Chapter 2
Five Generations of MT

Introduction (2007)

There is an ancient Chinese curse that dooms recipients to live in an interesting age, and by those standards MT workers are at present having a bad time. The reason things are interesting at the moment is that there is a number of conflicting claims in the air about how to do MT, and whether it can, or indeed has already, been done. Such a situation is unstable, and we may confidently expect some kind of outcome – always cheering for the empiricist – in the near future. This chapter was initially published as (Wilks, 1979) as an exploration of the relationship of MT to Artificial Intelligence (AI) in general, and should be seen as providing a snapshot of that intellectual period.

What happened initially was threefold. First, the "brute force" methods for MT, that were thought to have been brought to an end by the ALPAC (1966) Report have surfaced again, like some Coelacanth from the deep, long believed extinct. Such systems were sold for many years under such trade names as LOGOS, XYZYX, SMART, Weidner and SYSTRAN; and the last, and best known, has been used for thirty years by the EU in Paris (Van Slype, 1976) and Luxembourg.

Secondly, some large-scale, more theoretically based, MT projects continued, usually based in Universities, and have been tested in use, though sometimes on a scale smaller than that originally envisaged. METEO, for example, in Montreal (Chandioux, 1976), which was to have translated official documents from English to French, is still in use for the translation of the more limited world of weather reports.

Thirdly, workers in natural language in the field known as Artificial Intelligence (AI) began to make distinct claims about the need for their approach if there is ever to be general and high quality MT (Wilks, 1973a; Charniak, 1973; Schank, 1975a). Small pilot systems illustrating their claims were programmed, but their role in MT discussion was mainly of a theoretical nature.

However, these are not merely three complementary approaches, for they seem to be making different claims, and, unless we take the easy way out and simply define some level of MT appropriate to each of the enterprises, it seems they cannot all be right, and that we may hope for some resolution before too long.

These three correspond very roughly, to what are called the "generations" of MT, but to that I think we can add one or two more. The fourth is certainly the revival of empirical statistical methods in MT, which began around 1989 and lost momentum in the 90s when the early systems, like Jelinek's (see Chapter 1) failed to beat SYSTRAN decisively. However, empirical methods then colonised the whole of NLP, area by area, and now in the new millennium have returned to tackle MT itself (see Chapter 16 below). A possible fifth is that of hybrid methods, where researchers are seeking combinations of empirical methods with intelligent revivals of, earlier conceptual AI approaches (again, see Chapter 16 below).

It is interesting to notice that the reactions of Bar-Hillel and AI workers like Minsky were in part the same: Minsky (1968) argued that MT clearly required the formalization of human knowledge for a system that could be said to understand, or as Bar-Hillel reviewed the situation in 1971 (Lehmann and Stachowitz, 1971, p. 72):

> "It is now almost generally agreed upon that high-quality MT is possible only when the text to be translated has been understood, in an appropriate sense, by the translating mechanism".

In this chapter I want to explore some of the early connections between machine translation (MT) and artificial intelligence (AI). We all feel we understand the first phrase, and the second will be explained as we go along: for the moment, I ask the reader to accept some such working definition as: the use of computational methods for the simulation of distinctively human intellectual behaviour, by means of complex knowledge structures and their manipulation.

In what follows I shall first sketch some AI systems, and my argument will be that AI is relevant to, and important for the future of, MT, but that one can hope that a limited AI will be what will help in the foreseeable future, rather than those Al systems which appear to claim that they can express "all the knowledge in the universe" whatever that may mean.

A good place to start is Bar-Hillel's argument (1962) that MT impossible: for it has striking resemblances in terms of its premises (though not its conclusions) to some AI views.

Knowledge and MT

Bar-Hillel argued, even at the height of the early and finally disastrous MT period, that machine translation was not only practically but theoretically impossible, where "impossible" meant just that, and not merely difficult. "Expert human translators use their background knowledge, mostly subconsciously, in order to resolve syntactical and semantical ambiguities which machines will either have to leave unresolved, or resolve by some 'mechanical' rule which will ever so often result in a wrong translation. The perhaps simplest illustration of a syntactical ambiguity which is unresolvable by a machine except by arbitrary or ad hoc rules is provided by a sentence, say '. . . slow neutrons and protons. . .' whereas, in general, though by no means always, the human reader will have no difficulty in resolving the ambiguity through utilisation of his background knowledge, no counterpart of which could possibly stand at the disposal of computers" (1962).

