THE CLOWNS MICROWORLD*

Robert F. Simmons
Department of Computer Science
University of Texas

ABSTRACT

About fifteen years of active research in natural language question-answering systems has provided reasonably concise and elegant formulations of computational semantics for understanding English sentences and questions about various microworlds. These include the Woods Lunar Data Base, the Winograd world of a pictured hand and blocks, the Heidorn world of a fueling station, the Hendrix, Slocum, Thompson world of transactions, John Seely Brown's power circuit and Schank's sketches of motivated humans. (See Woods et al 1972, Winograd 1972, Hendrix et al 1973, Heidorn 1972, Schank 1975 and Brown et al 1974.) In each of these worlds, a natural language processor is able to understand an ordinary subset of English and use it conversationally to accept data and to respond to commands and questions.

Ignoring early work largely lost in the archives of corporate memos, Winograd's language processor is essentially a first reporting of how to map English sentences into diagrammatic pictures. Apart from potential applications, the pictures are of great value in providing a universally understood second language to demonstrate the system's interpretation of the English input. While we are still struggling in early stages of how to compute from English descriptions or instructions, there is much to be gained from studying the subset of English that is picturable. Translation of English into other more general languages such as predicate calculus, LISP, Russian, Basic Engish, Chinese, etc. can provide the same feedback as to the system's interpretation and must suffice for the unpicturable set of English. But for teaching purposes, computing pictures from language is an excellent instrument.

We began with the notion that it should be quite easy to construct a microworld concerning a clown, a pedestal and a pole. The resulting system* could draw pictures for such sentences as:

A clown holding a pole balances on his head in a boat.

A clown on his arm on a pedestal balances a small clown on his head.

Figure 1 shows examples of diagrams produced in response to these sentences.

----

*Supported in part by NSF Grant GJ509E
*(see Simmons & Bennett-Novak 1975 for grammar and semantics of this system.)

We progressed then to sentences concerning movement by adding land, water, a lighthouse, a dock and a boat. We were then able to draw pictures such as Figure 2 to represent the meanings of:

A clown on his head sails a boat from the dock to the lighthouse.

In the context of graphics, two dimensional line drawings are attractive in their simplicity of computation. An object is defined as a LOGO graphics program that draws it (see Papert 1971). A scene is a set of objects related in terms of contact points. A scene can be described by a set of predicates:

(BOAT ABOVE WATER)
(ATTACH $BOAT_{xy}$ $WATER_{xy}$)
(DOCK ABOVE WATER) (DOCK LEFTOF WATER)
(BOAT RIGHTOF DOCK)
(ATTACH $DOCK_{xy}$ $WATER_{xy}$)
(ATTACH $BOAT_{xy+ky}$ $DOCK_{xy}$)

Orientation functions for adjusting starting points and headings of the programs that draw the objects are required and these imply some trigonometric functions. A LISP package of about 650 lines has been developed by Gordon Bennett-Novak to provide the picture making capability.

What is mainly relevant to the computation of language meanings is that a semantic structure sufficient to transmit data to the drawing package is easily represented as a property list associated with an artificial name for the scene. For example, "A CLOWN ON A PEDESTAL" results in the following structure:

(C1, TOK CLOWN, SUPPORTBY C2, ATTACH(C1 FEETXY C2 TOPXY))

(C2, TOK PEDESTAL, SUPPORT C1, ATTACH(C2 TOPXY C1 FEETXY))

(CLOWN, EXPR(LAMBDA()...) FEET XY, SIZE 3, STARTPT XY, HEADING A)

(PEDESTAL, EXPR(LAMBDA()...) TOP XY, SIZE 3, STARTPT XY, HEADING A)

A larger scene has more objects, more attach relations, and may include additional relations such as INSIDE, LEFTOF, RIGHTOF, etc. In any case the scene is semantically represented as a set of objects connected by relations in a graph (i.e. a semantic network) that can easily be stored as objects on a property list with relational attributes that connect them to other such objects.

A small grammar rich in embedding capabilities is coded in Woods' form of Augmented Transition Net (Woods 1970) for a set of ATN functions to interpret. As each constituent is completed the operations under the grammar arcs create portions of property list structure. When a clause is completed, semantic routines associated with verbs and prepositions sort the various Subject Object and Complement constituents into semantic roles and connect them by semantic relations. A verb of motion creates a net of relations that are valid in all timeframes and in addition encodes a process model that changes the semantic net from one timeframe to another.

Nouns such as "clown", "lighthouse", "water", etc. are programs that construct images on a display screen. Other nouns such as "top", "edge", "side", etc. are defined as functions that return contact points for the pictures. Adjectives and adverbs provide data on size and angles of support. Prepositions and verbs are defined as semantic functions that explicate spatial relations among noun images. Generally, a verb produces a process model that encodes a series of scenes that represent initial, intermediate and final displays of the changes the verb describes.

The system is programmed in UTLISP for CDC equipment and uses an IMLAC display system. It currently occupies 32K words of core and requires less than a second to translate a sentence into a picture.

DISCUSSION

Nouns such as "circus", "party", "ballgame" etc. have not yet been attempted. They imply partially ordered sets of process models and are the most exciting next step in this research. More complex verbs like "return" or "make a roundtrip" imply a sequence of interacting process models. Thus, "a clown sailed from the lighthouse to the dock and returned by bus" offers interesting problems in discovering the arguments for MOVE*-return as well as in the design of a higher level process model whose intermediate conditions include the models of MOVE*-sail and MOVE*-return.

