

Controlling Inference in Story Understanding

Richard E. Cullingford
Department of Engineering and Applied Science
Yale University
New Haven, Connecticut 06520

I. A Story Understander

This paper sketches a method of pattern-matching and inference applied by a program called SAM [1] to read newspaper articles about car accidents, state visits, train wrecks, etc. Here is an example of a story SAM has processed, which illustrates the depth of comprehension this method enables it to achieve:

Sunday morning Enver Hoxha, the Premier of Albania, and Mrs Hoxha arrived in Peking at the invitation of Communist China. The Albanian party was welcomed at Peking Airport by Foreign Minister Huang. Chairman Hua and Mr Hoxha discussed economic relations between China and Albania for three hours.

Summary:

PREMIER ENVER HOXHA, THE ALBANIA GOVERNMENT HEAD, AND CHAIRMAN HUA KUO-FENG, THE CHINA GOVERNMENT HEAD, DISCUSSED CHINA ALBANIA ECONOMIC AFFAIRS IN PEKING, CHINA TWO DAYS AGO.

Question-Answering:

Q1: Who went to China?

A1: NADIA AND ENVER HOXHA WENT TO COMMUNIST CHINA.

Q2: How did Enver Hoxha and Mrs Hoxha get to China?

A2: THEY FLEW TO IT.

Q3: Why did they go to China?

A3: ENVER HOXHA AND HUA KUO-FENG WANTED TO DISCUSS CHINA ALBANIA ECONOMIC AFFAIRS.

SAM reads stories like this by introducing a frame-like data structure called a Script [2] when the first sentence is analyzed, and by finding subsequent inputs in this context via the expectations that are progressively aroused. The recognition process is driven by a pattern-match of the input conceptualization against a template stored in the Script. (SAM works internally with meaning structures coded in the Conceptual Dependency system [2].)

II. Making Connections

SAM constantly makes inferences during this process, both to fill in events which have been left out between the lines of a story, and to

This work was supported in part by the Advanced Research Projects Agency of the Department of Defense and monitored under the Office of Naval Research under contract N00014-75-C-1111

reconcile discrepancies between an input and the active expectations. Some of these inferences are illustrated by the example story.

The process which interconnects events explicitly contained in a text is called causal chain instantiation: the answer to Q1 above, e. g., depends on the information that arriving in a city must be preceded by entering the country the city is part of. Sometimes a role instantiation inference is needed to specify an actor or object not explicitly mentioned: e. g., that dignitaries usually fly to foreign countries. Reference specification is a standard Inference SAM makes: to decide, e. g., that "the Albanian party" can be the same as "Enver and Mrs Hoxha". The time- and place-setting of a story is also used: here, to infer that the official talks probably took place in Peking, on the day of the arrival. (SAM inserts the phrase "two days ago" because it is arranged, by convention, to be reading stories on Tuesday.)

Other inferences are available to iron out small differences between what was expected and what was actually received, two of which are illustrated in the story. SAM computes immediate results of Script actions, e. g., to understand that being "at Peking Airport" is a result of being flown there. A related class of movement inferences specifies the probable means by which people and their possessions travel around. In the story, these inferences would be responsible for asserting that the Hoxhas arrived in Peking on some sort of commercial conveyance, later discovered to be an airline; and that their possessions, e. g., their money, arrived with them.

III. Conclusions

What do we believe this work has accomplished? First of all, we were able to define a computationally usable notion of context, the Script; and to devise a computer program, SAM, which uses a data base of Scripts to achieve a reasonable depth of understanding of certain kinds of newspaper stories. Along the way, several types of inference had to be incorporated to enable SAM to make connections between the ideas in a story; and to recognize successfully story inputs which deviate from the expected in various ways. These kinds of problems crop up in stories of all kinds. Therefore, SAM's pattern-match-and-inference cycle provides a model for one kind of processing a general story understander will have to perform.

References

- (1) Cullingford, R. E., Organizing World Knowledge for Story Understanding by Computer, Ph.D. Thesis, Dept. of Engineering and Applied Science, Yale University, 1977.
- (2) Schank, R. C., and R. P. Abelson, Scripts, plans, goals and understanding, Hillsdale, New Jersey: Lawrence Erlbaum Press, 1977.