The immediate historical context of Bar-Hillel's argument was the performance of early syntax analysers, which, according to legend, were capable of producing upwards of 10 grammatical parsings of sentences like "Time flies like an arrow". With respect to standard dictionary information, any of the first three words in the sentence could be taken as a possible verb. To see "time" as the verb, think of the sentence as command with the accent on the first word; to see "like" as the verb, think of the sentence as expressing the tastes of a certain kind of fly, and so on.

The standard reaction to such syntactic results was to argue only showed the need for linguistic semantics, so as to reduce the "readings" in such cases to the appropriate one. Bar-Hillel's response was to argue that it was not a matter of semantic additions at all, but of the, for him, unformalisable world of human knowledge.

The contrast can be seen by looking at our everyday understanding of so simple a sentence as "He paddled down the river in a canoe". The standard machine parser, working only with grammatical information, will not be able to decide whether the clause "in a canoe" attaches to "paddled" or "river". The first reading, the correct one of course, tells you how he went down the river. The second implies that we went down a river that happened to be inside a canoe – the same structure that would be appropriate for "He paddled down the river in an unexplored province of Brazil". The purely syntactic parser has no way of distinguishing these two possible "readings" of the sentence and, furthermore, there is a difference of opinion as to how the information that would resolve the problem should be described. Those who take a more "linguistic semantics" view would say that it is part of the meaning of "canoe" that those objects go in rivers and not vice versa; whereas those of an Al persuasion would be more likely to say that it is merely a fact about our world that canoes are in rivers. At bottom, there is probably no clear philosophical distinction between these views, but they do lead to different practical results when attempts are made to formalise and program such information.

Bar-Hillel went further and produced an example (the best-known in the history of MT) proving, as he thought, the impossibility of MT. In brief, Bar-Hillel's example was the following children's story:

> Little John was looking for his toy box.
> Finally he found it.
> The box was in the pen.
> John was very happy.

Bar-Hillel's focus is on the third senteiipnce, "the box was in the pen", whose last word we naturally interpret in context as meaning playpen and not writing pen. Bar-Hillel argued persuasively that to resolve this correctly requires knowledge of the real world, in some clear sense: at least in the sense that the difficulty cannot be overcome in terms of some simpleminded "overlap of concepts", by arguing that the concepts of "baby" and "playpen" can be seen, by lexical decomposition of some sort, to be related in a way the concepts of "baby" and "writing pen" are not. Bar-Hillel argued that that would not do, because the story would have been understood the same way if the third sentence had been The inkstand was in the pen, where the semantic "overlap of concepts" would now be between inkstand and writing pen which would yield the wrong answer on the same principles.

Bar-Hillel thought that the absolute impossibility of high-quality machine translation had been demonstrated, whereas Minsky believed that the task had now been defined, and the job of Al was to get on with it.

The contrast is clear between the views of Bar-Hillel and Minsky on one hand, and the views of linguists on the other: Chomsky's generative theories are also, in a clear sense, a reaction to the failure of the early machine translation work, in that they state the case, with great force, for a solid theory of the syntax of natural languages as a precondition for any advance with machines and language. Katz and Fodor's (1963) semantics, conjoined to a Chomsky grammar, represent, as it were, the linguistic analogue to those in machine parsing who thought that purely semantic information would be enough to resolve the multiple analyses of the notorious "Time flies like an arrow".

The essence of the Katz-Fodor method was algorithms based on the repetition of "semantic markers". We can see the limitations of that method for MT by looking at a complex (and realistic) noun phrase like the following:

> "Analyse d'une méthode dynamique specifique d'établissement de balance materiel d'une installation de retraitement de combustion nucleaire par simulation".

Roughly : Simulation analysis of a dynamic specific to the establishment of a balance of material in an installation for the treatment of nuclear combustion. In English the ambiguity of dependence of ≪ Simulation ≫ does not exist.

The problem in French concerns the dependence of "par simulation", which is at least a two-way ambiguity (of dependence on the heads "analyse" and "établissement", where the former is correct), but one could argue that it might conceivably depend (semantically speaking, that is) on either "retraitement" and "combustion" as well, since they are both processes that could be simulated, just like the earlier two. One might argue as follows: Semantics means attaching markers and looking for their repetition, so we might attach a marker, say, PROCESS to "analyse", "établissement", "simulation" (and perhaps "combustion" and "retraitement", as well).

The trouble with this course of action should be obvious: we would have attached the same plausible semantic marker to all the possibilities, and so there can be no discrimination (of the correct dependence, in this case, on "analyse"). The reader should appreciate the force of this example for if the, essentially semantic, dependence of "par simulation" on "analyse" cannot be determined by rules (stronger, as we shall see, than mere repetition of semantic markers) then a unique syntactic structure for the phrase cannot be obtained either.

