. This trend is especially noticeable in aircraft assembly where aircraft are sometimes manufactured for a number of decades using one assembly system. Here you generally find product-specific jigs and fixtures that not only ensure correct joining of the assembled objects by defined geometrical, functional elements but also support proper shaping of the components.

High investment costs and the long product life cycles of the assembly systems hinder the continuing development of new assembly technologies. In part, assembly systems are used for several decades. The introduction of new assembly technologies is frequently tied to the introduction of new products or variants, the result of which, in aircraft production, is long innovation cycles. But these, in most cases rigid systems are unsuitable for new market demands like individually matched and quickly available product variants. Instead a flexible system is needed that can respond to customer requirements without huge investment.

Motivated by the drawbacks of jig and fixture-oriented assembly of large components, new approaches are sought to make the assembly of large components more flexible and be able to automate it. This is a context in which the aerospace industry assumes a leading role.

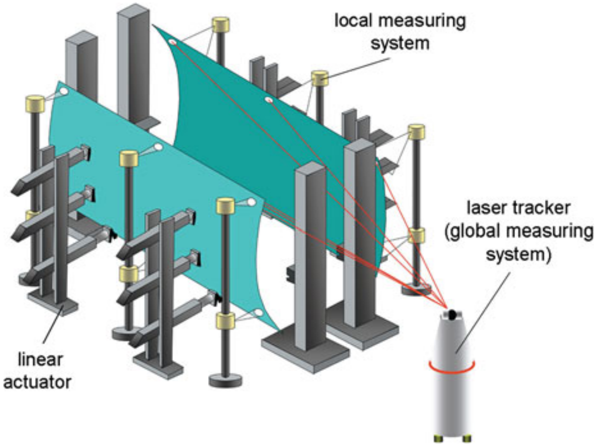


Fig. 6.134 Positioning system in structural assembly of aircraft fuselage sections. (Wollnack and Stepanek 2004)

One approach to reducing product-specific jigs and fixtures, means of fastening, clamping and the like in aircraft structural assembly is the use of linear actuators (Fig. 6.134). The shell elements of the aircraft fuselage are taken up by a number of linear actuators and fed to the joining station. This increases flexibility compared to jig and fixture-oriented assembly systems. The configuration of the actuators is specified however, and workspace is limited.

The use of robot systems can further increase flexibility for automated assembly of large components in changing applications. Cooperating robots enable the implementation of highly flexible assembly cells. By resetting and reprogramming it is then possible to use cooperating industrial robots for different purposes. Inflexible jigs and fixtures to handle large components can be omitted because the component can be grasped and supported at different points by multiple robots.

But there are still a number of drawbacks associated with the use of cooperating robots. In addition to the high price of industrial robots there is, in particular, the time taken up by programming, or shutdowns needed for conversion and new programming.

Two concepts have emerged to date to control cooperating robots. Some manufacturers use a system in which a master controller consisting of a control computer and drive components is expanded by additional drive components for each extra robot (Fig. 6.135). The advantage of having just one control computer (i.e. single controller solution or SCS) is that the extra cost for each additional robot is very much less than if each robot had a fully featured controller with computer and drive modules (Bredin 2005). The disadvantage is that the system cannot be composed in any random way because a powerful central processor is always needed. There are also restrictions when it comes to achieving single operations of the handling modules, for feeding purposes for example.

Contrasting with this is a control concept in which each robot possesses a fully featured controller and there are consequently a number of computers available in

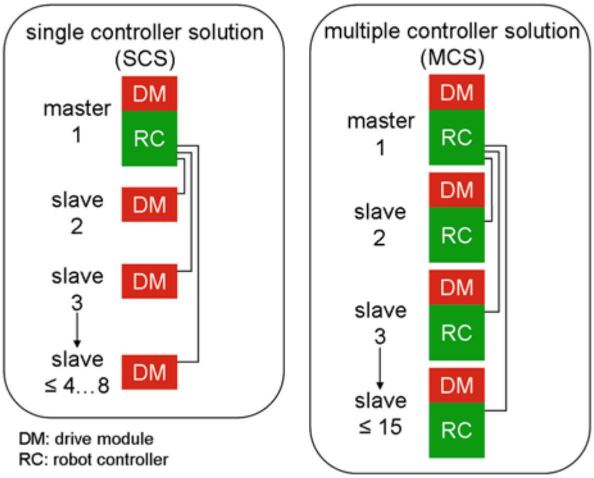


Fig. 6.135 Hardware concepts to control cooperating robots

the robot network (i.e. multiple controller solution or MCS). This concept can involve higher investment costs because each robot has its own complete controller. The additional computing power means that significantly more robots are able to cooperate. Despite a high-speed Ethernet connection the number of robots is still limited to 15 units (Kuka 2011a).

In an application with cooperating robots three operating modes are distinguished. In the simplest mode all robots act independently of one another but in overlapping workspace. One example is pick & place tasks where there is no direct interaction between the robots.

Other applications may require synchronised motion, i.e. where robots synchronise at certain points along a programmed trajectory. This makes sense, for instance, when one robot brings a component into a position for machining by a second robot. The situation requires coordination between the robots so that one can tell the other that a component has reached the position for machining, for example, or that machining is completed.

Precise path control of all robots is essential for coordinated motion, e.g. when they jointly move an object, because path deviations could result in deformation of or damage to a component. So the response of robots must be modelled as precisely as possible in the controller system. The single paths must also be generated at a fast interpolation rate to minimise path deviations between the computed interpolation points (Stoddard et al. 2004; Feldmann et al. 2007).

The programming of cooperating robots is independent of the hardware concept and similar for all operating modes. Each robot is invested with its own program, which is partly or entirely independent of the programs of the other robots. The robot can change between the operating modes within the program and coordinates with the other robots through the coordinate system of a master robot.

The disadvantage of this programming concept is that each robot possesses its own application-specific program that has to be configured for each new application. A high number of robots and frequent changes in work content mean a large programming effort, calling for more efficient concepts.

A further problem is the accuracy of robot systems. The position errors that appear are often kinematic in nature. They are caused by deviations between real system performance and the calculation model that is used.

The kinematics of a robot are frequently identified and modelled with inadequate precision. Instead of robot-specific calculation models, nominal models are used that do not correspond accurately to the genuine kinematics.

The length of robot components can vary as a result of production and measurement errors. A further error source is the position of rotary axes in relation to one another. There may be deviations in position if robot axes according to the model are not precisely parallel or at right angles to one another. In addition to axial position errors and deviations in length, another source of error is the zero positions of robot axes, which depending on kinematics may lead to a position and/or orientation error.

CAD data is used to model the other resources of the assembly system. Here too, assembly and manufacturing tolerances are not taken into consideration. Special importance attaches in reconfigurable systems to the position of the industrial robot, of the process station and of the handling object. When setting up heavy resources in particular, deviations of several millimeters may occur, as a result of which functional tolerances are no longer maintained.

Consequently these error effects must be measured and identified after assembly, a modification or repair, and taken into account appropriately in the model or for programming.

6.7.3 Research Questions

The aim is consequently to develop an assembly platform for reconfigurable self-optimising component handling that can be composed for differing handling purposes by a modular, versatile structure. It consists of multiple simple handling modules that are compact in shape and low in weight to allow fast reconfiguration. Use in a multiple-arm combination permits in particular the efficient handling of large-area components.

Two scientific questions, describing the individual tasks in more detail and allowing a target-oriented approach, were outlined for each of the four focal subjects presented in Fig. 6.136.

Under the heading “Design of versatile assembly units” those questions are resolved relating to the development of a reconfigurable handling system, covering kinematic reconfiguration of the hardware and the management of mechanical, electronic and software interfaces. To keep interface management as simple as possible, the approach consists in designing each module so that it is mechatronically self-contained, and in featuring it with ready defined interfaces so that it can be part of a mechatronic building set.

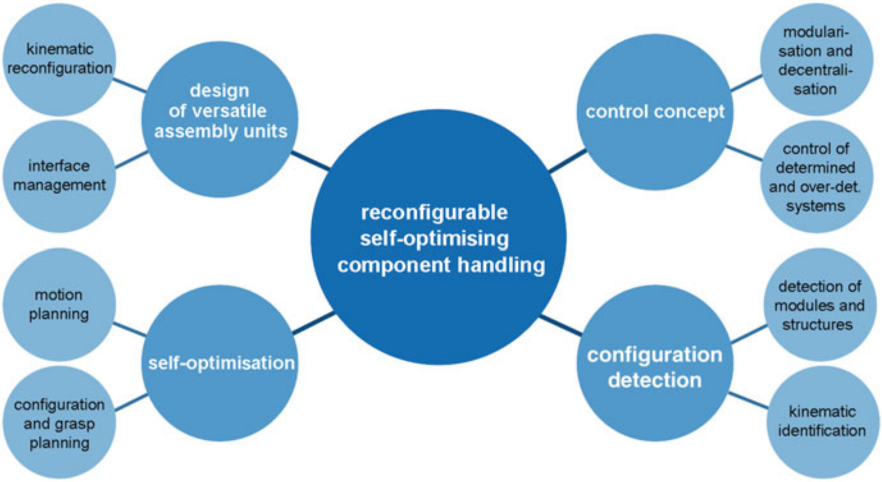


Fig. 6.136 Scientific questions in subproject

The subject "Control concept" deals with controlling and programming the modular handling system. This covers considerations of data storage within control, which must be managed decentrally at module level (e.g. transaction sets) on the one hand and centrally in a superordinate controlling instance (e.g. coordination of modules) on the other. Depending on the drive configuration the handling system can work both statically determined, i.e. with a total of six driven axes, and over-defined with further axes. A concept is devised by which both over-defined and exactly defined handling and multi-robot systems are simply programmed by specifying a path for the component from which the movements of every single handling module are indirectly derived.

The subject "Self-optimisation" looks at featuring the handling system with additional functions that support the user in planning the right movement and creating a suitable system configuration. To optimise movement, the trajectory of the component is analysed with computer support and subsequently adapted by the user interactively. Optimal system configurations can be calculated by an algorithm that optimises different weighting criteria, e.g. maximum drive speed, moment or accuracy, with reference to specified limits and thus determines the positions of the basic or reach coordinates for each handling unit.

"Configuration identification" means identifying the particular system configuration by means of measurement technology. After executing special movement patterns, the positions of the individual handling devices can be determined and deviations that appear in structuring the system can be compensated in control. Calibration of the handling devices is also provided to determine the exact kinematic parameters and match the calculation models. This is necessary to achieve maximum positioning accuracy of the robots.

6.7.4 Results

6.7.4.1 Idea of a New Handling System

Robot systems are often chosen to create the flexibility needed in an automated assembly used for changing applications. Cooperating robots are a solution for implementing flexible assembly cells. By reconfiguring and reprogramming it is possible to use industrial robots for different purposes. Drawbacks are the high purchase price of industrial robots and the complex programming. The need for synchronised path control, which also results in substantial setup time, means that such robot systems are often not an economical solution.

Against this background a new kind of handling strategy was developed that enables cooperative handling by simple handling modules and without elaborate setup procedures.

The handling method produced by this concept distinguishes three movement phases: coordinated movement, approaching movement and retracting movement. During coordinated movement, analogous to coordinated robots, multiple handling devices are joined to the component in a multiple-arm combination. Unlike conventional cooperating robots, the wrist axes are passive. During approaching and retracting movement the handling devices work autonomously and behave like separate mechanisms.

To implement this handling concept, and based on groundwork performed by the Department of Mechanism Theory and Dynamics of Machines (IGM), a suitable handling device was developed with six degrees of freedom (Fig. 6.137) (Müller et al. 2010).

The kinematic structure of the handling device corresponds to that of a vertical jointed-arm robot. A parallel double crank drives the third axis to apply the weight and inertia of the motor to the robot base. The wrist takes the form of a central hand and, unlike industrial robots, consists of three passive joints. In addition to the monetary saving by reducing the drive concept to only three active principal axes and three passive secondary axes, the handling device is lightweight and can be shifted by one operator without extra aids. The active degrees of freedom of the robot make it possible to join to the component by a certain grasp movement through a suitable contact element (magnetic or vacuum). The drives of the principal axes are actuated

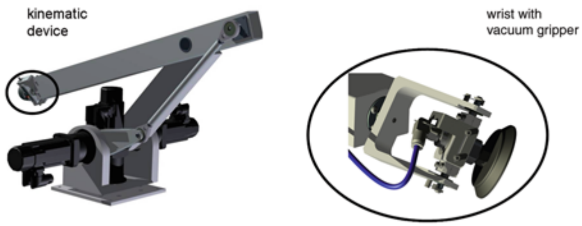


Fig. 6.137 Handling device with three active principal axes and three passive secondary axes

during the cooperating movement. The movement of the joint axes results from the kinematic structure of the multiple-arm combination, i.e. the wrists track passively.

In assembly platforms with multiple handling devices in particular, the economic advantage that results from reducing the active drives is considerable, because both the number of motors and the inertia plus complexity of the robot hand are very much reduced.

6.7.4.2 Steps of Reconfigurability

The wide-ranging field of use and the adaptability of the assembly platform to changing applications are achieved by multiple steps of system reconfigurability (Fig. 6.138). The usual boundaries such as workspace and choice of component or application are dissolved to produce an adaptable system that is reconfigurable and universally applicable.

The first step of reconfigurability is realised by altering the grasp-points on the object. This may be necessary if the planned object path leads to collisions between individual handling modules. Furthermore, rearrangement of the grasp-points is called for when the component changes. The second step of reconfiguration is relocation of the handling units within the assembly platform to design the workspace. The third step of reconfiguration involves enlarging or reducing the system by a number of handling devices. If more arm units are available in the layout than object handling requires, the object can be passed from one arm combination to the next. This re-grasping enables to re-design the workspace even during continuous object movement. Large translational object manipulations become possible, as well as unlimited object rotations about any arbitrary axis.

