

2 Pillars of artificial intelligence

The major goal of this elaboration is to work out how specific aspects of the human mind, namely those responsible for higher cognitive functions, function, and to figure out which hardware is required to process an artificial mind with the same, similar or superior capabilities. For this, I will focus on any approach that allows us to reproduce cognitive capabilities, but not necessarily achieve this target by the same means as evolution did. It makes sense, at this point, to provide an introduction to artificial intelligence in order to understand which areas of the neocortex are subject to AI research and development. This chapter notabene provides only a brief overview of the pillars of AI, as even the detailed elaboration of just one sub-area of each of those would suffice to fill books. In the later chapters I will mostly focus on artificial neural networks, logic, knowledge representation, and speech in order to illustrate how I think the human thought processes can be rebuilt artificially. I will also give a brief introduction to one of my research fields, autonomously acting cars, which should help to understand what it takes to create intelligent, autonomous, social and adaptive agents; rebuilding collective and intelligent behavior in artificial systems not only allows us to understand a significant amount of brain evolution, but also how intelligence makes us a highly complex species in terms of thought processes.

An early definition of artificial intelligence on the part of the IEEE Neural Networks Council was 'the study of how to make computers do things at which, at the moment, people are better'²⁹. This is still valid today, but the research also focused on letting software do things better, in which computers have always been better, as the analysis of large data sets. Data forms the basis for the development of artificial intelligent software systems that will not only collect information, but is able to

- learn,
- understand and interpret information,
- adapt its behavior,
- plan,
- conclude,
- solve problems,
- think abstract,
- come up with ideas, and
- understand and interpret language.

After having worked for about 10 years in the field, I feel more comfortable with the more general definition:

Artificial intelligence is a machine based form of intelligence that is not distinguished from actual natural intelligence, covering amongst others:

- cognitive intelligence
- emotional intelligence
- social intelligence
- tactic knowledge (everyday-life like)
- collective intelligence (swarm)

29 Rich E., Knight K. (1990): Artificial Intelligence, 5

I defined both in separate publications (the latter still being subject to secrecy obligations), but a quick literature review will show that both definitions have been published in similar ways by numerous other researchers. Generally, I pursue the holistic approach towards artificial intelligence, which is concerned with incorporating all or at least several pillars of AI. People very often talk about AI in scenarios where they applied machine learning or some other sophisticated approaches to solve a specific problem, i.e. market prediction. This is not what I consider to be AI – it's just an algorithm performing calculations. For me, there exist five pillars of AI, which is of course subject to discussion and sometimes also to dispute.

2.1 Machine learning

On the coarsest level algorithms in machine learning (ML) can be divided into two classes – supervised and unsupervised – depending on whether the respective algorithm requires the specification of a target variable or not.

2.1.1 *Supervised learning algorithms*

Supervised learning techniques require, in addition to the input variables (predictors), the known target values (labels) of a problem. To train a ML model on the identification of road signs on cameras, vectorized road signs are required as input variables images of, ideally in different configurations. Lighting, angle, pollution, etc. are summarized as noise or blur in the data, nonetheless a street sign in the rain should be equally recognized as good in the sunshine. The labels (the correct names) are typically assigned manually. This correct set of input variables and their correct classification account for a training data set. Although we only have one image per training data set in this case, we are talking about several input variables, since ML algorithms find relevant properties (features) in the training data and learn the relationship between these features and the class assignment for the example cited classification task. Supervised learning is mainly used to predict numeric values (Regression) and classification (prediction of class membership), where the data are not limited to a specific format; the processing of images, audio files, videos, numeric data and text provides for ML algorithms no problem. Some examples of classification are about object recognition (street signs, object ahead of the vehicle, etc.), face detection, credit risk assessment, voice recognition, customer churn to just to name a few.

Examples for regression are the determination of continuous, numerical values on several (sometimes hundreds or thousands) input variables, such as the optimum speed based on road and environmental conditions in an autonomous series vehicle, the determination of a financial measure such as gross domestic product based on a variable number of input variables (using agricultural land, education level of a population, industrial production, etc.) or the identification of potential market share by introducing new models. Each of these issues is highly complex and cannot be expressed by simple, linear relationships in simple equations; it may also happen that the expertise does not exist.