What are AI Systems?

The attempt by AI research to respond to Bar-Hillel's challenge is of a different sort. It is an attempt not only to admit from the beginning the need for "knowledge structures" in an understanding system, but also to formulate theories and systems containing processes for the manipulation of that knowledge. "Processes" here is not to be taken to mean merely programming a computer to carry out a task, for many interesting AI systems have either not been programmed at all or made to do only partial demonstrations. The word "process" means that a theory of understanding should be stated in a symbol-processing manner, one in which most linguistic theories are not. This is a contentious position, because generative grammar has also been in some sense a description of a process since the earliest descriptions of transformational theory. The AI case is that it never quite comes up to scratch in processing terms.

But what is an AI theory of language, and how might it help machine translation?

AI has been concerned, for some forty years now, with the problems of human intelligence seen from a particular point of view: what would it be like to program a computer to perform intelligent tasks that we do without even thinking about them; such as seeing and understanding what we see, understanding language, and inferring from what we understand? Some choose to investigate machine performance of tasks, like chess playing, that even humans find difficult, but the "unconscious tasks" remain the heart of AI.

As applied to the field of natural language understanding this has meant constructing elementary programs to carry out written commands, translate into another language, make inferences, answer questions, or simply carry on a dialogue – all of which are presented as written responses at a teletype or video screen.

As can be seen, machine translation is by no means a typical Al language program, but no difference of principle arises between the different sorts of task, especially if we accept a slogan like Steiner's that, in some sense, all acts of understanding are acts of translation (1975).

What almost all AI language programs have in common – though they differ widely over other assumptions – is strong emphasis on the role of knowledge in understanding, and on the presentation a theory as a possible process. In some programs – like the well-known one constructed by Winograd (1972) – this last assumption leads to writing the syntactic analysis part of the program in a special "grammar programming language" PROGRAMMAR, rather than as the normal battery of grammar rules like $S \rightarrow NP + VP$. This rule appears in all grammars and simply means that a noun phrase (NP) followed by a verb phrase (VP) is a well-formed sentence (S). In Winograd's system that grammar rule exists only as tiny program in PROGRAMMAR.

Winograd's program accepted dialogue and commands about a miniature world consisting only of a few blocks and a box, which it could appear to move about on the video screen. He wanted to show the role of knowledge of this microworld of blocks as a tool for resolving syntactic ambiguities in input to the system. So, for example, when it saw the sentence "Put the pyramid on the block in the box", it would immediately resolve the surface syntactic ambiguity of the command: that is, does it refer to a particular pyramid (on a block) to be picked up, or to a particular place to put it (on the block in the box), according to whether there actually was a block under a pyramid, or already in the box, in the small blocks scene that it understood.

Winograd's program could be called pure Al, in that it was motivated by classic problems of Al: plans (how to pick up the blocks) and theorem proving (how to show which is under the pyramid at a given moment), rather than being motivated by problems left over from the 1966 failure of machine translation, such as word-sense ambiguity, and correctly referring pronouns in discourse. Another group of Al language programs, such as the work Charniak (1973), Schank (1975a) and myself (1973, 1975) was directed more at those left-over questions: at meaning representation, and the use of inference rules: not about microworlds of blocks, but about the more general world in which we live.

Consider a simple sentence like "The soldiers fired at the women and I saw several fall", where we may be sure that any native speaker of English will understand that sentence (out of any further context, which may change matters, so let us leave it to one side) in such a way that "several" refers to the women and not to the soldiers. That cannot be done on any simple semantic (or syntactic) grounds since both soldiers and women are capable of falling. Mere proximity, in that "women" is closer to "several" in the sentence than the alternative, is known to be an unreliable guide in general. Correct reference of the pronoun "several" – and this might be vital in translation into a language where "soldiers" and "women" had different genders, for example – must almost certainly be done using some general inference rule like "If animate things have an object projected at them, they may well drop downwards". If the reader finds that implausible, he should ask himself just how he refers the pronoun correctly in that sentence.

The type of knowledge expressed in that rule is what one might call, following McCarthy partial: it is an inference that is not always true. It is a kind of knowledge that has no place in the very limited Winograd blocks world, but is central to the understanding capacities of the Charniak, Schank and Wilks systems. The three systems differ strongly in other respects: for example, the Schank and Wilks systems emphasise knowledge that can be expressed in very general terms, like the inference rule above, and develop notations of semantic primitives (actions like CAUSE, and CHANGE; entities like THING, and MAN for human being) in order to express this. In Charniak's systems, on the other hand, the knowledge is quite specific to a certain topic, like present giving.