6.7.4.3 Mechatronic Modularisation of the Handling System

The prerequisite for these reconfiguration possibilities is seamless mechatronic modularisation of the handling system. Initially, module limits were defined to mark the

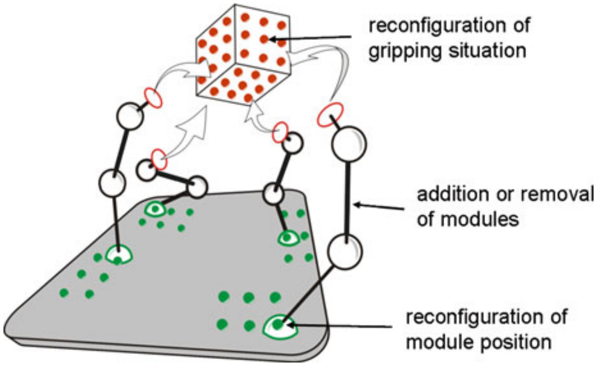


Fig. 6.138 Multiple steps of reconfigurability of an assembly platform

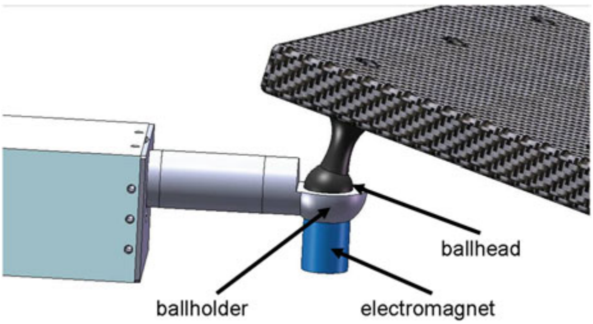


Fig. 6.139 Ball joint for handling non-magnetic large-area components

capabilities of a handling unit. In general terms, a module should exhibit a certain degree of autonomy, i.e. not be dependent on other modules, to prevent any meshed associations between modules and enable them to be exchanged as quickly as needed. Associated with this is the requirement to be able to assemble and start up a module in advance to ensure a fast restart upon replacement or reconfiguration. To satisfy these requirements it is consequently important that each handling device should be mechanically independent and also possess its own control, meaning mechatronic modularisation of the handling device. In addition to mechanical interfacing, simple control interfaces in particular are needed to ensure reconfigurability and scalability of the overall system.

To create appropriate mechanical interfacing of the handling module, the first three driving axes of the prototype were configured for simple conversion. The drives of the first axes were arranged coaxially to the particular rotary axis and the drive of the first principal axis positioned between the members of the handling device so that the latter can be set up on a baseplate. This is attached in the assembly cell by a drilling pattern. So the handling devices can be positioned in the same plane and extra degrees of freedom are provided for reconfiguration.

A further mechanical interface is the design of the wrist. Depending on the application, different solutions are conceivable for implementation of the handling module. Examples in the field of aircraft structure assembly are vacuum grippers (Fig. 6.139) or ball joints to handle large components.

The control of a handling module must possess functionality to control the drives, a local safety system, path interpolation and communication functions. Given the decentral control functions, the handling modules are designed so that they can be operated both autonomously and decentrally in association with a central controller. The assembly platform can thus be configured to requirement and reconfigured after completing an application.

Each handling module consists of a decentral control, a handling device with the mechanical structure, the servo drives and servo amplifiers.

The central module coordinates the assembly system and is responsible for cross-board functions such as communication and safety functions. Figure 6.140 is a schematic diagram of the related control system with the central module and the

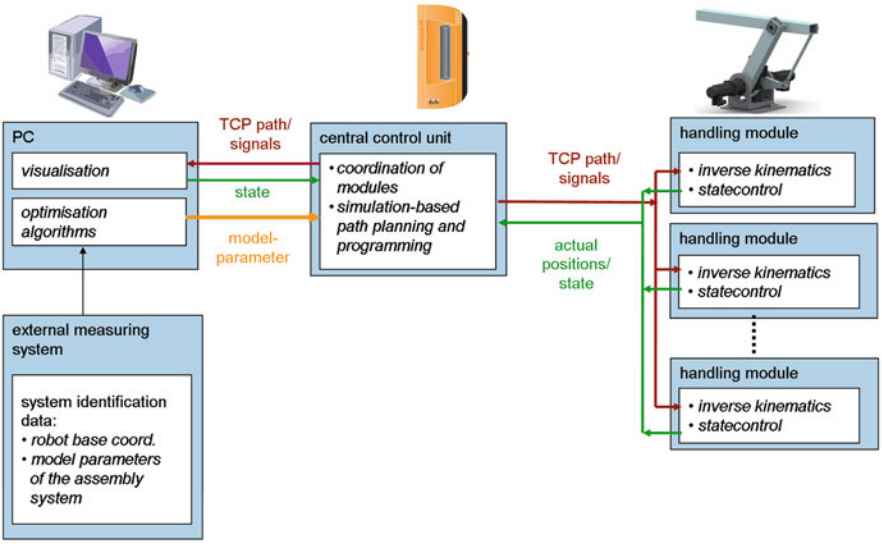


Fig. 6.140 Control architecture of the assembly platform

decentral handling devices. A simulation PC is also integrated that is responsible for optimisation functions in configuration planning and visualisation.

In addition to the mechanical interfaces of the handling device, the electrical and software interfaces were designed so that the system can be reconfigured fast without elaborate wiring (Fig. 6.141). The electric circuitry of each module was encapsulated

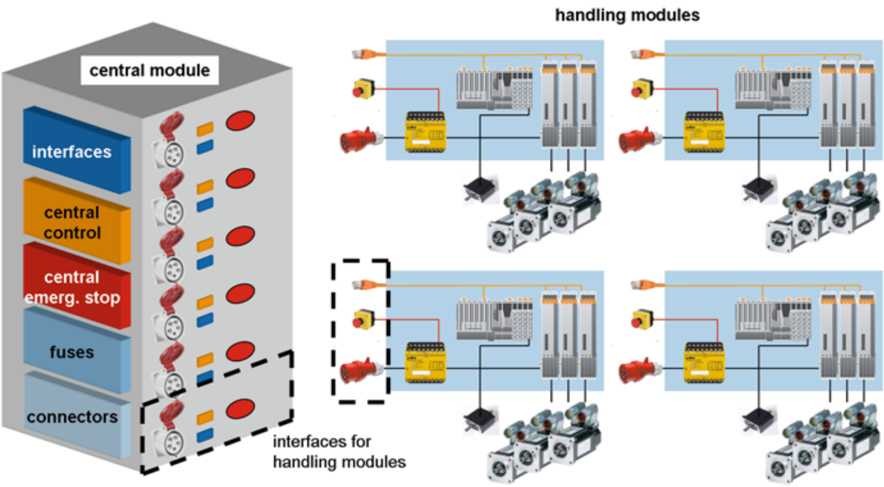


Fig. 6.141 Control concept of assembly platform with central module and decentral handling modules

in a separate cabinet and is self-contained. To connect the modules, different connectors were integrated in the central module and in the decentral handling modules to supply the decentral modules with power, enable communication with the decentral modules and connect the different emergency-off systems.

6.7.4.4 Grasp and Configuration Planning

Modularisation and reconfigurability generate new degrees of freedom. The reconfiguration steps give the user a variety of possibilities for configuring the assembly platform. But this major advantage compared to conventional *rigid* systems will only produce benefits in terms of efficiency and cost-effectiveness if both reconfiguration itself and the planning of an optimal configuration can be performed simply and fast. In addition to the accessibility of the grasp-points, stiffness, accuracy, load capacity and velocity transmission along the object motion must be considered as optimising criteria to find the optimal robot configuration.

A number of optimising tools were developed for this purpose, to automatically determine the additional adaptable degrees of freedom of the system, and thus to propose a possible configuration to suit the given application. In this way the assembly system is optimally reconfigured for a new application without prior testing.

Determination of grasp-points is automatically performed before each manipulation and has a large influence on the workspace of the object as well as the kinetostatic performances of the handling system, e.g. load capacity, stiffness or accuracy.

The handling devices used here possess a degree of freedom of $F = 6$, but unlike conventional industrial robots only three degrees of freedom are actuated. To be able to position and orient an object in the complete workspace, a handling system must have at least six actuated degrees of freedom. The handling concept is based on the combined working of multiple handling devices, so the system is even over-defined during cooperating component handling with three handling devices and a total of nine active driven joints. Further reduction of the number of drives at each individual handling device makes little sense because these three active degrees of freedom are needed to be able to grasp the object independently. The economic advantages compared to three cooperating robots with 18 active axes are nevertheless considerable because the concept used here saves nine drive units in this configuration.

This under-actuation of a separate handling device must be considered in grasp planning and performance however. The grasping pattern was consequently analysed in detail. The wrist features a spherical architecture with three serially arranged rotational joints whose axes intersect at one point. Figures 6.142 and 6.143 show a detailed view of the kinematic structure including terms and angles plus the prototype of a magnetic gripper.

Alignment of the contact element, e.g. an electromagnet or vacuum gripper, is passive upon contact with the surface of the object. Until contact with the object the joint is positioned centrally by spring forces. The sequence of a grasping operation with passive joint movement about the hand axes α_R and α_S is shown in Fig. 6.144.

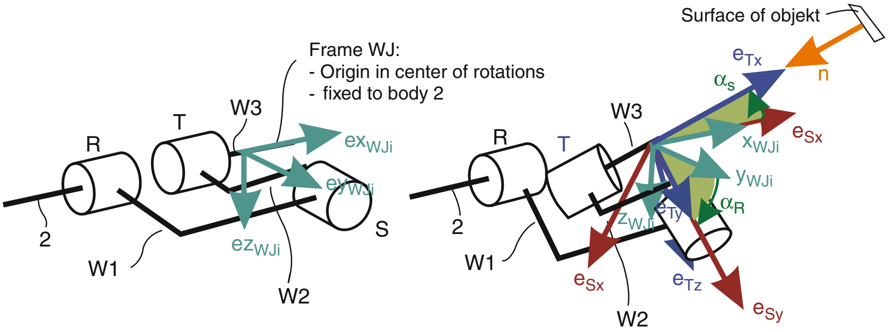
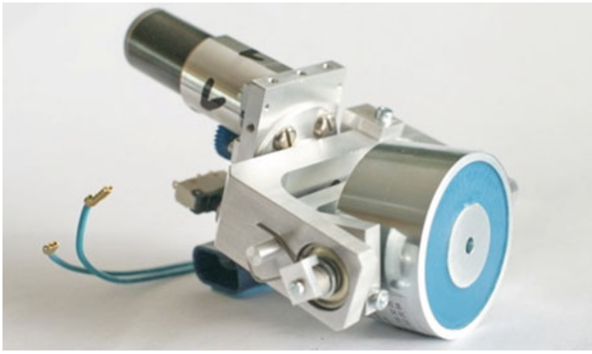


Fig. 6.142 Wrist structure: central position, deflected position

Fig. 6.143 Wrist with electromagnetic contact element



By experiments on a test bench for grasping movements and suitable mathematical models, it was possible to identify and optimise the parameters that are decisive for a good grasp pattern. In (Riedel et al. 2010a) it is shown that the area where the object could be placed on the ground and be grasped successfully can be multiplied almost tenfold by several optimisations concerning the grasp-process and the design of the wrist (Fig. 6.145).

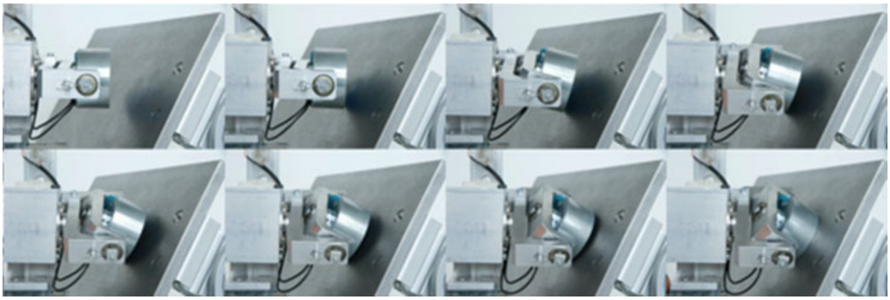


Fig. 6.144 Passive alignment of wrist joint when grasping

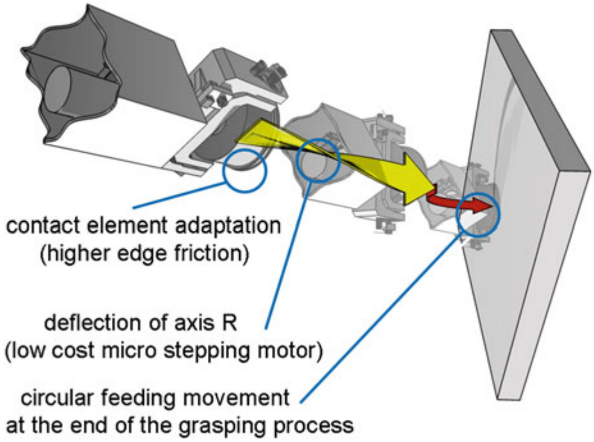


Fig. 6.145 Improvement of the grasp process

Taking a configuration for handling a cube-shaped object with magnetic grippers as an example, it was possible to expand the grasp area, which initially was only in the middle of the workspace (Fig. 6.146a), to the entire workspace at the level of the ground layer (Fig. 6.146b).

The major influencing variables are edge friction of the contact element and the grasp movement of the wrist. Spatial circular guidance of the wrist by the regional structure contributes very much to improving the complete grasp process. This circular path begins where the contact element sets down on the surface and is perpendicular to the joint axis e_{SY} of the wrist. The entire path of the wrists is shown in Fig. 6.147.

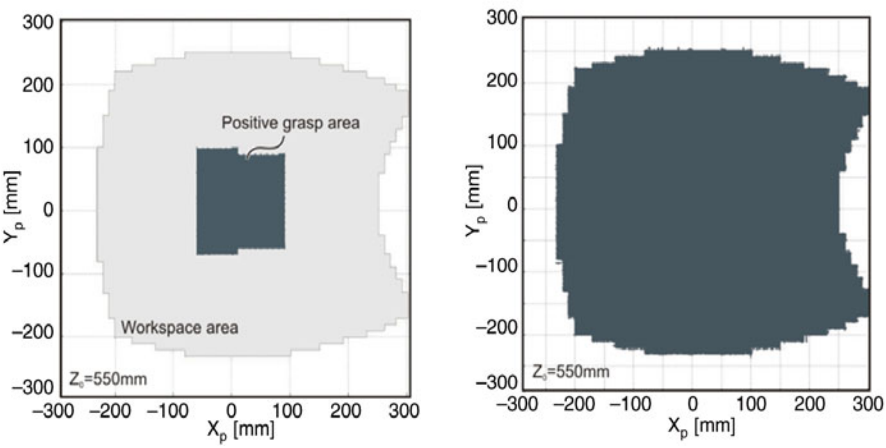


Fig. 6.146 Comparison of areas in which an object can be properly grasped. *Left:* grasp area before optimisation; *right:* grasp area after optimisation

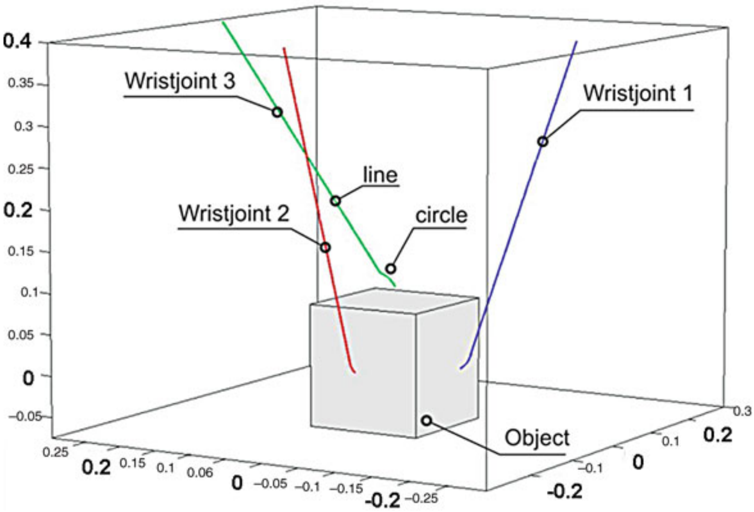


Fig. 6.147 Visualisation of grasp movement with circular track

Searching for and defining suitable grasp-points is performed by an optimising algorithm that is detailed in (Riedel et al. 2010b) and satisfies the following requirements:

- Optimisation of the grasp-points must consider the specific grasping pattern of the passive wrist.
- Grasp-points can only be selected and assessed correctly and meaningfully in a combination of the arms.
- Specification of a set of possible grasp-points or areas must be variable so that new objects can be integrated fast and simply.
- The user must be able to specify the objective of optimisation, e.g. minimal cycle time or maximum accuracy.