2.1.2 *Unsupervised Learning Algorithms*

Unsupervised learning techniques single out any individual target variable, but it is the aim to achieve a general characterization of an instance. Frequently, unsupervised ML algorithms are

used to group records (clusters), so relationships between individual data points are found, which can be composed of an arbitrarily high number of attributes, and merged into groups (clusters). In some cases, the output of unsupervised ML algorithms can be used again as an input for supervised learning algorithms. Examples of unsupervised learning are about the formation of groups of customers based on purchase behavior or demographics, or time series clustering in order to group millions of sensor time series into not previously obvious groups.

Machine Learning (ML) is therefore the part of the artificial intelligence (AI), in which computers are given the ability to learn from data without being explicitly programmed. In ML, the focus is on the development of programs, which are capable of teaching themselves in order to grow and change as they are provided with new data. Processes that can be mapped as a flow chart are therefore not candidates for machine learning – everything requiring dynamic, changing strategies and that cannot be limited to static rules, is potentially suitable to be solved by means of ML. It is applied when

- human expertise does not exist,
- people cannot express their expertise,
- the solution changes over time,
- the solution has to be adapted to specific cases.

In contrast to the statistics, which is used to track the inference from a sample, it is of interest to develop efficient algorithms for solving optimization problems and a representation of the model for evaluation of inference in computer science. Commonly used methods for optimization known in this context are evolutionary algorithms (genetic algorithms, evolution strategies), whose basic principles are based on the natural evolution.³⁰ These methods are very effective in the application to complex non-linear optimization problems.

ML is not the same as data mining, although ML is generally applied in this field – the goal of both is to analyze data in order to find patterns. Rather than to extract data for human understanding as is the case in the data mining, ML method can be used to improve one's understanding of a program on the data provided. Software implementing ML methods detects patterns in data, and can adapt its behavior based on what it finds. When talking about autonomously acting vehicles (or the software that interprets visual signals from divers input channels, i.e. a camera), an execution block, consisting of the output of the AI algorithm and logic on this output, initiating a braking maneuver in case of suddenly appearing passersby in front of the vehicle needs to function both with small, large, thick, thin, disguised, coming from the left, coming from the right, etc. people, which is called generalization. The term generalization is used here, as the ML-approach must be able to generalize on all persons from the limited amount of training examples provided. It must not, on the other hand, be too generalizing, as the vehicle should not slow down at a stationary dustbin on the roadside.

The complexity in the world will often be greater than the complexity of a ML-model which is tried in most cases, to divide into sub-problems and problems apply ML models on these sub-problems. The output of these models is then integrated to allow complex tasks such as autonomous acting of vehicles in structured and unstructured environments.

30 Bäck T., Fogel D.B., Michalewicz Z. (1997): Handbook of Evolutionary Computation, Institute of Physics Publishing, New York

2.2 Computer Vision

Computer Vision (CV) is a very broad field of research in which scientific theories from different fields (as so often in the AI) are merged, ranging from biology, neuroscience and psychology to computer science, mathematics and physics. To start with it is important to understand how an image is physically created. Before light impinges on a two-dimensional array of sensors, it is refracted, absorbed, scattered or reflected, and an image is formed through the measurement of the intensity of the light beams through each element in the image (pixels). The three basic perspectives on CV are:

- The reconstruction of a scene and the point, from which the scene is observed, based on an image, an image sequence or a film.
- The imitation of biological, visual perception, in order to better understand the physical and biological processes are involved, such as how the wetware works, and how the interpretation and understanding of what has been observed function.
- In technical research and development, the focus is on efficient algorithmic solutions – in CV software often problem-specific solutions are developed, which have only limited overlaps with the optical perception in biological organisms.

All three have similarities and influence each other. If emphasis is on obstacle detection in order to initiate an automated braking maneuver ahead of the vehicle as a passersby appears, it is primarily important to recognize the passer as an obstacle; an interpretation of the entire scene, such as to understand that the vehicle moves toward a family having a picnic in a meadow, is not necessary required in this case. In contrast, the understanding of a scene is then a requirement if context is a relevant input (or output), such as the development of robotic household assistants, who need to understand very well that a person lying on the floor is not only a sleeping obstacle, but is in all likelihood a medical emergency.