Machine translation has traditionally been much preoccupied with the problem of finding the topic in a text: in the "Time flies like an arrow" example, with it's three classic readings, we would have the correct reading immediately if we could find out, from wider context, that the sentence is about, say, time, and not about flies or liking. The semantic system of Charniak tried to detect the topic by specific cues, while the Schank and Wilks systems did so by general rules ranging over semantic representations expressed in primitives. In the Winograd system, on the other hand, topic can never be a problem because it is always the blocks world.

There is no doubt that Al systems can be brought to bear upon the problems of machine translation: the system described in Chapter 3 actually translated English into French and resolved word-sense and pronoun ambiguities that could only be resolved with the aid of the sort of partial knowledge used in the soldiers and woman example. There is enough capacity in such systems to express knowledge about protons and neutrons so as to have no difficulty with Bar-Hillel's phrase "slow neutrons and protons". If he were to protest that it was ad hoc for the system to code only one of those entities, say, as being potentially slow, then one could reply by asking how he could know that humans do not understand this sentence with precisely such a specific coding of knowledge.

But much may depend on one's choice of examples: it is not clear that the difficulty has been eased by these Al systems for old favourites like Time Flying. The partial knowledge systems I described might well know that things that flew were normally birds or planes, rather than time, and so they would have no reason to pick out the reading where time does the flying on such preference grounds. Given that flies can indeed be timed, such systems might well decide that the "imperative reading" was the one most suited to their general knowledge about the world with which they had been programmed. This is a melancholy conclusion, because it suggests that our competence with such examples can only be credited to an ability to read them off a list of prestored clichés, together with the interpretation "we feel as if time moves quickly". This would be a sad conclusion for all theoretically motivated work, and an awful fate for a long cherished example.

Frame Systems in Al

After the systems just described, a new complexity was introduced under the influence of a proposal of Minsky (1975) that the knowledge structures in use in AI – and he was writing about machine vision as well, but here we shall concentrate only on language – should be higher-order structures that he called frames. One can see the sort of thing he was getting at by considering the statement: "John went into a supermarket and put some soap into his basket. On impulse he put a bar of chocolate in his pocket as well, but when he reached the cash desk his face went red and he said 'I didn't mean to take it".

The question that might come up in machine translation is how we know that the "it" refers to the chocolate, and not to the soap. The two words might have different genders in some output language, and so we would have to get the decision right, and in a general and plausible manner. It is easy to see that one might need to have access, even for this apparently simple task, to some complex formalised structure expressing what normally went on in a supermarket, so that one could infer from it that putting buyable items in one's pocket was not normal behaviour. Notice that it would have to be very specific information too, because it would not be enough to know that, in a supermarket, one normally puts buyables into a container, for a pocket is certainly a container. On so general a description of the activity of shopping the "abnormal" act would slip through unnoticed. Only the specification of (cart or basket) will do the trick.

It is just such highly complex and specific knowledge structures that Minsky argued should be called frames, which, in some formalised version, would be essential to any computerised language understanding system.

Let us begin with the standard quotation from Minsky that best captures the general notion of "frame": "A frame is a data-structure for representing a stereotype situation, like a certain kind of living room, or going to a children's birthday party". Attached to each frame are several kinds of information. Some of this information is about how to use the frame. Some is about what one can expect to happen next. Some is about what to do if these expectations are not confirmed.

> "We can think of a frame as a network of nodes and relations. The top levels of a frame are fixed and represent things that are always true about the supposed situation. The lower levels have many terminals . . . 'slots' that must be filled by specific instances or data. Each terminal can specify conditions its assignments must meet."

Under the influence of Minsky's proposal, Charniak (1975) produced a frame for shopping in a supermarket (to deal with examples like that about soap and chocolate), while Schank (1975b) produced similar structures but called them scripts. Schank defines a script as "a predetermined causal chain of conceptualisations that describe a normal sequence of things in familiar situation", by which he means some account, capable of simple formalisation, of the normal order of events when visiting a restaurant. He sketches a restaurant script as follows:

Script:		
Restaurant roles:		Customer; waitress; chef; cashier
Reason:		to get food so as to go down in hunger and up in pleasure
scene 1 entering		
PTRANS	-	go into restaurant
MBUILD	-	find table
PTRANS	-	go to table
MOVE	-	sit down
scene 2 ordering		
ATRANS	-	receive menu
ATTEND	-	look at it
MBUILD	-	decide on order
MTRANS	-	tell order to waitress

and so on for scenes 3 eating and 4 exiting. For the reader to get the general idea, we need not go into the precise definitions of the associated primitive actions: entities like PTRANS on the left-hand side (indicating physical movement, in this case) that Schank uses in his underlying semantic conceptualisations of sentences in the computer. Schank's students wrote a program which to take a paragraph-length restaurant story and produce a longer story with the "missing parts" filled in from the script above; and do this in a number of output languages, thus producing a rather new definition of machine translation.