The possible grasp areas are given in the form of a set of discrete points or over continuous surfaces. Just nine points per face of a cuboid produce 148824 (54!/51!) possibilities of combination for the handling devices. Analysing each combination is not efficient. Consequently a smart and highly effective method of pre-selection was developed that first analyses only geometric information of the individual points. These include:

- It must be possible to grasp the point regarding the characteristics of the passive grasp process.
- The point must be within range of the handling device during the entire movement.
- All drive and joint angles must meet their permissible motion range during object manipulation.

Working through this sequence soon leads to exclusion of many of the points. Those points which fulfil the workspace conditions are comprised into a new set of possible combinations. The suitability of the combinations is examined at discrete points along the path, in that either the kinetostatic criteria allow a relative comparison between the solutions, or absolute performance data are evaluated by means of the dynamic model. If a combination of grasp-points satisfies all performance requirements, it goes into the final selection group. From this group the solution is then chosen that profiles best the user's needs. Figure 6.148 illustrates the flowchart for preselection for grasp-point optimisation.

If an object is mapped over continuous surfaces, a gradient-based method of optimisation can be used from dimensional synthesis in slightly modified form. The entire procedure of planning grasp-points is illustrated in Fig. 6.149.

The results of the different approaches to optimising are presented in Table 6.16. The first four solutions are selected manually, the next three (E through G) from a set of discrete points, and the last three (H through J) were optimised over a continuous surface.

The unsupported selection of grasp-points can easily produce unsatisfactory results. The cases show that grasp-points which were manually selected require too high driving torques (A), result in an undesirably large positioning error (B), can be grasped but violate the workspace condition during movement (C), or cannot even be grasped at all (D).

The automatic planning of grasp-points prevents these cases. In (E) all 148.824 combinations were evaluated, taking more than 12 min. of calculation. (F) uses smart preselection, so it considers all possible combinations and finds the best solution in the field after just half a second. Applications that are more time-critical can use search method (G), which halts at the first possible solution. Here the time to compute is negligible. But this approach only finds suitable solutions and not necessarily the best available. By its principle, optimising method (H) takes longer than a discrete search (E through G) because it takes at least 120 (6!/3!) optimizations if there is no preselection of surfaces. Furthermore, the optimising problem is not convex, so there is no assurance of finding a global optimum. That is why different starting values are given for surface optimisation, as a result of which the computing effort increases exponentially. Analogous to approaches (E through G) the time to compute is 4,757 s. for the complete run (H), 84 s. for optimisation with intelligent preselection, and 0.75 s. for optimisation with interrupt (J) (Table 6.17).

Despite the longer time to compute, these approaches are interesting if the performance of the handling system is to be improved further. In this way solutions are found that lie between the discrete points of the setting for (E through G), and further reduce the maximum driving torque from 8.6 Nm (F) to 7.5 Nm (I) or the maximum position error from 1.2 mm (F) to 1.0 mm (I). The choice of optimising approach here is user-specific or application-specific.

If a system embodies more than three handling devices, the combination of handling devices can also be optimised and selected for the movement application. Furthermore, the individual handling devices of the demonstrator can also be freely configured on the baseplate, resulting in a large number of possible combinations. If three handling devices can be set up at 100 base points, this alone means almost one million configuration possibilities–combined with the possible grasp-points the reconfigurability of the system is theoretically unlimited. For this purpose the approaches and algorithms of grasp-point selection were revised for efficiency and then

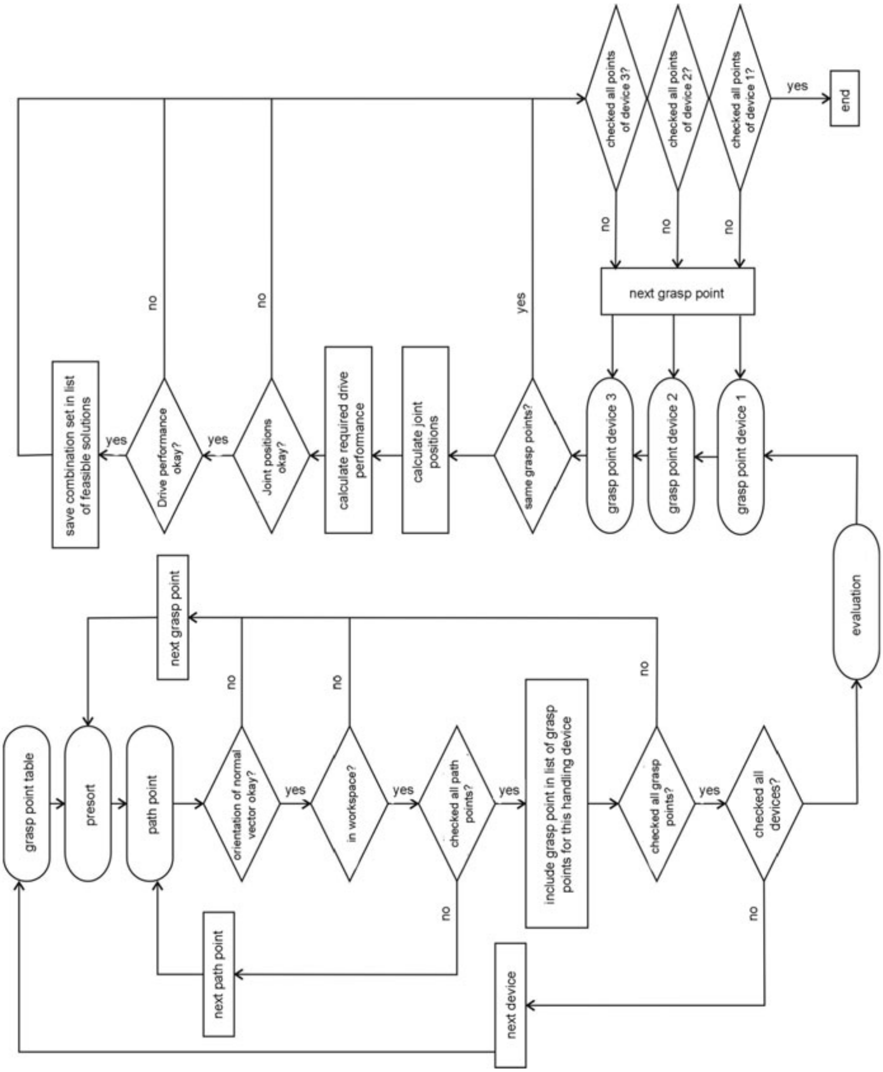


Fig. 6.148 Preselection to optimise grasp-points

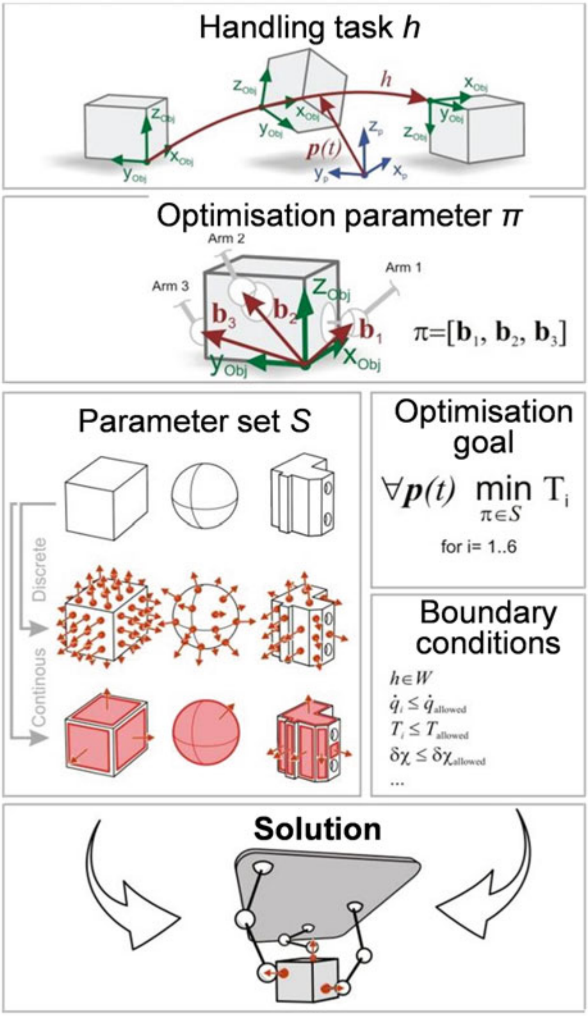


Fig. 6.149 Optimising procedure

implemented. With reduced grasping possibilities the automatic planning algorithm finds an optimal solution in about five seconds.

In reduced terms, the search for base points works as depicted in Fig. 6.150. After specification of the optimising parameters, goals and limits, a first fast selection searches the specifications field and reduces it by the points or areas that cannot be reached because of the range of the handling devices. At the same time a second preselection examines whether the point or area underruns a certain direction angle to the grasp surface normal during the movement. If this is so, this surface cannot be managed from the particular frame point and is not examined further. Experience shows that the specifications field shrinks to 2–5% of the initial size, and it is then newly sorted and combined. The possible combinations of handling devices are subsequently analysed dynamically or kinetostatically and selected analogously to the planning of grasp-points.

Table 6.16 Specifications and boundary conditions for optimising task

Task definition				
Path	*cubic spline*	P1	P2	P3
	$[x,y,z,]^T$ in [m]	$[-0,15;0,1;0]^T$	$[-0,05;0,15;0,1]^T$	$[0,1;0,15;0]^T$
	$[\varphi,\phi,\psi]^T$ in [deg]	$[0,0,0]^T$	$[0,45,0]^T$	$[0,90,0]^T$
Trajectory	law of motion	*5th order polynomial*		
Motion	duration	*1*	s	
	length	*0,48*	m	
	max. transl. velocity	*0,91*	m/s	
	max. transl. acceleration	*6,52*	m/s²	
Object	Shape	*Cube*		
	Size	*150x150x150*	mm x mm x mm	
	Mass	*1*	kg	
Optimisation goal				
	minimize drive torque along path			
Boundary conditions				
Restrictions	Grasp criterion	*true*		
	Workspace criterion	*true*		
Limits	Maximal drive torque	*15*	Nm	
	Maximal drive velocity	*1000*	rpm	
	Maximal object error δX	*3*	mm	

6.7.4.5 Path Planning and Programming

Programming of the assembly platform and/or handling system must also match the requirements of a decentral and reconfigurable control system. In conventional controllers for cooperating robots, application-specific programs are generated for each robot that map the particular movements and synchronise with the other robots through a network. The coordination effort is very high, for which reason online methods such as teaching and playback are unsuitable for programming multi-robot applications or for frequently changing purposes.

A more suitable method is offline programming where programs are created without using the real machine. The programs needed are generated and validated first in a simulation environment, and then transferred to the controller. For this purpose the simulation environment requires a model of the assembly system.

Table 6.17 Comparison of optimising approaches

Results										
Optimization method	Case	Grasp point vectors b_i [x;y;z] in respect to frame F_OBJ in [m]			Optimization Time in [s] @2,66GHz	Grasp criterion	Workspace criterion	max T in [Nm]	max q in [rpm]	max δX in [mm]
		b_1	b_2	b_3						
non	A	-0,075 -0,060 0,030	0,030 0,075 0,060	0,000 0,030 0,075	0	true	true	**31,0**	784,4	2,2
non	B	-0,075 0,060 0,010	0,060 0,075 0,060	0,000 0,030 0,075	0	true	true	11,9	690,1	**6,8**
non	C	-0,075 0,060 -0,060	0,075 0,060 0,060	0,075 0,000 0,000	0	true	**false**	-	-	-
non	D	-0,075 0,060 -0,060	-0,075 0,060 0,060	0,075 0,000 0,000	0	**false**	**false**	-	-	-
discrete	E	-0,075 0,060 -0,060	0,060 0,075 0,060	0,000 0,000 0,075	734,59	true	true	8,6	647,1	1,2
discrete	F	-0,075 0,060 -0,060	0,060 0,075 0,060	0,000 0,000 0,075	0,51	true	true	8,6	647,1	1,2
discrete	G	-0,075 -0,060 -0,060	-0,060 0,075 -0,060	0,000 0,000 0,075	0,10	true	true	13,3	647,1	2,6
continuous	H	-0,075 -0,049 -0,060	0,043 0,075 -0,024	0,060 -0,035 0,075	4757	true	true	7,5	575,6	1,0
continuous	I	-0,075 -0,049 -0,060	0,043 0,075 -0,024	0,060 -0,035 0,075	84,50	true	true	7,5	575,6	1,0
continuous	J	-0,075 -0,060 -0,060	-0,060 0,075 -0,060	0,060 0,029 0,075	0,75	true	true	8,2	567,3	1,3

To shorten the setup time, the programming concept (Fig. 6.151) developed in this project implies that handling devices are controlled directly from the simulation environment. In this way no programs have to be exchanged on handling modules after a reconfiguration. The handling modules are simply reconfigured, the simulation model is matched appropriately, and handling modules are given setpoint values to match the new model.