Visual perception in a biological organism is understood as an active process, which includes the control of the sensor and is closely linked to the successful completion of the action.³¹ As a result,³² CV-systems are mostly not passive, that is, the system has to

- be continuously supplied with data via sensors (streaming), and
- act on this data stream.

Apart from that the goal of CV systems is not the understanding of scenes in images – it has to primarily extract the relevant information for a specific task from the scene. Rather, a “region of interest” must be determined, which is used for processing. Finally, the system must have a short response time, because it is likely that a scene is changing over time and a delayed action may not result in the desired effect. For object recognition (“what” is “where” in a scene), many different methods have been proposed, including:

- Object detectors, which move a window over the image and determine a response for each position of any such filter by matching template and sub-picture (box contents). Each new object parameterization requires a separate scan. More sophisticated algorithms

31 Bajcsy R. (1988): Active perception, *Proceedings of the IEEE*, 76:996-1005

32 Crowley J. L., Christensen H. I. (1995): *Vision as a Process: Basic Research on Computer Vision Systems*, Berlin: Springer

expect the same at different scales and apply filters that have been learned from a large number of images.

- Segment based techniques extract a geometric description of an object by grouping of pixels that define the extent of an object in an image. Based on this, an immutable feature set is calculated, that is, the features contained therein retain the same values under various image transformations such as changes in the light conditions, scaling or rotation. These features are used to identify objects or object classes clearly. An example is the previously mentioned identification of road signs.
- Orientation-based methods use parametric object models that are trained on data.^{33,34} Algorithms search for parameters such as scaling, translation or rotation, which optimally fit a model on corresponding features in the image. An approximate solution can be found through a reciprocal process, i.e. by features such as contours, corners, or other characteristic points in the image “voting” for parameter solutions that are compatible with the feature found.

For conducting object recognition, it must first be decided whether algorithms are working on 2D- or 3D-representations of objects, whereby 2D-representations are often a good compromise between accuracy and availability. Current research in deep learning artificial neural networks shows that even distances between two points based on two 2D images, taken from different points, can be determined accurately. In daylight and reasonably good visibility this input can be used in addition to data from laser and radar in order to increase the accuracy – a mono-vision camera is sufficient to generate the necessary data. Unlike 3D-objects 2D-images do not encode any information such as shape, depth or orientation directly. Depth coding can be done in many ways, such as with the aid of laser or stereo cameras (like human perception) and structured light approaches (like Kinect from Microsoft). Currently, the most intensively pursued research direction relates super squares – geometric shapes defined by formulas that use any exponent to identify structures such as cylinders, cubes and cones with rounded or sharp corners. Using a small set of parameters, a wide variety of different base forms can be described. If 3D-images are determined by means of stereo cameras, due to poorer data quality compared with laser scans, statistical methods (such as generating a stereo point cloud) are applied instead of the previously mentioned form methods. Further research is done regarding tracking,^{35,36} contextual scene understanding,^{37,38} and surveillance³⁹, but these are currently for the automotive industry of secondary relevance.

-
- 33 D. P. Huttenlocher, S. Ulman: Recognizing solid objects by alignment with an image, *International Journal of Computer Vision*, 5: 195-212, 1990
 - 34 K. Frankish, W. M. Ramsey: *The Cambridge handbook of artificial intelligence*, Cambridge: Cambridge University Press, 2014
 - 35 Chaumette F., Hutchinson S. (2006): Visual servo control I: Basic approaches, *IEEE Robotics and Automation Magazine*, 13(4): 82-90
 - 36 Dickmanns E. D. (1991): *Dynamic Vision for Perception and Control of Motion*, London: Springer, 2007
 - 37 T. M. Straat, M. A. Fischler: Context-based vision: Recognizing objects using information from both 2D and 3D imagery, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13: 1050-65
 - 38 Hoiem D., Efros A. A., Hebert M. (2006): Putting objects in perspective, *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2137-44
 - 39 Buxton H. (2003): Learning and understanding dynamic scene activity: A review, *Vision Computing*, 21: 125-36

2.3 Logic and reasoning

Known as “Knowledge Representation & Reasoning” (KRR) in the literature, the focus in this field this field of research is on the design and development of data structures and inference algorithms. Problems that are solved by reasoning, are frequently found in applications requiring interaction with the physical world (as humans), such as the making of diagnoses, planning, natural language processing, answering questions, etc. KRR forms the basis for human-level AI.

