Frames for machine translation

This second of two articles on machine translation discusses a recent concept that may help the inanimate computer "understand" the often apparently irrational utterings of human language.

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 is information 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, Eugene Char-
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 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 would infer only that a physical object was ordered.

Frames or scripts, therefore, will certainly help in determining topic or overall context, provided that we can reliably decide in advance what is the appropriate frame with which to analyse a 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 enormous 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"). Later, problems arise as to when to stop following a script and get rid of it in favour of another.

Claims by frame users

The real issue, though, is not technical but concerns what claims are being made by frame users. They are, I think, making a plot line hypothesis: "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." As Charniak 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—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 exorts CROWD (on proper behaviour)
(e) (general condition)
Bad behaviour by CROWD → Activity halted
(f) ELDER checks if CHILD properly purified
(g) (special condition)
CHILD not purified activity halted
(h) ELDER marks CHILD'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 "story sentences": "Little Kimathis's mother accidentally dropped her shoga looked away and touched his arm during the puberty rite. The crowd drew back in horror." If we wish to "understand" this story, do we need the frame above to do it? The frame covers the story with line (e) in some sense, given an adequate list defining bad behaviour accessible 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 Kimathis during the ceremony was a 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 that 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 a special 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 an important point, this 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 contained explicitly all and only the knowledge in the frame about X, than if, say, only 5 per cent of texts about X did so.

Another possibility that I have put forward is the use for understanding of more static frames, such as one about the functioning and use of say, a car. These would be used to deal with the frequent occasions in text where the writer ignores the normal "semantic preferences" of words. This phenomenon is often called metaphor, as in "My car drinks petrol" which we understand perfectly well, even though any semantic description of "drink" would declare it to be normally an action done by animate things. Yet this "boundary breaking" sentence is perfectly comprehensible and of a type we might expect to find two or three times in any randomly chosen newspaper paragraph.

The suggestion is that the "car frame" might contain some formalised expression of such facts as "humans inject liquid (petrol) into a car" and "the car's engine uses liquid". Given a sophisticated matching process for semantic representations we might be able to match the representation for "My car drinks petrol" to one or both of these frame statements, and infer that, in understanding the sentence, we would project the (frame) sense of "use" onto that of "drink" in the sentence, thus changing the underlying representation of it to something like "My car uses petrol". This representation would now be closer to a literal meaning that the computer system could manipulate. An application of frames along these lines is being programmed, and might conceivably be brought to bear upon metaphor, a long intractable problem in the world of machine translations.