The question is what exactly frames are for in language-understanding systems; what hypothesis their use implicitly appeals to; and whether the benefit they confer could be obtained by other simpler means? There is no doubt they express the dynamic order of events that is part of the meaning of certain concepts, in some intuitive sense.

Moreover, the frame is potentially a powerful device for defining topic context, a problem that has plagued all formal work with language since the earliest machine translation. So, for example, if we see the sentence "John ordered an omelette", we know that it is the "ordering food" sense rather than the "order people about" sense (and these are expressed by different words in French and German, for example, so that for MT the right sense would have to be found). If we are processing a particular text with the aid of the "restaurant script" this problem will already have been settled because the Schankian MTRANS (in the last line of scene 2) will be tied only to the appropriate sense of "order".

This point may be clearer if we think of a language understanding system encountering a word it did not know: suppose it encountered "John ordered scampi", although "scampi" was not in its dictionary. Suppose the system had no restaurant script, but just representations of the senses of "order", including the appropriate one in which ordering was normally done by humans and of physical objects. These normal objects and agents we can call the preferences of the action, because they are not absolute – we can all understand children's stories with sentences like "The dog ordered a bone in the doggy shop" – but they do enable important semantic choices to be made. In "John ordered the numbers", for example, we can reasonably say that we select the mathematical sense of "order" because numbers fit the preferred object for that particular sense, though not the preferred physical object of the sense of "order" appropriate to "ordering things commercially".

Now we can see the payoff from the restaurant script: if we are analysing our sentences with it then we know that even the unknown "scampi" is almost certainly a food, just because that is the preferred object of the sense of the action tied into the script at that point. If we had only the general sense of "order" we could infer only that a physical object was ordered.

Frames or scripts, therefore, will certainly help in determining individual word context, provided that we can reliably decide in advance what is the appropriate frame with which to analyse the given input. This assumes reliable cues (the word "restaurant" for example) which will not always be present ("They stopped off to eat at a little place he knew"), and a way of deciding which of these large-scale information structures to use when several have been cued by a single sentence ("On the way home from the cinema, they stopped off at the supermarket before dropping into Luigi's restaurant"). Problems will also arise as to when to stop following one script and get rid it in favour of another.

The real issue, though, is not technical but concerns what claims are being made by frame users. They are, I think, making what I will call a plot line hypothesis as follows: Humans, or computer understanding systems, can only understand a particular story by seeing how far it follows, or diverges from (as did the chocolate and soap story), the stereotypical story of that type. Or, as Charniak (1975) puts it:

"The primary mechanism in understanding a line of a story is to see it as instantiating one or more frame statements".

The trouble is that the claim is not obviously true, as we can see by making up an imaginary frame about a more remote cultural activity. I have jotted down the following for a male puberty rite in Charniak's (1975) notation one which is more or less self-explanatory:

Frame: male puberty rite
Roles: male child, village elder, helpers, crowd
Reason: placing ritual incisions on back of child
(a) Goal: CHILD is tattooed
(b) HELPERS hold CHILD (by both arms)
(c) ELDER obtains TOOLS
(d) ELDER exhorts CROWD (on proper behaviour)
(e) (general condition) Bad behaviour by CROWD => Activity halts
(f) ELDER checks if CHILD properly purified
(g) (special condition) CHILD not purified => activity halted
(h) ELDER marks CHlLD's back
(I) (method suggested) do for all CUT-MARKS

and so on. Again the general idea is clear, and the choice of a remote, and imaginary, culture is not accidental, as I shall now try to show.

Suppose we have three stories contained in three sentences:

"Little Kimathis's mother (looked away accidentally)"
"Little Kimathis's mother (dropped her shoga)"
"Little Kimathis's mother (touched his arm) during the puberty rite. The crowd drew back in horror".

If we wish to understand these stories, do we need the frame above to do it? The frame covers the end of the story with line (e) in some sense, given an adequate 1ist defining bad behaviour accessib1e from the frame.