The starting point for this programming is the component or the handling application. In a first step the handling system is modelled in the simulation environment. A collision-free component path is then defined by an initial pose, a final pose and the mode of movement. Proceeding from this component path, configuration and grasp

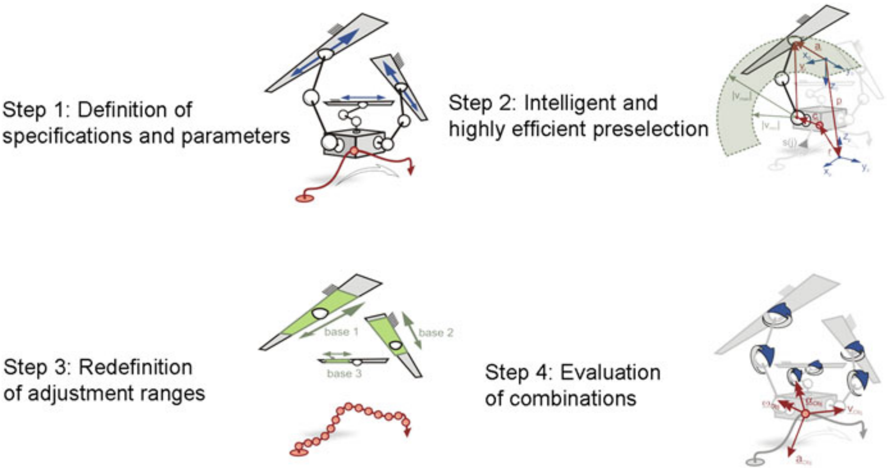


Fig. 6.150 Schematic of search for frame points

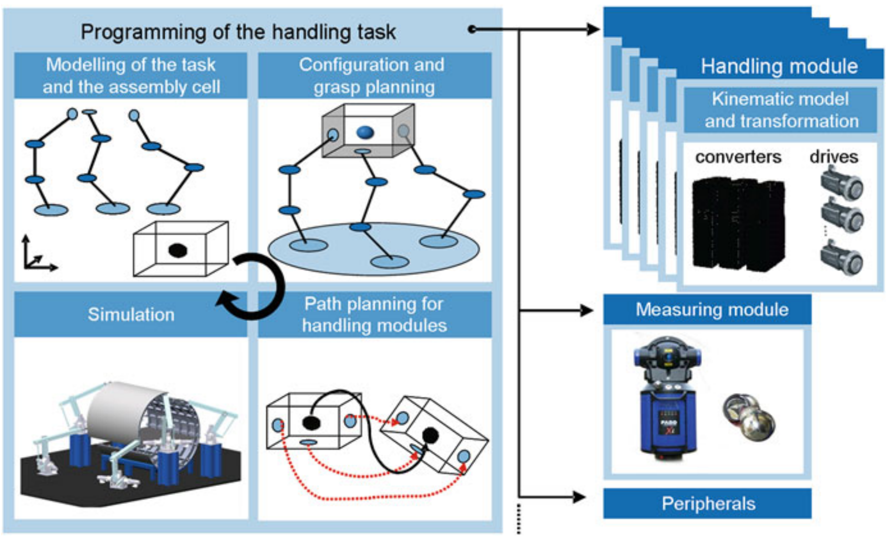


Fig. 6.151 Programming concept for reconfigurable assembly system

planning is performed, and the handling modules are matched appropriately in the simulation environment and the real assembly system.

The grasp-point paths of the handling modules can then be derived from the component path and the defined grasp-points on the component by simple transformation from the component coordinate system to the grasp-point coordinate system. Intermediate points of the component path are calculated in the interpolation cycle. Then the positions of the grasp-points are calculated by the fixed transforms to each intermediate point, and sent in the interpolation cycle on an interface as a setpoint to the handling modules (Fig. 6.152).

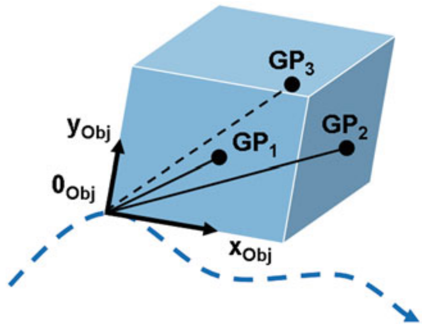


Fig. 6.152 Location of grasp-points (GP) and of component coordinate system as function of component path

The decentral handling modules read this interface in the interpolation cycle and use inverse kinematics to calculate the joint angles for the particular grasp-point, which are then set by the servo amplifiers. If necessary, finer interpolation of the set-points is possible decentrally. This kind of setpoint generation very much simplifies the complex programming of cooperating robots.

Movement of the multiple-arm combination can be very complex, so a possibility was created of visualising the programs first. The setpoints on the controllers are calculated and then sent not to the servo amplifiers but instead to an interface for visualisation. The simulation cell is then updated by the real controllers in the interpolation cycle. A collision check supports the operator in judging movements.

By avoiding application-specific programs on the individual modules it is possible to exchange or add single modules. Thus, for example, industrial robots or other available handling devices could adopt the task of a handling module.

6.7.4.6 System Identification

An exact description of the system pattern is necessary for direct transfer of programs generated offline. Handling devices are calibrated after installation to significantly reduce the unavoidable gap between simulation and reality.

Here the real kinematic parameters of the handling devices are determined as accurately as possible. An external measuring system registers different poses and the associated joint angles. From the joint angles and the nominal model it is possible to calculate the set poses of the handling devices and compare them to the measured poses. Then, by a numerical method, the descriptive model parameters are altered with reference to gradients until the error between the measured pose and the calculated pose is minimal (Fig. 6.153). These parameters are subsequently returned to the kinematic model and stored on the controller. The real kinematics of the handling system documented in this way are then used for direct and inverse kinematic calculation. The additional setup time for matching simulated programs is thus minimised—so the possibility of real offline programming is given.

For further reduction of setup times, handling devices are automatically calibrated after reconfiguration. It is sufficient to position the units roughly in the assembly

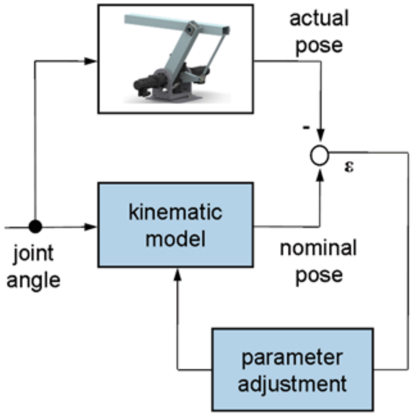


Fig. 6.153 Identification of optimal model parameters by recursive parameter matching

cell according to the configuration planning. They are then identified by short test movements and the appropriate transforms in the calculation model are matched to the real cell configuration.

By this system identification it is possible to determine the relevant model parameters of the assembly platform and of other objects in the assembly cell with high accuracy.

6.7.5 Industrial Relevance

Faced with constantly increasing cost pressure, a broader spectrum of types and shorter innovation cycles, there is greater demand for more flexible assembly and handling systems aimed at individualised production. As a cost-effective approach for adapting component-dependent applications, the concept of cooperating robots was further developed, based on grasping and moving objects with multiple handling devices. On this basis a reconfigurable, modular assembly platform is being developed which, beyond its handling devices, also allows the integration of measuring, testing and joining modules. In addition, the modular concept fulfils requirements for flexible and demand-driven configuration of multiple handling devices. An appropriate control system enables common handling of large components across several grasp-points. This also responds to a call from industry for harmonisation of production systems, i.e. the use of uniform systems for different components at all company locations.

One possible example of use of such an assembly platform is in aircraft structure assembly. The fuselage structure of a passenger aircraft usually consists of a number of sections (Fig. 6.132). The individual sections are composed of shell elements that are reinforced by ribs and stringers (Engmann 2008). The scope of work and the joining technologies for the sections are basically the same, but the sizes and shapes of the sections can differ. Today, to position the fuselage shells to join an aircraft

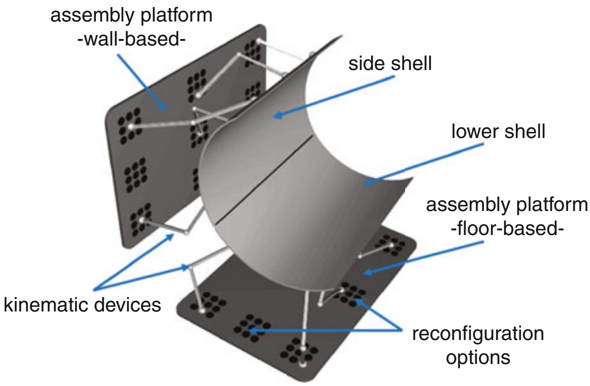


Fig. 6.154 Shell assembly

section, large rigid jigs and fixtures are used that map the component geometry. To maintain the required tolerances, the jigs and fixtures need a high degree of rigidity, which is usually only possible by a solid construction. As a result, the jigs and fixtures are large, heavy and costly.

An economically attractive alternative is reconfigurable assembly platforms, which are configured with handling devices to meet the demands of the specific application, and that will be able to replace the rigid jigs and fixtures in future. Here the assembly platform is configured according to the shell.

Since the workspace of the handling devices is very small compared to the dimensions of the shells, the shells must be conveyed by an additional handling system. While the transport system executes the macro movement of the component to the assembly station and roughly positions it there, the multiple-arm combination executes the smaller and more complex joining movement (Fig. 6.154). Shape and position corrections are possible by additional measuring means.

By separating the macro and micro movement and through the possibility of reconfiguring the multiple-arm combination, very different components can be handled and assembled.

For the purpose of combining large and smaller structures, there are different solutions such as wall and overhead gantries or Omnimove from KUKA (Kuka 2011b). The focus of the work is consequently on efficient, simple and cost-effective reconfiguration of the multiple-arm combination. The exact joining movement is executed by the handling devices. These position the two shell elements so that they can be joined manually or by automated means.

The advantage of this system is that the assembly platforms can always be configured to match the application, and the handling devices can be re-used for other purposes and products. This also makes the pilot series production of new aircraft types economical.

Another application example for this assembly platform is stringer integration. CFRP stringers of more than 20 m in length are laid in a shell and then pressed on for adhesion. Figure 6.155 shows a possible automated solution for this purpose. The

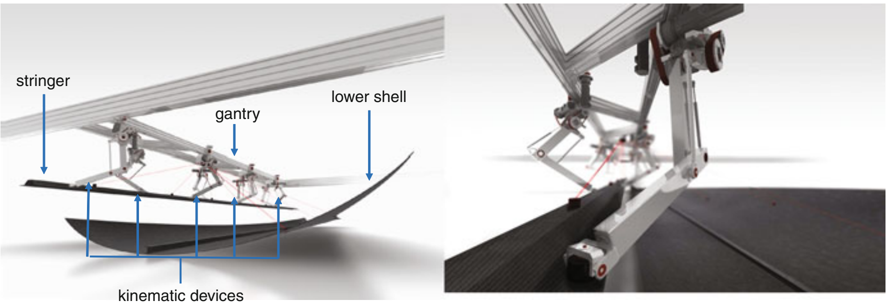


Fig. 6.155 Stringer integration

shell is positioned under an overhead gantry. Attached to the gantry are a number of handling modules that all engage with the stringer to prevent inadmissible sag.

Modularisation enables the number of handling devices to be matched to the length of the stringers. The low weight of the handling devices (approx. 25 kg) and of the stringer enables very simple and cost-attractive design of the gantry. In addition to laying stringers, the handling devices can lay further items of equipment needed for the autoclave process.

To achieve the previously mentioned scaling of the assembly system by addition and removal of handling modules with as little effort as possible, modularity must be consistent and interdisciplinary, from the mechanical system through the controller hardware to the controller software. The described controller concept can be implemented not only for handling devices but also for more extensive production systems that have to be reconfigurable. What is decisive is the modularisation and use of standard interfaces.

The presented, reconfigurable handling concept can be used to great advantage in other sectors of industry, wherever free configurability of large handling solutions can replace rigid jigs and fixtures or show the way to partial or full automation of small series.

6.7.6 Future Research Topics

Increasingly complex products, a growing number of variants and shorter product life cycles mean that automated production systems have less time to pay back.

As a result there is a continuing development in the automation of assembly towards flexible and adaptable systems accompanied by a variety of new applications. The greater functionality and configurability of assembly systems create new degrees of freedom. To enable optimum use of all this, self-optimising functions are aimed at reducing the effort and expense of reconfiguration as well as programming and planning complex applications.

A future research topic will be the integration of additional measurement and sensor technology. In the context of self-optimisation in particular, it is important to quickly identify system states. For this purpose, various methods are to be developed that not only determine the positions of robots but also the position of the component. The position data will then be used to monitor the component path and correct it if necessary. The correction data will be sent over a sensor input to the central controller, which automatically corrects the individual robots. In addition to monitoring of the component path, force/moment sensors can be integrated, to check and control the forces applied to a component. Forces could be intentionally introduced by tensioning of the robots to correct the shape of a component.

The description of a movement used in this project by reference to a component path very much simplifies the programming of cooperating robots. The programming concept, where the entire application is described in a central sequential program and the handling devices only possess simple standard programs to evaluate commands, also reduces the programming and modification effort for frequently changing tasks.

Programming of the handling system is nevertheless still elaborated in terms of the current state of development because individual movements of the handling devices outside the mode of cooperation have to be specified by the user. This is the case with the grasping movements for example, when every single handling device must be taken from the initial position to the grasp-point by a sequence of single movements.

Semi-automated path planning makes sense especially for these frequently recurring sequences, to support the user in planning and programming application-specific movements.

A future research topic is consequently the development of semi-automated path planning, by which a further function of self-optimisation can be integrated in the assembly platform, operation is simplified, and user acceptance can be enhanced. Existing methods of automated path planning will therefore be analysed, judged for suitability for the handling system and further developed.

Dynamic collision areas, which can be both machines, e.g. other robots, and humans, are a far greater challenge. In the case of machines and other controlled cell components, the position can be read from the controller and considered in path planning. Detecting the position of other dynamic obstacles requires external measuring technology (e.g. imaging), which means a much higher investment and stricter safety stipulations.