Reasoning in the field of KRR is where data-based answers have to be found without human intervention or assistance, whereby the data are usually presented in a formal system with clear and precise semantics. Since about 1980, it is assumed that the data involved are a mixture of simple and complex structures, where the former are in a low degree of computational complexity and provide the research basis for large databases. The latter are presented in a more expressive language, which requires less space for representation, and correspond with generalizations and fine-granular information.

Decision-making is a kind of reasoning, in which answering questions about preferences between activities is paramount, such as when an autonomously acting agent tries to perform a task for a human being. Very often this decision-making is conducted in a dynamic domain, which changes by the execution of actions and the lapse of time. One example is the autonomously acting vehicle that must react to changes in traffic.

Mathematical logic is the formal basis for many applications in the real world, such as the theory of computation, our legal system and related arguments, and theoretical developments and evidence regarding research and development. The initial vision was to represent any kind of knowledge in the form of logic and to reason with universal algorithms, but some challenges have arisen; not any kind of knowledge is easy to represent. In addition, it can be very complex to compile the required knowledge for complex applications, and in addition it is not easy to learn this knowledge in from an expressive, logical language.⁴⁰ Moreover, it is not easy to reason with the required, expressive language – in the extreme case, such a scenario is computationally not feasible, even if the first two challenges can be solved. Currently three debates are conducted in this regard:

- First there is the statement that logic cannot represent many concepts, such as space, analogy, shape, uncertainty, etc., and therefore they cannot as an active part in the development of AI on human level are counted. The counter argument is that logic is just one of many tools. Currently, the combination of representative expressiveness, flexibility and clarity with any other method or other system is achieved.
- The second debate revolves around the statement that logic is too slow for inference and will therefore never play a role in a production system. The counterargument here is that ways for approximation of inference by logic exist, so that processing is approaching time limits and progress on logical inference is thus made.
- The third debate revolves around the argument that it is extremely difficult or even impossible to develop systems of logical axioms in real world applications. The counter-

40 Lavarac N., Dzeroski S. (1994): *Inductive Logic Programming*, vol. 3: *Nonmonotonic Reasoning and Uncertain Reasoning*, Oxford University Press: Oxford

arguments are mainly based on the research of those seeking techniques for learning logical axioms from natural language texts.

A distinction is made between four different types of logic⁴¹ that are not discussed further here:

- Propositional logic
- First-order logic
- Modal logic
- Non-monotonic logic

At this point it is also required to mention automated decision making, i.e. autonomously acting robots (vehicles), automated agents in the internet, or mapping decision finding processes I logic in order to automate those. Very often, such a decision process considers the dynamics in the surrounding world, such as when a transport robot has to avoid another in a production plant. However, this is not a prerequisite, for example when a decision-making process is carried out without well-defined direction in the future, i.e., the decision to rent a warehouse at a specific price at a specific location. Decision-finding as research discipline spans multiple domains, such as computer science, psychology, economics and all engineering. Some fundamental issues need to be addressed for the development of automated decision making systems:

- Is the domain is dynamic in that a sequence of decisions is required, or is it static, so that a single or multiple simultaneous decisions must be made?
- Is the domain deterministic, non-deterministic or even stochastic?
- Should benefits be optimized or a goal be reached?
- Is the domain known to the full extent or only partially at any time?

Logical decision problems are non-stochastic in nature, which includes planning and conflictory behavior. Both require that the available information on the initial and intermediate states is complete, and that actions have only deterministic, known effects and that there is a specific target definition. These types of problems are found in the real world frequently, such as in the control of robots, logistics, complex behavior on the internet, and computer and network security.