And yet it is clear that we understand the sentences perfectly well without the frame. In commonsense terms we could say that we infer from the sentences that the mother touching Kimathi during the ceremony was bad thing. We do not need that information in order to understand.

One might argue that, in order to understand the above, a program should tie two parts of its representation together with some rule equivalent to:

human display alarm => other human has performed bad action.

A Martian lacking any earthly frame could understand the stories so long as he understood this rule and the constituent words. That is, of course, why I chose a puberty rite rather than a restaurant as a frame topic, for most of us are Martians where puberty rites are concerned. If we do understand the stories (and we do) it cannot be from our associated frame, because we do not have one. So we must understand it on the basis of knowledge organised on some simpler principles.

At present there is a tension between those who believe that frames are necessary for language understanding, and those who think whatever is necessary can be provided by a system of cues and inference rules no more complex than the "humans show alarm" rule So, to return to the "ordering scampi" example, provided we had a restaurant cue (which even a frame needs, as we saw) we could have an inference rule tied to that cue that said "ordering is now normally of food". The reply from frame advocates is that these inference rules would be too numerous to be accessed but, as we saw, there are also enormous problems about access to and manipulation of frames, so that this question is not settled, either by argument or by the performance of programs.

Some frame advocates are not urging the "plot line hypothesis" (PLH) in the strong form of "you must have structure X to understand" but are claiming that it is more efficient to understand text from the topmost level down in that way.

However, an efficiency-PLH cannot be assessed in the absence of frame application procedures. Moreover, and this is the important point, an efficiency-PLH almost certainly rests on some statistical assumption about the degree to which texts do in fact follow the frame norms: the PLH would clearly be more plausible if, say, 90 per cent of texts about X were consistent with the knowledge in the frame about X, than if, say, only 5 per cent of texts about X did so.

The Argument that Less-than-Frame Al Systems have a Role in MT

I want to argue that, although Bar-Hillel was wrong, and everyday knowledge can be manipulated in Al systems, nevertheless we may not need to go as far as the frame systems just described in order to alleviate some of the pressing problems of MT. Let me briefly recap the notation of my own semantics-based natural language understanding system (NLUS) (Wilks, 1973a, 1973b).

In other places I have described an NLUS in which rules operate on semantic word-sense descriptions to build up text descriptions. The rules that insert sense descriptions into text descriptions are what I have called "preferential": they seek preferred entities, but will accept the less preferred if necessary. A sense description for the action "drink" might be the formula:

((*ANI SUBJ)(((FLOW STUFF)(OBJE)((*ANI IN)(((THIS(*ANI (THRU PART)))TO)(BE CAUSE)))))

Figure 2.1 is a formal structure of semantic primitives expressing the meaning of the action "drink" (see Wilks 1973a): that drinking is a CAUSing to MOVE, preferably done by an ANlmate SUBJect (=agent) and to a liquid (FLOW STUFF), TO a particular ANImate aperture (THRU PART), and INto the SELF (=the animate agent). For short we will write Fig. 2.1 as [drink]. The text structures in this system are semantic templates (together with semantic ties between them): a template is a

((*ANI SUBJ)(((FLOW STUFF)(OBJE)((*ANI IN)(((THIS(*ANI (THRU PART)))TO)(BE CAUSE)))))

Fig. 2.1 A semantic formula for the action of drinking

network of formulas, containing at least an agent, action and object formula. Thus the template for "The adder drinks water" will be written: [the+adder drinks water] for short where the whole of Fig. 2.1 is in fact at the central (action) node of that structure.

The process of setting up the templates allows the formulas to compete to fill nodes in templates. Thus the formula for the (snake-) adder goes to the agent node in the template above in preference to the (machine-) adder because Fig. 2.1 specifies, by (ANI SUBJ), that it prefers to be accompanied in a template by an animate agent formula. However, in the sentence:

My car drinks gasoline

the available formula for the first template node, namely [car], is not for an animate entity, yet it is accepted because there is no competitor for the position.

An important later process is called extraction: additional template-like structures are inferred and added to the text representation even though they match nothing in the surface text. They are "deeper" inferences from the case structures of formulas in some actual template. Thus, to the template for [My car drinks gasoline] we would add an extraction (in double square parentheses in abbreviated form):

[[gasoline in car]]

which is an inference extracted from the containment subformula of Fig. 2.1, (SELF IN). Analogous extractions could be made for each case primitive in each formula in the template for [my car drinks gasoline].