References

Abele E, Liebeck T, Wörn A (2006) Measuring flexibility in investment decisions for manufacturing systems. Ann CIRP 55(1):433–436

Adelt P, Donath J, Gausemeier J, Geisler J, Henkler S, Kahl S, Klöpper B, Krupp A, Münch E, Oberthür S, Paiz C, Porrmann M, Radkowski R, Romaus C, Schmidt A, Schulz B, Vöcking H, Witkowski U, Witting K, Znamenshchykov O (2009) Selbstoptimierende Systeme des Maschinenbaus. In: Gausemeier J, Rammig F, Schäfer W (eds) HNI-Verlagsschriftenreihe, vol 234. Westfalia Druck GmbH, Paderborn

Agrawal G, Jin R, Yang G (2004) Shared memory parallelization of data mining algorithms: techniques, programming interface, and performance. IEEE Trans Know Data Eng 16(10):1247–1262

Ahsan M, Hasin A, Pandey P (1996) MRP II: should its simplicity remain unchanged? Ind Manag 38(3):19–21

Akgün M, Alpay HR, Süle G Eren R (2010) Influence of warp tension on breking strength and strain of woven fabrics. Tekstil ve Konfeksiyon 1:30–36

Albert C, Fuchs C (2007) Durchblick im Begriffsdschungel der Business-Software. Universität Würzburg, Germany

Alred GJ, Brusaw CT, Oliu WE (2003) Handbook of technical writing. Bedford/St. Martin's, Boston

Andreasen M, Ahm T (1988) Flexible assembly systems. Springer, Berlin, p 44

Arburg (2010) http://www.arburg.com/com/COM/de/products/ac. Accessed 20 July 2010

Askin RG, Goldberg JB (2002) Design and analysis of lean production systems. Wiley, New York

Assem D (1986) Sensitivity analysis of linear systems. Springer, Berlin

Bainbridge L (1987) Ironies of automation. In: Rasmussen J, Duncan K, Leplat J (eds) New technology and human error. Wiley, Chichester

Balve P (2002) Ein Rahmenkonzept zur Gestaltung wandlungsfähiger Auftragsmanagementsysteme. Dissertation Universität Stuttgart. Jost-Jetter, Heimsheim

Bangor A (2000) Display Technology and Ambient Illumination Influences on Visual Fatigue at VDT Workstations. Published PhD thesis, Virginia Polytechnic Institute and State University, Virginia

Bannat A, Wallhoff F, Rigoll G, Friesdorf F, Bubb H, Stork S, Müller HJ, Schubö A, Wiesbeck M, Zäh MF (2008) Towards optimal worker assistance: a framework for adaptive selection and presentation of assembly instructions. In: Proceedings of the 1st international workshop on cognition for technical systems, Cotesys 2008, Technische Universität München

Barachini F (1990) Match-time predictability in real-time production systems. In: Proceedings of the 1st international Workshop on Expert Sys. Engineering, Lecture Notes in Computer Science, vol 462. pp 190–203

Bauer T (2008) Aufbau und Verbindungstechnik für Mikrooptische Systeme und Module, Jenoptik Polymer Systems GmbH, Jena, 6 Mai 2008

Becker M (1991) Anwendung von höheren Optimierungsmethoden in der Umformtechnik. Dissertation, Institut für Bildsame Formgebung, RWTH Aachen, Mainz, Aachen

Becker T (2008) Prozesse in der Produktion und Supply Chain optimieren. Springer, Berlin

Beckert E (2005) Ebene Keramiksubstrate und neue Montagetechnologien zum Aufbau hybridoptischerSysteme, Jena

Beckmann I (2009) Entrepreneurship-Politik: Neue Standortpolitik im politischen Spannungsfeldzwischen Arbeitsmarkt und Interessengruppen. VS Verlag für Sozialwissenschaften, Wiesbaden

Beer S (1973) Kybernetische Führungslehre. Herder und Herder, Frankfurt

Beer S (1979) The heart of enterprise. Wiley, Chichester

Beetz M, Buss M, Wollherr D (2007) Cognitive technical systems – what is the role of artificial intelligence? Adv Artific Intell 4667:19–42

Beierle C, Kern-Isberner G (2006) Methoden wissensbasierter Systeme Grundlagen – Algorithmen – Anwendungen. Vieweg, Wiesbaden

Belforte D (2009) Keeping the economy in perspective. Ind Laser Sol 24(1):4–11

Best E (2010) Process Excellence – Praxisleitfaden für erfolgreiches Prozessmanagement. Gabler, Wiesbaden

Blecker T, Kaluza B (2004) Produktionsstrategien – ein vernachlässigtes Forschungsgebiet? In: Braßler A, Corsten H (eds) Entwicklungen im Produktionsmanagement. Vahlen, München

Bleicher K (2004) Das Konzept Integriertes Management: Visionen- Missionen- Programme. Campus, Frankfurt a. M.

Borgelt C, Klawonn F, Kruse R, Nauck D (2003) Neuro-Fuzzy-Systeme: Von den Grundlagen künstlicher Neuronaler Netze zur Kopplung mit Fuzzy-Systemen. Vieweg, Wiesbaden. ISBN 3-528-25265-0

Brecher C, Fayzullin K, Possel-Dölken F, Valkyser B (2007) Optimierung der flexiblen Fertigung. Werkstattstechnik 97(6):464–470
Brecher C, Fayzullin K, Possel-Dölken F, Valkyser B (2008a) Optimierung flexibler Produktionsabläufe flexibel gestalten. ATP 5(50):60–69
Brecher C, Fayzullin K, Possel-Dölken F (2008b) Intelligent operations control. Prod Eng Res Dev 2(3):293–302
Brecher C, Herfs W, Fayzullin K (2008c) Operative Flexibilität – Ein workflow-basierter Ansatz für Fertigungsleittechnik, A&D Kompendium 2008/2009: Das Referenzbuch für industrielle Automation. Publish-Industry, München
Brecher C, Klocke F, Schmitt R, Schuh G (2008d) Wettbewerbsfaktor Produktionstechnik – Aachener Perspektiven. Apprimus, Aachen
Brecher C, Pyschny N, Loosen P, Funck M, Morasch V, Schmitt R, Pavim A (2009) Self-optimising flexible assembly systems. Workshop on self-X in mechatronics and other engineering applications, Paderborn, pp 23–38
Bredin C (2005) Team-mates – ABB multiMove functionality heralds a new era in robot applications. In: ABB Review 1/2005
Brennan RW (2003) From FMS to HMS. In: Deen SMH (eds) Agent-based manufacturing – advances in the holonic approach. Springer, Berlin
Buhmann MD (2003) Radial basis functions – theory and implementations. In: Cambridge monographs on applied and computational mathematics, no 12. ISBN 9780521633383
Busch A, Dangelmaier W (2004b) Integriertes Supply Chain Management – Ein koordinationsorientierter Überblick. Gabler, München
Castellini C, Giunchiglia E, Tacchella A (2001) Improvements to sat-based conformant planning. In: Proc. of the 6th European conf. on planning
Chryssolouris G (2005) Manufacturing systems – theory and practice. Springer, New York
Chryssolouris G, Mourtzis D (eds) (2004) Autonomous control of processes. In: IFAC conference manufacturing, modelling, management and control
Cisek R, Habicht C, Neise P (2002) Gestaltung wandlungsfähiger Produktionssysteme, ZWF – Zeitschrift für wirtschaftlichen. Fabrikbetrieb 97(9):441–445
Coit DW, Jackson BT, Smith AE (1998) Static neural network process models: considerations and case studies. Int J Prod Res 36(11):2953–2967
Cremers B (1991) Expertensysteme für die Planung der Produktion. Verlag TÜV Rheinland, Köln
Dangelmaier W, Warnecke H (1997) Fertigungslenkung – Planung und Steuerung des Ablaufs der diskreten Fertigung. Springer, Berlin
Denkena B, Shpitalni M (2007) Knowledge management in process planning. CIRP Ann Manuf Technol 56(1):175–180
DIN 32564 (2003) Fertigungsmittel für Mikrosysteme – Begriffe – Teil 1–3. Beuth, Berlin
DIN EN 61131-3 (2003) Speicherprogrammierbare Steuerungen. Teil 3: Programmiersprachen. Beuth, Berlin
DIN EN 62264 (2008) Integration von Unternehmensführungs- und Leitsystemen. Beuth, Berlin
Ding H, Kain S, Schiller F, Stursberg OA (2008) Control architecture for safe cognitive systems. In 10. Fachtagung Entwurf komplexer Automatisierungssysteme, Magdeburg
Dittmar R, Pfeiffer B-M (2004) Modellbasierte prädiktive Regelung: Eine Einführung für Ingenieure. Oldenbourg, München
Dohner JL, Lauffer JP, Hinnerichs TD, Shankar N, Regelbrugge M, Kwan C, Xu R, Winterbauer B, Bridger K (2004) Mitigation of chatter instabilities in milling by active structural control. J Sound Vib 269(1/2):197–211
Dören J (2007) Qualitätsmanagement und Neuronale Netze
Drabow G (2006) Modulare Gestaltung und ganzheitliche Bewertung wandlungsfähiger Fertigungssysteme. PZH Produktionstechnisches Zentrum, Garbsen
Draper NR, Smith H (1998) Applied regression analysis. Wiley series in probability and statistics. ISBN 978-0-471-17082-2

Drumwright E, Ng-Thow-Hing V, Mataric M (2006) Toward a vocabulary of primitive task programs for humanoid robots. In: Proceedings of the international conference on development and learning (ICDL). Bloomington, IN

Dyckhoff H (2003) Grundzüge der Produktionswirtschaft – Einführung in die Theorie betrieblicher Wertschöpfung. Springer, Berlin

Dyckhoff H, Spengler TS (2007) Produktionswirtschaft – Eine Einführung für Wirtschaftsingenieure. Springer, Berlin

Efron B, Gong G (1983) A leasurly look at the bootstrap, the jackknife, and cross-validation. Am Stat 37(1):36–48

Elatta AY, Gen LP, Zhi FL, Daoyuan Y, Fei L (2004) An overview of robot calibration. Inf Technol J 3(1):74–78

ElMaraghy HA (2005) Flexible and reconfigurable manufacturing systems paradigms. Int J Flex Manuf Syst 17(4):261–276

ElMaraghy HA, Azab A, Schuh G, Pulz C (2009) Managing variations in products, processes and manufacturing systems. CIRP Ann Manuf Technol 58(1)

Engesser H (1993) Eintragungen "Künstliche Intelligenz" und "Expertensystem". Duden "Informatik". Bibliographisches Institut und F.A. Brockhaus AG, Mannheim

Engmann K (2008) Technologie des Flugzeuges. Vogel, Würzburg

Enríquez Díaz JA, Frieling E, Thiemich J, Kreher S (2010) Auswirkung eines Chaku-Chaku-Montagesystems auf die älteren Beschäftigten am Beispiel der Abgasanlagen-Montage. In: GFA (ed) Neue Arbeits- und Lebenswelten gestalten. GfA-Press, Dortmund

Erdim H, Lazoglu I, Ozturk B (2006) Feedrate scheduling strategies for freeform surfaces. Int J Mach Tools Manuf 46(7/8):747–757

Espejo R (1996) Organizational transformation and learning – a cybernetic approach to management. Wiley, Chichester

Espejo R, Harnden R (1989) The viable system model – interpretations and applications of Stafford Beer's VSM. Wiley, Chichester

Eversheim W (1998) Organisation in der Produktionstechnik. Konstruktion. Springer, Berlin

Eversheim W, Schuh G (1996) Betriebshütte – Produktion und Management. Springer, Berlin

Eversheim W, Pfeifer T, Weck M (2006) Hundert Jahre Produktionstechnik. Springer, Berlin

Fayzullin K (2010) Autonome Planung und Entscheidungsoptimierung in der Ablaufsteuerung flexibel automatisierter Fertigungssysteme. Dissertation RWTH Aachen

Feldmann K, Ziegler C, Michl M (2007) Bewegungssteuerung für kooperierende Industrieroboter in der Montageautomatisierung. Werkstattstechnik 97(9):713

Ferry WB, Altintas Y (2008) Virtual five-axis flank milling of jet engine impellers part II: feed rate optimization of five-axis flank milling. ASME J Manuf Sci Eng 130:1

Field A (2005) Discovering statistics using SPSS. Sage, London

Finkler M (2006) Advanced planning and scheduling – die stille ERP-revolution. PPS Manag 3

Fleisch E (2008) High-resolution-management. Konsequenz der 3. IT-Revolution auf die Unternehmensführung. Schaeffer-Poeschel, Stuttgart

Fleisch E, Österle H (2004) Auf dem Weg zum Echtzeit-Unternehmen. In: Alt R, Österle H (eds) Retail-Time Business – Lösungen, Bausteine und Potenziale des Business Networkings. Springer, Berlin

Frank U, Gausemeier J (eds) (2004) Selbstoptimierende Systeme des Maschinenbaus: Definitionen und Konzepte/Sonderforschungsbereich 614. Heinz-Nixdorf-Institut, Univ. Paderborn, Paderborn

Frank U, Giese H, Klein F, Oberschelp O, Schmidt A, Schulz B, Vöcking H, Witting K (2004) Selbstoptimierende Systeme des Maschienenbaus. Definitionen und Konzepte. Sonderforschungsbereich 614. Bonifatius, Paderborn

Fraunhofer IOF (2006) Produktion von mikrooptischen Bauteilen und Systemen Status und Perspektiven, Fraunhofer Institut angewandte Optik und Feinwerktechnik, Jena

Frieling E, Sträter O (2009) Folgeantrag im Rahmen des Schwerpunktprogramms 1184 "Altersdifferenzierte Arbeitssysteme"

Fries E (1999) Anwendung neuronaler Netze zur Verschleißerkennung beim Fräsen. Dissertation, TU Berlin
Fu MC (2002) Optimization for simulation: theory vs. practice. INFORMS J Comput 14:3
Funke J, Frensch PA (2006) Handbuch der Allgemeinen Psychologie – Kognition. Hogrefe, Göttingen
Garvin DA (1984) What does "product quality" really mean? Sloan Manag Rev 1984(Fall)
Gat E (1998) On three-layer architectures. In: Kortenkamp D, Bonnasso R, Murphy R (eds) Artificial intelligence and mobile robots. AAAI Press, Menlo Park, pp 195–211
Gausemeier J (2005) Von der Mechatronik zur Selbstoptimierung – Herausforderungen an die domänenübergreifende Zusammenarbeit. Produkt Daten J, 1
Gausemeier J, Frank U, Giese H, Klein F, Schmidt A, Steffen D, Tichy M (2005) A design methodology for self-optimizing systems. Automation, Assistance and Embedded Real Time Platforms for Transportation (AAET2005), 16, 2005
Gausemeier J, Ramming FJ, Schäfer W (2006) Selbstoptimierung im Maschinenbau – Auf dem Weg zu den Maschinen von übermorgen. Forschungsforum Paderborn 10(1)
Gausemeier J, Adelt P, Donath J, Geisler J, Henkler S, Kahl S, Klöpper B, Krupp A, Münch E, Oberthür S, Paiz C, Porrmann M, Radkowski R, Romaus C, Schmidt A, Schulz B, Vöcking H, Witkowski U, Witting K, Znamenshchykov O (2009a) Selbstoptimierende Systeme des Maschinenbaus. HNI-Verlagsschriftenreihe
Gausemeier J, Rammig F, Schäfer W (eds) (2009b) Selbstoptimierende Systeme des Maschinenbaus. Definition, Anwendungen, Konzepte. HNI-Verlagsschriftenreihe, vol 234. Westfalia Druck GmbH, Paderborn
Gazzola V, Rizzolatti G, Wicker B, Keysers C (2007) The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. NeuroImage 35:1674–1684
Gehrke J, Ramakrishnan R, Ganti V (2000) Rainforest – a framework for fast decision tree construction of large datasets. Data Min Know Dis 4(2/3):127–162
Gesatzki RP (2002) Elemente von advanced planning und scheduling-systemen zur optimierung der supply chain. SCM Newsletter
Gladkov EA, Maloletkov AV, Perkovski RA (1998) Predicting the quality of TIG butt welded joint using neural network models. Weld Int 3(12):215–219
Gläßner J (1995) Modellgestütztes Controlling der beschaffungslogistischen Prozesskette. VDI, Düsseldorf
Gloy Y-S, Muschong C, Schneuit H, Gries T (2009) Innovative Simulationstools zur Webmaschinenoptimierung. In: Küppers B (ed) Proceedings of the 3rd Aachen-Dresden International Textile Conference, Aachen, November 26–27, 2009, Aachen: DWI an der RWTH Aachen e.V.
Göhner P, Urbano PGA, Wagner T (2004) Softwareagenten – Einführung und Überblick über eine Alternative art der Softwareentwicklung. Teil 3: Agentensysteme in der Automatisierungstechnik. Automatisierungstech Prax 46(2):42–51
Golden RM (1996) Mathematical methods for neural network analysis and design. MIT Press, Cambridge
Gomez P (1978) Die kybernetische Gestaltung des Operations Managements – Eine Systemmethodik zur Entwicklung anpassungsfähiger Organisationsstrukturen. Haupt, Bern
Gorrisen D, De Tommasi L, Crombecq K, Dhaene T (2009) Sequential modeling of a low noise amplifier with neural networks and active learning. Spr Neural Comput Appl 18(5):485–494
Gramann W (1999) Industrielle Montage mikrooptischer Halbleiterbauelemente. In: Reinhart G (ed) Iwb-Seminarberichte Nr. 44: Automatisierte Mikromontage. Utz, München
Grauel A (1992) Neuronale Netze – Grundlagen und mathematische Modellierung. BI Wissenschaftsverlag, Mannheim
Grešovnik I, Rodič T (1999) A general-purpose shell tor solving inverse and optimization problems in material forming. In: Proceedings of the 2 ESAFORM conference on material forming guimararäes, Portugal