In general, a planning problem involves an initial (known) situation, a target definition and a set of permitted actions or transitions between steps. The result of a planning process is a sequence or set of actions, the correct execution of which moves the performer from an initial state to a state that satisfies the goal conditions. Planning is a computationally difficult problem, even if simple problem specification languages are used. The search for a plan can, even with simpler problems, not traverse through the entire state space graph, because it is exponentially larger in the number of states defining the domain. Therefore, it is attempted to develop efficient algorithms mapping subgraphs in order to browse these in the hope of achieving the goal. Recent research has focused on the development of new search methods and new representations of actions and states that allow easier planning. Especially when taking into account one or more agents working against each other, it is important to find a

41 Frankish K., Ramsey W. M. (2014): The Cambridge handbook of artificial intelligence, Cambridge: Cambridge University Press

balance between learning and decision – exploration for the learning's sake, while decisions are taken, can lead to undesirable results.

Many problems in the real world are problems whose dynamic is of stochastic nature – purchasing a vehicle whose properties are unknown to us and affect its value is an example. These dependencies affect the purchase decision and it is therefore necessary to incorporate risks and uncertainties. Stochastic domains are practically more difficult with respect to decision-finding, but also more flexible with regard to approximations than deterministic domains – the simplification of practical assumptions also makes automated decision-finding practicable. There are lots of problem formulations which can represent different aspects and decision making processes in stochastic domains, under the most famous decision networks and Markov decision-making processes.

Many applications require the combination of logical (non-stochastic) and stochastic elements, such as the control of robots requiring high-level specifications in logic and low-level representations with respect to a probabilistic sensor model. Natural language processing is another area to which applies this assumption, since high-level knowledge in logic is combined with probabilistic low-level models of text and voice signals.

2.4 Language and communication

Processing of speech is considered as fundamental in AI, and two fields are distinguished: Computational Linguistics (CL) and natural language processing (NLP). In summary, the difference that in CL research the use of computers for language processing is conducted, and NLU consists of all applications, such as machine translation (MT), Q&A, document summary, information extraction, etc. NLP therefore requires a specific task and is not per se a research discipline. NLP includes:

- Part-of-Speech Tagging
- Natural language understanding
- Natural language generation
- Automated summarization
- Entity recognition
- Parsing
- Voice recognition
- Sentiment analysis
- Voice-, topic- and word-segmentation
- Co-reference resolution
- Discourse analysis
- Machine translation
- Word sense disambiguation
- Morphological segmentation
- Answers to questions
- Relationship extraction
- Sentence separation

The central vision of the AI states that a version of first-order predicate logic (first-order predicate calculus, FOPC), supported by appropriate to the problem mechanisms for the representation of language and knowledge is sufficient. This assumption states that logic can

and should provide the semantics underlying natural language. Although experiments in AI and linguistics to leverage a form of logical semantics as a key for the representation of content have progressed, such were crowned with little success concerning a program translating English into formal logic. Even in psychology, it has not yet been proven that such translations in logic corresponds to the way in which people store and manipulate “meaning”. The translation from one language FOPC is therefore currently still a goal that has not yet been achieved. No doubt there are NLP applications must which establish logical inferences between sentence representations, but if these only account for a part of an application, it is not obvious that they have something to do with the underlying meaning of natural language (and thus with CL/ NLP), since the primary task of logical structures was inference. Different positions have emerged also in this area:

- Logical inferences are very closely linked to the meaning of sentences, because to know their meaning is the same as deriving inferences, and logic is the best way to do that.
- There is a meaning outside the logic that posits a number of semantic markers or primitives that are hung on the words to express their meaning. Today, this is commonly referred to as annotation.
- The predicates of logic and formal systems generally appear only different from human language, but their terms are in fact the words than they appear.

The introduction of statistical and AI methods is the latest trend in this regard in the field. The general strategy is to learn how language is processed, ideally in the way in which humans do, whereby this does not constitute a prerequisite. Regarding ML, this is learning based on very large corpora, which were translated manually by humans. This often means that it must be learned (algorithmically) how annotations are assigned, how corpora are provided with part-of-speech categories (allocation of words and punctuation of a text to speech), semantic markers or primitives, and all this based corpora prepared by humans (which are therefore correct). In supervised ML possible associations of part-of-speech tags with words that were annotated by humans on text can be learned so that the algorithms then can annotate new, hitherto unknown texts.⁴² This works weakly supervised or unsupervised as well, i.e. when no annotations were made from humans and only a text in one language with identical content in another language is presented, or even relevant clusters are found in thesaurus data, without having defined a goal.⁴³

Regarding AI and language information retrieval (IR) and information extraction (IE) play a major role, which very strongly correlate with each other. One of the main tasks of IR is grouping of texts based on their content, whereby IE extracts similar fact elements from texts, or is used to answer questions about text content. Therefore, these fields correlate with each other strongly, since individual records (not only long texts) can also be considered as documents. These methods are applied, i.e., in interactions of users with systems, such as when a driver asks the onboard computer a question about the operator manual while driving – after the linguistic input has been converted to text, based on the semantic content of the question an answer in the manual is searched, extracted and returned.