After the programmed version of the system, reported in (Wilks 1978), a structural change (Wilks 1976b) allowed a wider, and more specific, form of expression

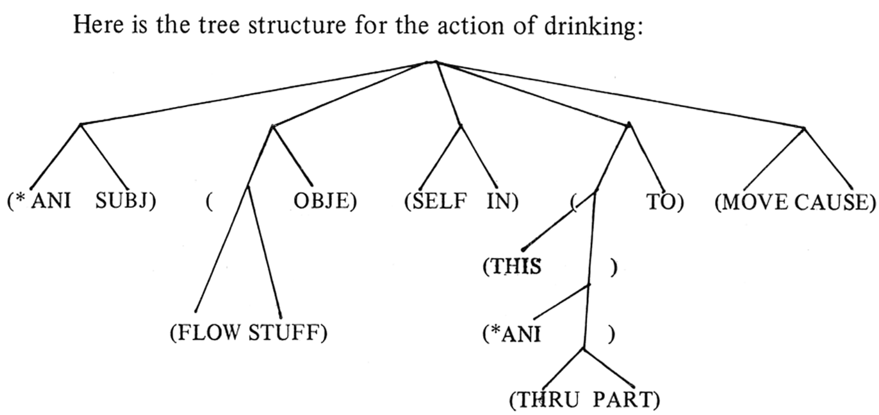

Fig. 2.2 The action formula for drinking installed at the central action node of a semantic template of formulas for "John drinks beer"

in formulas by allowing thesaurus items, as well as primitives, to function in them. No problems are introduced by doing this, provided that the thesaurus items are also themselves words in the dictionary, and so have their formulas defined elsewhere in their turn. One advantage of this extension is to impose a thesaurus structure on the whole vocabulary, and so render its semantic expression more consistent.

Let us now return to two of the examples introduced earlier: first, the complex French noun phrase. Let us imagine that its four key words “analyse”, “establish”, “simulate” and “understand” (we will ignore the problems of “combustion” and “retraitment”, since they are not in the heart of the matter) have as part of their meaning expression inference rules (and those need not be distinct – but merely types of label for nodes on a syntax tree).

1.	[Human analyses X]	→ [Human wants X]	1i
		→ [Human understands X]	1ii
		→ [X has components/parts]	1iii
		→ [Human uses a process/method]	1iv
2.	[Human establishes X]	→ [X does not change]	2i
		→ [Human uses a process/method]	2ii
3.	[Human simulates X]	→ [Human wants X]	3i
	[Human understands X]	→ [X is process/method]	3ii

If we think of these rules, written as template-to-template patterns, chaining across the text under the guidance of a reasonable semantic algorithm, it should be clear that “analyse” chains to “simulation”, on plausible grounds, but “etablissement” does not. The appropriate chaining is:

From “analyse” (the rules 1i, 1ii, 1iii are pointed to by that word) we infer by 1ii
[the human (analyser) understands (the method)]
also we infer (backwards up the text) from simulation via 3i and 3ii
[the human (simulator) understands X]

Hence, given this chain, leading to a close match of semantic patterns (after identifying X with method) we infer that it is the analysis that is simulated and hence make the prepositional phrase (“par simulation”) depend appropriately and so resolve the syntactic tree ambiguity.

Notice that both 1iv and 2ii (from “analyse” and “etablissement” respectively) are both patterns seeking the one yielded from simulation by 3ii. In other words, both seek a process for achieving their ends (this is equivalent to the simple “process” match which yields no discrimination between the candidates because both matches succeed). It is only the quite separate pattern matching via 1ii and 3ii that succeeds in discriminating, and this one is not reducible to simple marker repetition, hence our method is not reducible to that of Fodor and Katz.

We might even attempt to apply this method of semantic pattern manipulation to Bar-Hillel’s favourite example: the box in the pen, given earlier.

First, let us assume two sense formulas for “pen” in the notation given earlier, as trees or primitives.

Let us now consider the treatment of [The box is]/[in the pen] broken in to two templates as shown. The sentence will have been fragmented at the stroke by initial procedures and a template will be attached to each part: the first template having a dummy object place, and the second a dummy agent place, since a formula for “in” becomes the “pseudo-action” of the second template and has no agent.

Thus the parsing phase will have implicitly assigned formula trees to the slots of the two templates as follows:

FIRST TEMPLATE

Agent – formula	Action – formula	Object – formula
[The box]	[is]	Dummy

SECOND TEMPLATE

Agent – formula	Action – formula	Object – formula
Dummy	[in]	$\left\{\begin{array}{l}[\text{playpen}]\\ [\text{writingpen}]\end{array}\right\}$

There will, of course, be two second templates with the two different trees above at their respective third nodes.