Gruber TR (1993) A translation approach to portable ontology specifications. Knowl Acquis 5(2):199–220 (Academic Press, New York)
Günther H, Tempelmeier H (2007) Produktion und Logistik. Springer, Berlin
Gurney K (1997) An Iintroduction to neural networks. UCL Press, London
Haats C (2000) Produktionsplanung und -steuerung für Unternehmen mit linearen Prozessketten. Dissertation Universität Stuttgart, Jost-Verlag, Heinsheim
Haferkamp H (1999) Thermographic system for process monitoring of laser beam cutting. Europto conference on optical measurement systems for industrial inspection, Munich, Germany, SPIE vol 3824, June 1999
Han J, Kamber M (2006) Data mining: concepts and techniques. Kaufmann, San Francisco
Hand D, Mannila H, Smyth P (2001) Principles of data mining. Massachusetts Institute of Technology, USA
Hankes J (1999) Sensoreinsatz in der automatisierten Mikromontage, Fortschritt-Berichte VDI 2 Nr. 459, Düsseldorf, VDI
Hart PE, Nilsson NJ, Raphael B (1968) A formal basis for the heuristic determination of minimum cost paths. IEEE Trans Syst Sci Cybernet SSC 4(2):100–107
Hauck E, Gramatke A, Henning K (2008) Cognitive technical systems in a production environment. In: Proceedings of the fifth international conference on informatics in control, automation and robotics. ICINCO: Madeira, Portugal
Hayes-Roth B (1985) A blackboard architecture for control. Artif Intell 26:251–321 (Elsevier, Oxford)
Heidrich J (2001) Optimierung und Überwachung des MSG-Schweißprozesses mit Hilfe klassischer Modelle und KI-Methoden. Aachener Berichte Fügetechnik, vol 3. Shaker, Aachen
Hellmich K (2002) Wettbewerbsvorteile durch realisierte Kundenorientierung der Wertschöpfung. Ind Manag 18(5):27–30
Hellmich K (2003) Kundenorientierte Auftragsabwicklung – Engpassorientierte Planung und Steuerung des Ressourceneinsatzes. DUV, Wiesbaden
Henrich P (2002) Strategische Gestaltung von Produktionssystemen in der Automobilindustrie. Shaker, Aachen
Herold C (1991) Ein Vorgehenskonzept zur Unternehmensstrukturierung – Eine heuristische Anwendung des Modells lebensfähiger Systeme. Dissertation Universität St. Gallen, Difo-Druck, Bamberg
Hesselbach J, Weule H, Klocke F, Weck M (2002) mikroPRO – Untersuchung zum internationalen Stand der Mikroproduktionstechnik. Essen, Vulkan
Hichri H (2005) Überwachung und Optimierung des MSG-Schweißprozesses mit Hilfe von KIMethoden und Spracherkennungsalgorithmen, vol 2. Aachener Berichte Fügetechnik, Shaker, Aachen
Hill T (2000) Manufacturing strategy. McGraw-Hill, Irwin
Hoffmann JFF (2001) The fast-forward planning system. AI Magaz 22;57–62
Hoffmann J, Brafman R (2005) Contingent planning via heuristic forward search with implicit belief states. In: Proceedings of the 15th international conference on automated planning and scheduling, Monterey, USA
Hoitsch H (1993) Produktionswirtschaft. Grundlagen einer industriellen Betriebswirtschaftslehre. 2., völlig überarbeitete und erweiterte Auflage. Vahlen, München
Homem de Mello LS, Sanderson AC (1986) And/Or graph representation of assembly plans. In: Proceedings of 1986 AAAI national conference on artificial intelligence, pp 1113–1119
Hufgard A (ed) (2005) Business Integration mit SAP-Lösungen – Potenziale, Geschäftsprozesse, Organisation und Einführung. Springer, Berlin
Insperger T, Mann BP, Stépàn G, Bayly PV (2003) Stability of up-milling and down-milling, part 1: alternative analytical methods. Int J Mach Tools Manuf 43(1):25–34
Isermann R (2002) Mechatronische systeme: grundlagen. Springer, New York
Ismail F, Ziaei R (2002) Chatter suppression in five-axis machining of flexible parts. Int J Mach Tools Manuf 42(1):115–122

Jennings NR, Sycara K, Wooldridge M (1998) A roadmap of agent research and development. Auton Agent Multi Agent Syst 1:275–306
Johannaber F, Michaeli W (2004) Handbuch Spritzgießen. Hanser, München
Jones RM, Laird JE, Nielsen PE, Coulter KJ, Kenny P, Koss, FV (1999) Automated intelligent pilots for combat flight simulation. AI Mag 20:27–41
Jorden W (2007) Handbuch für Studium und Praxis. Form- und Lagetoleranzen. Hanser, München
Jurecka F (2007) Robust design optimization based on metamodeling techniques. Dissertation Technische Universität München
Kammel S, Ziegler J, Pitzer B, Werling M, Gindele T, Jagzent D, Schroder J, Thuy M, Goebl M, von Hundelshausen F, Pink O, Frese C, Stiller C (2008) Team AnnieWAY's autonomous system for the 2007 DARPA Urban Challenge. J Field Robotics 25:615–639
Kämpf R (2007) Grundlagen des Produktionsmanagements. In: Gienke H, Kämpf R (eds) Handbuch Produktion: Innovatives Produktionsmanagement: Organisation, Konzepte, Controlling. Hanser, München
Kaplan RS, Norton DP (2006) The balanced scorecard. Translating strategy into action. Harvard Business School Press, Boston
Karim S, Sonenberg L, Tan AH (2006) A hybrid architecture combining reactive, plan execution and reactive learning. In: Proceedings of the 9th Biennial Pacific Rim International Conference on Artificial Intelligence (PRICAI), China
Kaufman S, Wilson R, Jones R, Calton T, Ames A (1996) The Archimedes 2 mechanical assembly planning system. Proc IEEE Int Conf Robotics Automat 4:3361–3368
Kempf T (2010) Ein kognitives Steuerungsframework für robotergestützte Handhabungsaufgaben. Apprimus, Aachen
Kempf T, Herfs W, Brecher C (2008) Cognitive control technology for a self-optimizing robot based assembly cell. In: Proceedings of the ASME 2008 international design engineering technical conferences and computers and information in engineering conference, American Society of Mechanical Engineers, USA
Keuster J De (2005) Methods for monitoring of laser cutting by means of acoustic and photodiode sensors. Adv Mater Res 6–8:809–816
Keuster J De (2007) Real-time adaptive control and optimisation of high-power CO_2 laser cutting using photodiodes. In: Proceedings of the 5th lane conference, laser assisted Net Shape Engineering pages, 2007, pp 979–992
Kiener S (2006) Produktions-Management. Oldenbourg, München
Kim IS, Jeong YJ, Lee CW, Yarlagadda V (2004) Prediction of welding parameters for pipe-line welding using an intelligent system. Int J Adv Manuf Technol 22(9–10):713–719
Kim IS, Son JS, Park CE, Kim IJ, Kim HH (2005) An investigation into an intelligent system for predicting bead geometry in GMA welding process. J Mater Process Technol 159(1):113–118
Kinkel S, Friedewald M, Hüsing B, Lay G, Lindner R (2008) Arbeiten in der Zukunft. Strukturen und Trends der Industriearbeit. Sigma, Berlin
Kinnebrock W (1994) Neuronale netze: Grundlagen, Anwendungen, Beispiele. Oldenbourg, München
Kletti J (2006) MES – Manufacturing Execution System – Moderne Informationstechnologie zur Prozessfähigkeit der Wertschöpfung. Springer, Berlin
Klocke F, Kratz S, Veselovac D, Mtz. de Aramaiona P, Arrazola P (2008) Investigation on forcesensor dynamics and their measurement characteristics. ASME international mechanical engineering congress and exposition. Boston, USA
Klocke F, Veselovac D, Auerbach T, Kamps S (2009) Controller parameter analysis during the milling of aerospace alloys. SAE 2009 aerotech congress und exhibition, Seattle, USA, Nov 2009
Klocke F, Veselovac D, Auerbach T, Kamps S (2010) Kennwertgenerator für die automatisierte Versuchsdurchführung von Zerspanversuchen. In: Jamal R, Heinze R (eds) Virtuelle Instrumente in der Praxis -Begleitband zum 15. VIP-Kongress 2010. Hüthig, Heidelberg, pp 75–81. ISBN 978-3-8007-3235-7

Kobayashi I, Panskus G (eds) 20 Keys® Die 20 Schlüssel zum Erfolg im internationalen Wettbewerb. Adapt-Media-Verlag, Bochum
Kohonen T (1984) Self-organization and assoziative memory. Springer, Berlin
Korb W (2005) Mit Pepp zur wirtschaftlichen Mikromontage, Mikroproduktion 1/2005. Hanser, München, pp 21–22
Kötzle A, Göbel E (1997) Strategisches Management – Theoretische Ansätze, Instrumente und Anwendungskonzepte für Dienstleistungsunternehmen. Lucius und Lucius, Stuttgart
Krause F-L (2007) Innovationspotenziale in der Produktentwicklung. Hanser, München
Kudlik N (1997) Reproduzierbarkeit des Kunststoff-Spritzgießprozesses. Dissertation RWTH Aachen, ISBN: 3-89653-412-4
Kühne T (2005) What is a Model? In: Proceedings of Language Engineering for Model-Driven Software Development, Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI). Dagstuhl, pp 1–10
KUKA Robotics GmbH (2011a) http://www.kuka-robotics.com/germany/de/products/software/hub_technologies/. Stand: March 2011
KUKA Robotics GmbH (2011b) http://www.kuka-omnimove.com/en/ Stand: March 2011
Kurbel K, Endres A (2005) Produktionsplanung und -steuerung: Methodische Grundlagen von PPS-Systemen und Erweiterungen. Oldenbourg, München
Kurth J (2005) Flexible Produktionssysteme durch kooperierende Roboter. wt Werkstatttechnik online 3
Ljung L (1999) System identification. Prentice-Hall, Upper Saddle River
Laird J, Bates Congdon C (2008) The soar user's manual, 9 edn. The Regents of the University of Michigan, USA
Langley P, Cummings K, Shapiro D (2004) Hierarchical skills and cognitive architectures. In: Proceedings of the twenty-sixth annual conference of the cognitive science society. Chicago, IL
Langley P, Laird JE, Rogers S (2009) Cognitive architectures: research issues and challenges. Cogn Syst Res 10:141–160
Larose DT (2005) Discovering knowledge in data. Wiley, New Jersey
Latzke PM, Hesse R (1974) Textilien Untersuchen-Prüfen-Auswerten, Hrsg. Fach-verlag Schiele und Sohn GmbH, pp 123–125. ISBN 3 7949 0228 9
Laue T, Röfer T (2006) Getting upright: migrating concepts and software from four-legged to humanoid soccer robots. In: Pagello E, Zhou C, Menegatti E (eds) Proceedings of the workshop on humanoid soccer robots in conjunction with the 2006. IEEE International Conference on Humanoid Robots, Genoa
Lehman JF, Laird J, Rosenbloom P (2006) A gentle introduction to soar, an architecture for human cognition. Technical Report, University of Michigan, USA
Leiden K, Laughery KR, Keller J, French J, Warwick W, Wood SD (2001) A review of human performancer models for the prediction of human error. Prepared for: national aeronautics and space administration system-wide accident prevention program. Ames Research Center, Moffet Filed, CA
Levine DS (2000) Introduction to neural and cognitive modeling. Psychology Press, New York
Li ZZ, Zhang ZH, Zheng L (2004) Feedrate optimization for variant milling processes based on cutting force prediction. Int J Adv Manuf Technol 24(7/8):541–545
Liao TW, Chen LJ (1994) Neural network approach for grindind processes: modeling and optimization. Int J Mach Tools Manuf 37(7):919–937
Licha A (2003) Flexible Montageautomatisierung zur Komplettmontage flächenhafter Produktstrukturen durch kooperierende Industrieroboter. Meisenbach, Bamberg
Liepmann D, Beauducel A, Brocke B, Amthauer R (2007) I-S-T 2000 R Intelligenz-Struktur-Test 2000 R, 2., erweiterte und überarbeitete Auflage. Hogrefe, Göttingen
Lim TS, Loh WY, Shih YS (1997) An empirical comparison of decision trees and other classification methods. Technical Report 979, Department of Statistics, University of Wisconsin, Madison