42 Leech G., Garside R., Bryant M. (1994): CLAWS4: The tagging of the British National Corpus. In Proceedings of the 15th International Conference on Computational Linguistics (COLING 94) Kyoto, Japan, pp. 622-628

43 Spärck Jones K. (1999): Information retrieval and artificial intelligence, Artificial Intelligence 141: 257-81

2.5 Agents and actions

In classical AI research mainly focused on single, isolated software systems that operated relatively inflexible on predefined rules. New technologies and applications have created the demand for artificial entities that are more flexible, adaptive and autonomous, and act as social units in multi-agent systems. In classical AI (see “Physical Symbol System Hypothesis”⁴⁴, which was embedded in so-called “deliberative” systems) an action theory, thus how systems make decisions and act, is represented logically in individual systems needing to perform actions. The system shall, based on these rules, prove a theorem – the requirements for this are that the system has

- a description of the world in which it is currently located,
- the desired target state and a set of actions,
- along with conditions for execution of this and
- a list with the results of each action gets.

It has been found that even in simple problems by the computational complexity each time-limited system is unusable, which had a great impact on symbolic AI, resulting in the development of reactive architectures. Such follow if-then rules converting inputs directly into tasks. Such systems are extremely simple, but can nevertheless solve very complex tasks. A problem is that such systems learn procedures, but not declarative knowledge, that is, they are learning attributes that simply cannot generalize to similar situations. There have been many attempts to combine deliberative and reactive systems, however, it seems that one must either focus on impractical, deliberative systems or on very loosely developed reactive systems – both are not optimal.

2.5.1 Principles of the new agent-centered approach

The agent-oriented approach can be characterized by the following principles:

- Autonomous behavior: Autonomy refers to the ability of systems to make their own decisions and carry out tasks on behalf of the system designer. The aim is to let a system act autonomously in scenarios in which it is difficult to control it directly. Traditional software systems perform methods after they have been called – they have no choice; Agents, however opt for their views, desires and intentions (Belief, Desire, Intention – BDI).⁴⁵
- Adaptive behavior: Since it is impossible to predict all situations which agents will encounter, they must be able to act flexibly. They must be able to learn from their environment and adapt. This task is more difficult, if not only nature is a source of uncertainty, but the agent is also located in a multi-agent system. Only environments that are not static and closed allow meaningful application scenarios for BDI agents, e.g. a lack of knowledge of the world may be compensated by reinforcement learning: agents are acting in an environment that is described by a set of possible states. Each time an

44 Newell A., Simon H. A.: Computer science as empirical enquiry: Symbols and search, Communications of the ACM 19:113-26

45 Bratman M., Israel D. J., Pollack M. E. (1988): Plans and resource-bounded practical reasoning, Computational Intelligence, 4: 156-72

agent performs an action, it will be “rewarded” with a numerical value expressing how good or bad was his action. This leads to a number of conditions, actions and rewards. The agent is now encouraged to identify an approach that maximizes reward.

- Social behavior: In an environment in which different entities act, it is necessary that agents recognize their opponents and form groups, provided a common goal requires them to do so. Agent-oriented systems are used for personalization of user interfaces, middleware and in competitions such as the RoboCup. In a scenario in which only autonomously vehicles are on the roads, not only autonomy, but also the exchange of information between vehicles and the thereupon based acting as group is an indispensable component. Through coordination between the agents the traffic flow is optimized – congestion and accidents are virtually impossible.

In summary, that the agent-oriented approach is accepted as a forward-looking direction in the AI community.