Inference structures called paraplates, specific rules for prepositions, whose nature need not detain us (see WiIks 1978), then seek to link the two templates back together, the paraplate being in effect a case frame that resolves the “in” of this particular sentence as the introducer of the CONTAINMENT case. The application of this paraplate allows the dummy agent of the second template to be “repacked” in this particular containment case frame by the agent of the first template and thus we obtain, by a “repacking inference”, a template form in the representation equivalent to [box in pen], which is not, of course, an assertion explicitly present in the text. This “repacked template form” will have a formula for “box” at its agent node and, since we still have two formulas for “pen” in play, not having yet resolved between them, we shall in fact have two repacked templates at this point, both with a formula for “box” at their agent node, and with the two formulas for “pen” at their respective object nodes. Or, expanding the “square bracket”, or shorthand form.

[box in pen] and [box in playpen]

[playpen] stands for the formula for “playpen” which will contain the primitive WRAP (for “contain”), and [box] the formula for “box”, whose head is the primitive THING.

Lastly we have to remember another very general process within this semantic representation system. When the templates of formulas are first set up for sentence fragments, an attempt is made to see what “semantic preferences”, expressed by the formulas are in fact satisfied by the presence of neighbouring formulas in the templates (in which the formulas are now installed).

Thus, in a template for [John drinks gin] the formula [drinks] (Fig. 2.1 above) shows within its tree that drinking is normally done by animate beings (just as the formula tree for “playpen” showed that it normally contains children). So in [John drinks gin] the animate agent “preference” of [drinks] is satisfied by the presence of

[John] (which can be seen to be animate because its head is MAN) at the agent node of the template that has [drinks] at its action node.

The general preference rule of inference in the system is to take, as the provisional semantic representation at every stage, the template with the most satisfied preferences between its constituent formulas.

So now, let us envisage this process re-applied after the application of case paraplates and the consequential repacking. If we do this to the two competing templates we still have for "in the pen", and one turns out to have more preferences satisfied than the other then we shall, by this general rule of inference be able to discard the latter.

The system will consider whether the box is plausibly inside the playpen or writing pen from the inference rule (of the same type as used in the "combustion nucleaire" example above).

$$[\mathrm{X} \quad \mathrm{WRAP} \qquad \mathrm{Y}] \rightarrow [\mathrm{Y\ in} \qquad \mathrm{X}]$$

which links, as inverses, the internal structure of the [pen] formulas to the structure of the two templates.

Inside the formula tree for "playpen" we see that playpens prefer to contain children, while writing pens prefer to contain liquids. And, since a box is a physical object (the head of its formula is THING), and so is a child, while a liquid (the head of whose formula is STUFF for substances) is not a physical object. Thus it is clear that the first template with the "playpen" formula, is more satisfied that the other, and so is preferred. To put the matter simply: the application (backwards) of the inference rules prefers a wrapper of objects (the playpen) to a wrapper of liquids (writing pen).

This method can seem ad hoc, designed to deal with a classic cleaver example, and it is. But notice, it cannot be outmanoeuvred by the repeated-writing-marker "inkstand in pen" point of Bar Hillel's mentioned earlier, since the writing pen will also be dispreferred as a wrapper of an object like an inkstand, given that the writing pen prefers to wrap liquids. Furthermore, this solution is somewhat general, for the ambiguity is resolved by the application of the general rules of preference used in setting up the representation, and not in any sense by special rules for the example. Although nothing follows from any particular example in this field, this use of general principles of language that set up the representation itself is, I would argue, a more promising approach to the traditional MT problem than either (a) very large knowledge structures, like frames, that are difficult to motivate and manipulate or the other suggested alternative for this example which is (b) ad hoc size categorizations of physical objects on some 1–10 scale, so that, say, boxes would be 4, writing pens 2 and playpens 5 in which case unaided arithmetic could yield the correct result, since $4 > 2$ and $4 < 5$ (a solution once seriously suggested by Horace Enea).

The conclusion so far is that we should be optimistic, and that AI semantic systems based on general principles like "preference" may relieve at least some pressing MT difficulties, and without claiming to be able to represent all the knowledge in the universe. For about that and its intractability Bar-Hillel was surely right.

http://www.springer.com/978-0-387-72773-8

Machine Translation
Its Scope and Limits
Wilks, Y.
2009, X, 252 p., Hardcover
ISBN: 978-0-387-72773-8