Lin AC, Chang TC (1993) An integrated approach to automated assembly planning for three-dimensional mechanical products. Int J Prod Res 31(5):1201–1227 (Taylor & Francis)
Lindemann U (2006) Individualisierte Produkte – Komplexität beherrschen in Entwicklung und Produktion. Springer, Berlin
Litoiu M, Woodside M, Zheng T (2005) Hierarchical model-based autonomic control of software systems. ACM SIGSOFT Softw Eng Notes 30(4)
Lödding H (2008) Verfahren der Fertigungssteuerung – Grundlagen, Beschreibung, Konfiguration. Springer, Berlin
Loeffelholz FV (1991) Qualität von PPS-Systemen – Ein Verfahren zur Analyse des Informationsgehaltes. Dissertation RWTH Aachen, Springer, Berlin
Lopitzsch JR (2005) Segmentierte adaptive Fertigungssteuerung. Dissertation der Universität Hannover. PZH Produktionstechnisches Zentrum GmbH. Garbsen
Lotter B (2006) Montage in der industriellen Produktion. Ein Handbuch für die Praxis. Springer, Berlin
Luczak H, Eversheim W (1999) Produktionsplanung und -steuerung: Grundlagen, Gestaltung und Konzepte. Springer, Berlin
Lunze J (2008) Regelungstechnik 1, 7 edn. Springer, Berlin
Lutz S (2002) Kennliniengestütztes Lagermanagement. VDI, Düsseldorf
Nørgaard M, Ravn O, Poulsen NK, Hansen LK (2000) Neural networks for modelling and control of dynamic systems. Springer, London
Mälck H (2001) Zielsysteme in der Produktion: Das Führen mit Zielen als Grundlage für optimierte Technik-, Organisations- und Personalentwicklungsprozesse. RKW, Eschborn
Malik F (2006) Strategie des Managements komplexer Systeme – Ein Beitrag zur Management-Kybernetik evolutionärer Systeme. Haupt, Bern
Mane I, Gagnol V, Bouzgarrou BC, Ray P (2008) Stability-based spindle speed control during flexible work piece high-speed milling. Int J Mach Tools Manuf 48(2):184–194
Maqbool S, Smith JS, Lucas J (1998) Neural networks for the control of welding systems. 8th international conference on computer technology in welding, Maastricht, pp 125–133
Marshall SP (2008) Cognitive models of tactical decision making. In: Karowski W, Salvendy G (eds) Proceedings of the 2nd international conference on applied human factors and ergonomic (AHFE) 14–17 July 2008. Las Vegas, Nevada
Martin PJ (1994) Artificial neural networks in welding. Join Mater 6(2):62–67
Martin W (2003) Kundenbeziehungsmanagement im Echtzeit-Unternehmen. Strategic Bulletin CRM 2004. itresearch, Sauerlach bei München
Matteson A, Morris R, Tate R (1992) Real-time GMAW quality classification using an artificial neural network with air born acoustic signals as input. In: International conference on computerization of welding information IV, USA, pp 189–197
Mattke F, Berlekamp CG, Queren-Lieth W (1993) Die Darstellung von schweißtechnischem Wissen in Expertensystemen. DVS-Berichte, vol 156, DVS-Verlag Düsseldorf, pp 147–150
Mayer M, Odenthal B, Grandt M, Schlick C (2008) Anforderungen an die benutzerzentrierte Gestaltung einer Kognitiven Steuerung für Selbstoptimierende Produktionssysteme. In: Gesellschaft für Arbeitswissenschaft eV (ed) Produkt- und Produktions-Ergonomie – Aufgabe für Entwickler und Planer. GfA-Press, Dortmund
Mayer M, Odenthal B, Faber M, Kabuß W, Kausch B, Schlick C (2009) Simulation of human cognition in self-optimizing assembly systems. In: Proceedings of 17th world congress on ergonomics IEA 2009, Beijing
McCarthy J (1959) Programs with common sense. In: Proceedings of the teddington conference on the mechanization of thought processes, Her Majesty's Stationery Office London, pp 756–791
Menges G, Stitz S, Vargel J (1971) Grundlagen der Prozeßsteuerung beim Spritzgießen. In: Kunststoffe, Februar, pp 74–80
Menges G, Thienel P (1977) Pressure-specific volume-temperature behavior of thermoplastics under normal processing conditions. Polymer Eng Sci 17(10):758–763

Merget M (2004) Kostenoptimierung durch Toleranzvariation im Simultaneous Engineering. Shaker, Aachen
Meyer D, Steil T, Müller S (2005) Shared Augmented Reality zur Unterstützung mehrerer Benutzer bei kooperativen Montagearbeiten im verdeckten Bereich. In: Kuhlen T, Kobbelt L, Müller S (eds) Virtuelle und Erweiterte Realität. Shaker, Aachen
Meyer M (2006) Logistisches Störungsmanagement in Kundenverbrauchsorientierten Wertschöpfungsketten. Dissertation RWTH Aachen, Shaker, Aachen
Michaeli W (2006) Einführung in die Kunststoffverarbeitung. Hanser, München
Michaeli W, Lauterbach M (1989) Die pmT-Optimierung "Konsequenzen aus dem pvT-Konzept zur Nachdruckführung". Kunststoffe, Sept 1989, pp 852–855
Michaeli W, Schreiber A (2010) Der Weg zum geregelten Prozess. Swiss Plastics, Nov, pp 21–26, Aarau, AZ Fachverlage AG
Michaeli W, Breuer P, Hohenauer K, Von Oepen R, Philipp M, Pötsch G, Recker H, Robers T, Vaculik R (1992) Qualitätsgesichertes Spritzgießen. Kunststoffe, Dezember, pp 1167–1171
Michaeli W, Hopmann C, Gruber J (2004) Prozessregelung beim Spritzgießen. Kunststoffe, Januar, pp 20–24
Michie D, Spiegelhalter DJ, Taylor CC (1994) Machine learning: neural and statistical classification. Horwood, London
Middle JE, Li J (1993) Artificial neural networks applied to process modelling for robotic arc welding, modelling of casting, welding and advanced solidification processes, 6th international conference on modelling and welding process, USA, pp 127–134
Mintzberg H, Quinn BJ, Ghoshal S (2004) The strategy process. Prentice-Hall, Hemel Hempstead
Mitsuishi M, Ueda K, Kimura F (2004) Manufacturing systems and technologies for the new frontier, 41st CIRP conference
Möller N (2008) Bestimmung der Wirtschaftlichkeit wandlungsfähiger Produktionssysteme, Forschungsbericht N IWB, vol 212. Utz, München
Müller B, Reinhardt J (1995) Neural networks. An introduction. Springer, Berlin
Müller R, Esser M, Janßen C, Brecher C, Pyschny N (2009a) Flexibel automatisierte Montagesysteme – Toleranzoptimierte Montage von miniaturisierten Produkten. Proceedings der VDITagung Mechatronik 2009, Wiesloch, 12–13. Mai
Müller R, Buchner T, Gottschalk S, Fayzullin K, Herfs W, Hilchner R, Pyschny N (2009b) Montagetechnik und – organisation. Apprimus, Aachen
Müller R, Corves B, Hüsing M, Esser M, Riedel M, Vette M (2010a) Reconfigurable self-optimising handling system. In: Fifth international precision assembly seminar IPAS 2010, Chamonix
Müller R, Esser M, Janßen C, Vette M (2010b) System-Identifikation für Montagezellen – Erhöhte Genauigkeit und bedarfsgerechte Rekonfiguration. In: wt Werkstattstechnik online 100, 9, S 687–691. ISSN 1436-4980, Internet: www.werkstattstechnik.de. Springer-VDIVerlag, Düsseldorf
Murthy SK (1995) On growing better decision trees from data. Ph. D. thesis, Department of Computer Science, Johns Hopkins University, Baltimore
Nason S, Laird J (2005) Soar-RL: integrating reinforcement learning with soar. In: Cognitive systems research, special issue of cognitive systems research – the best papers from ICCM2004
Nauck D, Klawonn F, Kruse R (2002) Neuronale Netze und Fuzzy Systeme. Vieweg, Wiesbaden
Niemann G, Winter H, Höhn BR (2005) Maschinenelemente, Band 1: Konstruktion und Berechnung von Verbindungen, Lagern, Wellen. Springer, Berlin
Nise N (2009) Control systems engineering, 4 edn. Wiley, Hoboken
Normeey-Rico JE, Camacho EF (2007) Control of dead-time processes. Springer, London
Nyhuis P (2008) Produktionskennlinien: Grundlagen und Anwendungsmöglichkeiten. In: Nyhuis P (ed) Beiträge zu einer Theorie der Logistik. Springer, Berlin
Nyhuis P, Wiendahl H (2003) Logistische Kennlinien – Grundlagen, Werkzeuge und Anwendungen. Springer, Berlin

Nyhuis P, Heinen T, Reinhart G, Rimpau C, Abele E, Wörn A (2008) Wandlungsfähige Produktionssysteme – Theoretischer Hintergrund zur Wandlungsfähigkeit von Pproduktionssystemen, wt Werkstatttechnik Online pp 1–2
o.V. (2006) Gibt es eine verschwenundungsfreie Logistik? In: 3. Lean management summit. Aachener Management Tage, Aachen, pp 29–45
Odenthal B, Mayer M, Grandt M, Schlick C (2007) Concept of an adaptive training system for production. In: Hinneburg A (ed) Lernen – Wissen – Adaption, Workshop Proceedings. Martin- Luther-University, Halle-Wittenberg
Odenthal B, Mayer M, Grandt M, Schlick C (2008) Konzept eines Lehr-/Lernsystems einer kognitiven Steuerung für Selbstoptimierende Produktionssysteme. In: Gesellschaft für Arbeitswissenschaft e. V (ed) Produkt- und Produktions-Ergonomie – Aufgabe für Entwickler und Planer. GfA-Press, Dortmund
Odenthal B, Mayer M, Kabuß W, Kausch B, Schlick C (2009) Error detection in an assembly object using an augmented vision system. In: Proceedings of 17th world congress on ergonomics IEA 2009, Beijing
Odenthal B, Mayer M, Kabuß W, Kausch B, Schlick C (2011) An empirical study of disassembling using an augmented vision system. In: Proceedings of the HCII 2011
Ohno T (1993) Das Toyota-Produktionssystem. Campus, New York
Ong SK, Yuan ML, Nee AY (2008) Augmented reality in manufacturing: a survey. Int J Prod Res 46(10):2702–2742
Onken R, Schulte A (2010) System ergonomic design of automation. Dual-mode cognitive design of vehicle guidance and control work systems. Springer, Berlin
Orlicky J, Plossl G (1994) Orlicky's material requirements planning. McGraw-Hill, New York
Paetzold K (2006) On the importance of a functional description for the development of cognitive technical systems. In: International design conference 2006, Dubrovnik
Palani S (2010) Control systems engineering. McGraw-Hill, New Delhi
Peters A (2001) Benutzerorientierte Leitsoftware für automatisierte Fertigungssysteme in Komponentenstruktur. Dissertation RWTH Aachen
Petersen B (2003) Flexible Handhabungstechnik für die automatisierte Mikromontage. Aachen, Shaker
Pfeifer T, Schmitt R (eds) (2006) Autonome Produktionszellen: Komplexe Produktionsprozesse flexibel automatisieren. Springer, Berlin
Pfeifer T, Tillmann M, Wimmer M (2003) Prozesskettenoptimierung – Mit der IPO-Systematik innovativ zu ganzheitlichen Lösungen. In: REFA-Nachrichten 6/2003
Pfendler C, Schlick C (2007) A comparative study of mobile map displays in a geographic orientation task. Behav Inf Technol 26(6):455–463
Pfirrmann O, Astor M (2006) Trendreport MST 2020. Basel, Berlin
Pfohl H (2004) Logistiksysteme: Betriebswirtschaftliche Grundlagen. Springer, Berlin
Photonics 21 (2010) Lighting the way ahead. Düsseldorf
Porter ME (2004) Competitive strategy. Free Press, New York
Posieek S (2005) Einsatz kombinatorischer Optimierungsmethoden bei automatischer Optimierung von Umformprozessen. Shaker, Aachen
Possel-Dölken F (2006) Projektierbares Multiagentensystem für die Ablaufsteuerung in der flexibel automatisierten Fertigung. Shaker, Aachen
Pramujati B, Dubay R, Samaan C (2006) Cavity pressure control during cooling in plastic injection molding. Adv Polym Technol 25(3):170–181
Premium AEROTEC GmbH: http://www.premium-aerotec.com/Binaries/Binary5687/A350-Grafik_work_packages_DE.jpg
Priddy KL, Keller PE (2005) Artificial neural networks. SPIE – The International Society for Optical Engineering, Bellingham
Puppe F, Gappa U, Poeck K, Bamberger S (1996) Wissensbasierte Diagnose- und Informationssysteme Mit Anwendungen des Expertensystem-Shell-Baukasten D3. Springer, Berlin

Putzer HJ (2004) Ein uniformer Architekturansatz für Kognitive Systeme und seine Umsetzung in ein operatives Framework. Verlag Dr. Köster, Berlin
Rasmussen J (1986) Information processing and human-machine interaction. An approach to cognitive engineering. North-Holland, New York
Rehse M (1999) Flexible Prozessüberwachung bei der Bohr- und Fräsbearbeitung in einer autonomen Produktionszelle. RWTH Aachen, Dissertation
Reinhart G, Berlak J, Effert C, Selke C (2002) Wandlungsfähige Fabrikgestaltung, wirtschaftlichen Fabrikbetrieb 97(1/2):18–23
Reisgen U, Beckers M, Willms K, Buchholz G (2010) Einsatz und Vorgehensweise bei der Ersatzmodellierung beim Impulslichtbogenschweißverfahren. VS-Berichte 268:79–84
Reuber M (2001) Prozessüberwachung beim Schlichtfräsen von Freiformflächen. Dissertation RWTH Aachen
Riedel M, Nefzi M, Corves B (2010a) Grasp planning for a reconfigurable parallel robot with an underactuated arm structure. In: Proceedings of the 1st international workshop on underactuated grasping. Montreal, Canada, 19. Aug
Riedel M, Nefzi M, Corves B (2010b) Performance Analysis and Dimensional Synthesis of a Six DOF Reconfigurable Parallel Manipulator. In: Proceedings of the IFToMM symposium on mechanism design for robotics. Mexico City, Mexico, 28.–30. Sept , 2010
Rojas R (1996) Neural networks. Springer, Berlin
Romero V, Burkardt J, Gunzburger M, Peterson J (2006) Comparison of pure and "Latinized" centroidal Voronoi tessellation against various other statistical sampling methods. J Reliabil Eng Syst Safety 91(10/11):1266–1280
Roosen S (1997) Online-Prozessoptimierung beim MAG-Schweißen mit Hilfe eines Expertensystems, vol 4. Aachener Berichte Fügetechnik, Shaker
Rosenblatt F (1958) The perceptron. A probabilistic model for information storage and organization in the brain. Psychol Rev 65:386–408
Rüegg-Stürm J (2003) Das neue St. Galler Management-Modell: Grundkategorien einer integrierten Managementlehre; der HSG-Ansatz. Haupt, Bern
Rummery GA, Niranjan N (1994) On-line Q-learning using connectionist systems, Technical Report CUED/F-INFENG/TR 166, Cambridge University, GB
Russell SJ, Norvig P (2003) Artificial intelligence: a modern approach. Pearson Education, Upper Saddle River
Russer H, León FP (2007) Informationsfusion – eine übersicht. Tech Mess 74(3):93–102
Salomon P (2006) Micro sensors and microsystems. Eurosensors Göteborg, Sept 2006
Sauer W (2003) Prozesstechnologie der Elektronik, Modellierung, Simulation und Optimierung der Fertigung. Hanser, München
Savory S (1985) Künstliche Intelligenz und Expertensysteme. Oldenburg, München
Scherms E (1998) Optimierte Tolerierung durch Qualitätsdatenanalyse. Konstruktion 50: 31–36
Schild K, Bussmann S (2007) Self-organization in manufacturing operations. Commun ACM 50(12):74–79
Schlick C (1999) Modellbasierte Gestaltung der Benutzungsschnittstelle autonomer Produktionszellen. In: Luczak H (ed) Schriftenreihe Rationalisierung und Humanisierung, vol 61 (zugl. Dissertation RWTH Aachen). Shaker, Aachen
Schlick C, Odenthal B, Mayer M, Neuhöfer J, Grandt M, Kausch B, Mütze-Niewöhner S (2009) Design and evaluation of an augmented vision system for self-optimizing assembly cells. In: Schlick C (ed) Industrial engineering and ergonomics. Springer, Berlin
Schlick C, Mayer M, Odenthal, B (2010) MTM als Prozesslogik für die kognitiv automatisierte Montage. In: Britzke B (ed) MTM in einer globalisierten Wirtschaft – Arbeitsprozesse systematisch gestalten und optimieren. Wirtschaftsbuch, München
Schmidt C (2008) Konfiguration überbetrieblicher Koordinationsprozesse in der Auftragsabwicklung des Maschinen- und Anlagenbaus. Dissertation RWTH Aachen, Shaker, Aachen