2.5.2 *Multi-agent behavior*

There are different approaches pursued to implement multi-agent behavior, which differ mainly in terms of how much control the designer on individual agent has.^{46,47,48} Differentiation is made between

- distributed problem-solving systems (Distributed Problem Solving systems, DPS) and
- multi-agent systems (multi-agent system, MAS)

DPS allows the designer to control every individual agent in the domain in which the solution of the task is distributed to more than one agent. In MAS, on the contrary, there are several designers and each can only influence his own agent and has no access to the design of all other agents. In this case, the design of the interaction protocols is very important. In DPS, agents try to achieve a goal or solve a problem, whereas in MAS, each agent is individually motivated to reach its own target and wants to maximize his own benefit. The research in DPS has to achieve the goal of cooperation strategies for problem solving while minimizing the degree of communication. In MAS attention is paid to the coordinated interaction, thus how the autonomously acting agents can be made to find a common basis for communication and carry out uniform actions.⁴⁹ Ideally, a world that is only traversed by acting autonomously vehicles, a DPS, however, due to competition between the OEMs a MAS will most likely arise first. The communication and negotiations between agents are in the foreground (see Nash equilibrium).

46 Bond H. Ah., Gasser L. (1988): Readings in Distributed Artificial Intelligence, San Mateo, CA: Morgan Kaufmann

47 Durfee E. H. (1999): Coordination for Distributed Problem Solvers, Boston, MA: Kluwer Academic, 1988

48 Weiss G.: Multiagent Systems: A modern approach to distributed artificial intelligence, Cambridge, MA: MIT Press

49 Frankish K., Ramsey W. M. (2014): The Cambridge handbook of artificial intelligence, Cambridge: Cambridge University Press

2.5.3 *Multi-agent learning*

The multi-agent learning (MAL) is given a certain degree of attention only recently. Key issues are the definition of techniques that should be used and what exactly multi-agent learning is. Current ML approaches have been developed to train individual agents, whereas at MAL distributed learning is the target. Distributed does not necessarily mean that a neural network, in which many identical operations are run during exercise and this can therefore be parallelized, is executed, but

- that a problem is split into sub-problems and individual agents learn these in order to solve the main problem based on their combined knowledge, or that
- many agents independently try to solve the same problem by competing against each other.

Reinforcement Learning is an approach which takes in this respect application.⁵⁰

2.6 Summary

Artificial intelligence has already found its way into our daily lives and is no longer exclusively the subject of science fiction novels. At present, AI is mainly used in the following areas:

- Analytical data processing
- In domains in which quick, qualified decisions must be made on a large amount of (often heterogeneous) data
- At monotonous, but attention-requiring activities

In analytical data processing, over the next years we will no longer rely on decision support systems, but on systems that relieve us of decisions. Specifically, in data analysis we are currently developing individual, analytical solutions for specific problem statements, but those cannot be used across contexts – for example, a solution developed for anomaly detection in stock value behavior cannot be used to understand content of images. Also in the future, this will not be the case, but AI systems will integrate individual, interacting components and will so take over more complex tasks currently reserved for humans – a clear trend we can already observe now. A system not only processing the latest data from stock markets, but also analyzing the development of political structures on news texts or videos, extracting sentiments from from blogs or social networks, monitoring and predicting relevant financial ratios, etc., requires the integration of many different subcomponents – encouraging them to work together is the subject of current research, and progress is published weekly. In a world in which AI systems are able to improve themselves continuously and to control companies more effectively than humans, what remains there for the humans? Time to focus on expanding one's knowledge, the improvement of society, the eradication of hunger, elimination of diseases, propagation of our species over the boundaries of our own solar system. Some theories suggest that quantum computers are needed for the development of strong AI systems, and only someone who is very careless, would argue an effective quantum computer would be available within the next 10 years. And only someone who is careless

50 Busoniu L., Babuska R., De Schutter B. (2008): A comprehensive survey of multi-agent reinforcement learning, IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews 38: 156-72

would argue that that is not the case. Be that as it may, as with the majority of relevant, scientific achievements history shows that artificial intelligence must be used wisely – systems that can take exponentially more decisions in extremely short periods of time with increasing hardware performance, can cause many positive things, but also have the potential to be abused.

Reverse Engineering the Mind

Consciously Acting Machines and Accelerated Evolution

Neukart, F.

2017, XXXI, 383 p. 105 illus., 21 illus. in color., Softcover

ISBN: 978-3-658-16175-0