Schmitt R (2000) Aufbau flexibler Mess- und Prüfstationen für die automatisierte Montage. Shaker, Aachen
Schmitt R (2010) Qualitätsmanagement. Strategien, Methoden, Techniken, 4 edn. Hanser, München
Schmitt R, Beaujean P (2007) Selbstoptimierende Produktionssysteme. Zeitschrift für wirtschaftlichen Fabrikbetrieb 9:520–524
Schmitt R, Betzold M, Hense K (2007) Das Aachener Qualitätsmanagementmodell. In: Schmitt R, Pfeifer T (eds) Masing Handbuch Qualitätsmanagement. Hanser, München, pp 38–41
Schmitt R, Pavim A, Brecher C, Pyschny N, Loosen, P, Funck, M, Dolkemeyer J, Morasch, V (2008) Flexibel automatisierte Montage von Festkörperlasern: Auf dem Weg zur flexiblen Montage mittels kooperierender Roboter und Sensorfusion. In: wt – Werkstatttechnik, Ausgabe 11/12, 2008, Jahrgang 98, pp 955–960
Schmitt R, Stiller S, Beaujean P (2009a) The aachen quality management model – normative framework of the entrepreneurial quality philosophy
Schmitt R, Wagels C, Isermann M, Matuschek N (2009b) Cognitive optimization of an automotive rear-axle drive production process. J Mach Eng 9(1):78–90
Schneider F (2004) Überwachung, Regelung und Automatisierung beim Hochgeschwindigkeitsschneiden von Elektroblechen mit Laserstrahlung. Dissertation RWTH Aachen
Scholz-Reiter B, de Beer C, Bose F, Windt K (2007) Evolution in der Logistik Selbststeuerung logistischer Prozesse. In: VDI Berichte
Schotten M (1998) Beurteilung von EDV-gestützten Koordinationsinstrumentarien in der Fertigung. Dissertation RWTH Aachen, Shaker, Aachen
Schuh G (2005) Produktkomplexität managen: Sstrategien – Methoden – Tools. Hanser, München
Schuh G (2006) Produktionsplanung und -steuerung. Springer, Berlin
Schuh G, Orilski S (2007) Roadmapping for competitiveness of high wage countries. In: Proceedings of the XVIII. ISPIM conference. Warschau, Polen
Schuh G, Roesgen R (2006) Aufgaben. In: Schuh G (ed) Produktionsplanung und -steuerung. Springer, Berlin
Schulz W, Kostrykin V, Zefferer H, Petring D, Poprawe R (1997) A free boundary problem related to laser beam fusion cutting: ODE approximation. Int J Heat Mass Trans 40(12):2913–2928
Schulz W, Niessen M, Eppelt U, Kowalick K (2009) Simulation of laser cutting. In: Dowden JM (ed) The theory of laser materials processing: heat and mass transfer in modern technology. Springer Series in Materials Science, pp 21–69
Schulz W, Nießen M, Vossen G, Schüttler J (2010) Modelling and process monitoring of laser processing. In: Erscheint in proceedings of NUSIM multiphysics 2010 (und in The International Journal of Multiphysics)
Schüttler J, Lose J, Schmitt R, Schulz W (2009) Exploring process domain boundaries of complex production processes using a metamodeling approach, 12th CIRP conference on modelling of machining operations, Donostia-San Sebastiàn – Spain, vol 2. In: Arrazola P (ed) Conference on modelling of machining operations, vol II, Mondragon, 7.–8. Mai 2009, pp 835–841
Schwindt C, Trautmann N (2004) Advanced- planning- systeme zum supply – chain- management, Universität Karlsruhe, Karlsruhe
Shafer J, Agrawal R, Mehta M (2005) SPRINT: a scalable parallel classifier for data mining. In: Readings in database systems. Massachusetts Institute of Technology, USA, pp 668–679
Shepard RN, Metzler J (1971) Mental rotation of three-dimensional objects. Science 171(972):701–703
Sheridan T (2002) Humans and automation: system design and reserch issues. Wiley, New York
Shingo S (1992) Das Erfolgsgeheimnis der Toyota-Produktion: eine Studie über das Toyota-Produktionssystem – genannt die “Schlanke Produktion”. Verlag Moderne Industrie, Landsberg/Lech
Siciliano B, Khatib O (eds) (2008) Springer handbook of robotics. Springer, Berlin
Siegert H-J, Bocionek S (1996) Robotik: Programmierung Intelligenter Roboter. Springer, Berlin
Smud SM, Harper DO, Deshpande PB, Leffew KW (1991) Advanced process control for injection molding. Polymer Eng Sci 31(15):1081–1085

Spath D, Scholtz O (2007) Ideen gegen Verlagerung der Montage ins Ausland. Werkstattstechnik 97(1/2):2–4

Spath D, Klinkel S, Barrho T (2002) Auftragsabwicklung in dezentralen Strukturen: Erfolgsfaktoren und Probleme – Ergebnisse einer Studie. Zeitschrift für Wirtschaftlichen Fabrikbetrieb 97(3)

Steinberg DV, Lin DKJ (2006) A construction method for orthogonal Latin hypercube designs. Biometrika 93(2):279–288

Stepanek P (2007) Flexibel automatisierte Montage von leicht verformbaren großvolumigen Bauteilen. Shaker, Aachen

Stoddard K, Kneifel W, Martin D, Mirza K, Chafee M, Hagenauer A, Graf S (2004) Method and control system for controlling a plurality of robots. Patent No.: US 6,804,580 B1 (12 Oct 2004)

Stork A, Thole CA, Klimenko S, Nikitin I, Nikitina L, Astakhov Y (2007) Simulated reality in automotive design. International conference on cyberworlds, pp 23–27

Strohner H (1995) Kognitive Systeme. Eine Einführung in die Kognitionswissenschaft. Westdeutscher, Opladen

Sutton RS, Barto AG (1999) Reinforcement learning, Journal of Cognitive Neuroscience, Massachusetts Institute of Technology, USA

Taguchi G (1990) Quality Engineering

Tai YF, Scherfler C, Brooks DJ, Sawamoto N, Castiello U (2004) The human premotor cortex is “mirror” only for biological actions. Curr Biol 14:117–120

Takeda H (2009) QiP – Qualität im Prozess: Leitfaden zur Qualitätssteigerung in der Produktion. Finanz Buch Verlag GmbH, München

Tang A, Owen C, Biocca F, Weimin M (2004) Performance evaluation of augmented reality for directed assembly. In: Ong SK, Nee AYC (eds) Virtual and augmented reality applications in manufacturing. Springer, New York

Thiele M (1997) Kernkompetenzorientierte Unternehmensstrukturen – Ansätze zur Neugestaltung von Geschäftsbereichsorganisationen. DUV, Wiesbaden

Thiem I (1998) Ein Strukturmodell des Fertigungsmanagements. Dissertation Universität Bochum, Shaker, Aachen

Thomas U (2008) Automatisierte Programmierung von Robotern für Montageaufgaben. Shaker, Braunschweig

Thornton AC (2004) Variation risk management. Focusing quality improvements in product development and production. Wiley, Hoboken

Thrun S, Montemerlo M, Dahlkamp H, Stavens A, Aron J, Diebel P, Fong J, Gale M, Halpenny G, Hoffmann K, Lau C, Oakley M, Palatucci V, Pratt P, Stang S, Strohband C, Dupont LE, Jendrossek C, Koelen C, Markey C, Rummel J, van Niekerk E, Jensen P, Alessandrini G, Bradski B, Davies S, Ettinger A, Kaehler A, Nefian A, Mahoney P (2006) The robot that won the DARPA Grand Challenge. J Field Robot 23(9):661–692

Trumpold H, Beck CH, Richter G (1997) Toleranzsysteme und Toleranzdesign. Qualität im Austauschbau. Hanser, München

Unbehauen H (2008) Regelungstechnik I: Klassische Verfahren zur Analyse und Synthese linearer kontinuierlicher Regelsysteme, 15 edn. Vieweg+Teubner, Wiesbaden

Upton D (1992) A flexible structure for computer-controlled manufacturing systems. Manuf Rev 5(1):58–74

Urmson C, Anhalt J, Bae H, Bagnell JAD, Baker C, Bittner RE, Brown T, Clark MN, Darms M, Demitrish D, Dolan J, Duggins D, Ferguson D, Galatali T, Geyer C, M, Gittleman M, Harbaugh S, Hebert M, Howard T, Kolski S, Likhachev M, Litkouhi B, Kelly A, McNaughton M, Miller N, Nickolaou J, Peterson K, Pilnick B, Rajkumar R, Rybski P, Sadekar V, Salesky B, Seo YW, Singh S, Snider JM, Struble JC, Stentz AT, Taylor M, Whittaker WRL, Wolkowicki Z, Zhang W, Ziglar J (2008) Autonomous driving in urban environments: boss and the urban challenge. J Field Robot – Special Issue on the 2007 DARPA Urban Challenge 25:425–466

Völklein F, Zetterer T (2006) Mikrosystemtechnik. Vieweg und Sohn, Wiesbaden

Vossen G, Schüttler J (2010) Modelling and stability analysis for laser cutting. Submitted to Mathematical and Computer Modelling od Dynamical Systems

Vossen G, Nießen M, Schüttler J (2010) Optimization of partial differential equations for minimizing the roughness of laser cutting surfaces. In: Diehl M, Glineur F, Michiels W (eds) Recent advances in optimization and its applications in engineering. Springer, New York

Weck M (1998) Maschinenarten und Anwendungsbereiche, Werkzeugmaschinen Fertigungssysteme, vol 1. Springer, Berlin, p 20

Weck M, Brecher C (2006) Werkzeugmaschinen 4 – Automatisierung von Maschinen und Anlagen. 6 edn. Springer, Berlin

Weinmann A (1995) Regelungen II: Analyse und technischer Entwurf, vol 2. Springer, Wien

Westkämper E, Zahn E (2009) Wandlungsfähige Produktionsunternehmen – das Stuttgarter Unternehmensmodell. Springer, Berlin

Westkämper E, Bullinger H-J, Horvàth P, Zahn E (eds) (2001) Montageplanung – effizient und marktgerecht, Stuttgart. Springer, New York

Whitney DE (2004) Mechanical assemblies. Their design, manufacture, and role in product development. Oxford University Press, New York

Wibbeler J (2002) Frequenzselektive Vibrationssensoren mit spannungsgesteuerter Resonanzabstimmung in Oberflächenmikromechanik. Dissertation Technische Universität Chemnitz

Wiendahl H (2002) Situative Konfiguration des Auftragsmanagements im turbulenten Umfeld. Dissertation an der Universität Stuttgart, Stuttgart

Wiendahl H (2005) Betriebsorganisation für Ingenieure. Hanser, München

Wiendahl H (2006) Methodische grundlagen. In: Schuh G, Westkämper E (eds) Liefertreue im Maschinen-und Anlagenbau: Stand-Potenziale-Trends. Aachen und Stuttgart

Wiendahl H, Hernàndez R (2006) The transformable factory – strategies, methods and examples. In: Dashchenko AI (ed) Reconfigurable manufacturing systems and transformable factories. Springer, Berlin

Wiendahl H, ElMaraghy H, Nyhuis P, Zäh M, Duffie N, Brieke M (2007) Changeable manufacturing – classification, design and operation. CIRP Ann Manuf Technol 56(2)

Wiest U (2001) Kinematische Kalibrierung von Industrierobotern. Dissertation Universität Karlsruhe, Aachen, Shaker

Wight OW (1984) Production and inventory management in the computer age. Reinhold, New York

Witt S (2007) Integrierte Simulation von Maschine, Werkstück und spanendem Fertigungsprozess. Dissertation RWTH Aachen

Woods S, Pflanz T (2008) DPSS lasers rival fiber lasers for marking applications. LaserFocusWorld, Januar

von Witzleben A (2006) Workshop des Thüringer Ministeriums für Wirtschaft, Technologie und Arbeit Entwicklungsschwerpunkte Thüringens im Bereich Optische Technologien, 17. Juli 2006, am Fraunhofer IOF Jena

Wollnack J, Stepanek P (2004) Formkorrektur und Lageführung für eine flexible und automatisierte Großbauteilmontage. Werkstattstechnik Jahrgang 94(9):414–421

Womack JP, Jones DT (1996) Lean thinking. Simon und Schuster, New York

Yu K-W (2001) Terminkennlinie: Eine Beschreibungsmethodik für die Terminabweichung im Produktionsbereich. VDI, Düsseldorf

Zäh M (2005) A holistic framework for enhancing the changeability of production systems. In: 1st conference on changeable, agile and virtual procduction (CARV 2005b), München

Zaeh M, Wiesbeck M (2008) A model for adaptively generating assembly instructions using state-based graphs. In: Manufacturing systems and technologies for the new frontier. Springer, London

Zaeh MF, Beetz M, Shea K, Reinhart G, Bender K, Lau C, Ostgathe M, Vogl W, Wiesbeck M, Engelhard M, Ertelt C, Ruehr T, Friedrich M, Herle S (2009) The cognitive factory. In: El-Maraghy HA (ed) Changeable and reconfigurable manufacturing systems. Springer, Berlin

Zakharian S, Ladewig-Riebler P, Thoer S (1998) Neuronale Netze für Ingenieure. Vieweg Verlagsgesellschaft

Zäpfel G (2001) Grundzüge des Produktions- und Logistikmanagement. Oldenbourg, München

Zell A (1994) SNNS: stuttgart neural network simulator, Universität Stuttgart, Stuttgart
Zell A, Mache N, Sommer T, Korb T (1991) Design of the SNNS neural network simulator. Österreichische Artificial Intelligence Tagung, pp 93–102
Zheng HY (1990) An experimental study of the relationship between in-process signals and cut quality in gas assisted laser cutting. SPIE – CO_2 Lasers and Applications II 1276:218

http://www.springer.com/978-3-642-21066-2

Integrative Production Technology for High-Wage
Countries
Brecher, C. (Ed.)
2012, XXV, 1096 p., Hardcover
ISBN: 978-3-642-21066-2