As it stands, the CLOWNS system has served as a vehicle for developing and expressing our ideas of how to construct a tightly integrated language processing system that provides a clearcut syntactic stage with coordinate semantic processing introduced to reduce ambiguity. Two stages of semantic processing are apparent; the first is the use of prepositions and verbs to make explicit the geometric relations of "support", "leftof", etc. among the objects symbolized by the nouns; the second is the transformation of these geometric relations into connected sets of x-y coordinates that can be displayed as a scene. Schank's notion of primitive actions is reflected in our approach to programming high level verbs such as MOVE* to encompass the idea of motion carried in verbs such as "sail", "ride", etc. Woods' ATN approach to syntactic analysis is central to this system and in sharp contrast to the approach of Schank and Riesbeck who attempt to minimize formal syntactic processing. Our process model reflects the ideas developed by Hendrix (1974) in his development of a logical structure for English semantics.

The system is not limited to its present grammar nor to its present vocabulary of images. Picture programs to construct additional objects are easily constructed and the semantic routines for additional verbs and prepositions can be defined for the system with relative ease.

The system has been used successfully to communicate methods for natural language computaton to graduate students and to undergraduates. It appears to have immediate possibilities for teaching the structure of English, for teaching precision of Engish expression, and for teachng foreign languages through pictures. Eventually it may be useful in conjunction with very good graphic systems for generating animated illustrations for picturable text.

In my mind CLOWNS shows the power and value of the microworld approach to the study of Artificial Intelligence. By narrowing one's focus to a tiny world that can be completely described, one can define a subset of English in great depth. This is in contrast to the study of text where the situations described are so complex as to forbid exhaustive analysis. The translation into a visualized microworld provides an immediate display in a two-dimensional language of the interpretations dictated by the syntactic and semantic systems and thus a scientific measuring instrument for the accuracy of the interpretation.

Although there is potential for expansion of the system into the world of useful applications, I believe the primary value of this experiment with the CLOWNS world is to show that there exist orderly and straightforward ways of economically computing translations from subsets of English to procedures that do useful work. This is not a new finding but I believe the implementation is considerably simpler than most previous ones.

REFERENCES

Brown, Burton, R.R. & Bell, A.G., "SOPHIE: A Sophisticated Instructional Environment for Teaching Electronic Troubleshooting", BBN Report # 2790, April 1974.

Heidorn, George E., "Natural Language Inputs to a Simulation Programming System," NPS-55HD, Naval Post Graduate School, Monterey, Calif. 1972.

Hendrix, G., "Preliminary Constructs for the Mathematical Modeling of English Meanings." University of Texas, Department of Computer Sciences, Working Draft, April 1974. (not for distribution)

Hendrix, G.G., Thompson, Craig and Slocum, Jonathan. "Language Processing via Canonical Verbs and Semantic Models." _Proc. 3rd Int. Jt. Conference on Artificial Intelligence_, Stanford

University, Menlo Park, Calif., 1973.

Papert, S., "Teaching Children to be Mathematicians vs. Teaching About Mathematics." Int. J. Math. Educ. in Science & Tech., New York: Wiley & Sons, 1972; MIT, A.I. Memo. No. 249, July 1971.

Schank, Roger, Conceptual Information Processing, North-Holland Publishing Company 1975 (In Press).

Simmons, R.F. and Bennett-Novak, G., "Semantically Analyzing an English Subset for the Clowns Microworld", Dept. Comp. Sci. Univ. Texas, Austin, 1975.

Winograd, Terry, Understanding Natural Language, New York: Academic Press, 1972.

Woods, W.A., Kaplan, R.A., & Nash-Webber, B., "The Lunar Sciences Natural Language Information System: Final Report: BBN Report # 2378, June, 1972, Bolt Beranek and Newman Inc., Cambridge, MA.

Woods, Wm. A., "Transition Network Grammars for Natural Language Analysis," Comm. ACM, 13, Oct. 1970.

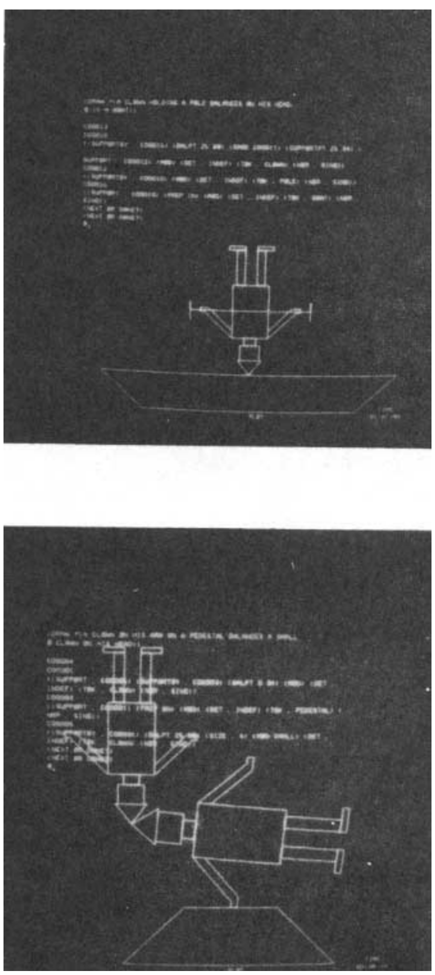

Figure 1. State Verbs

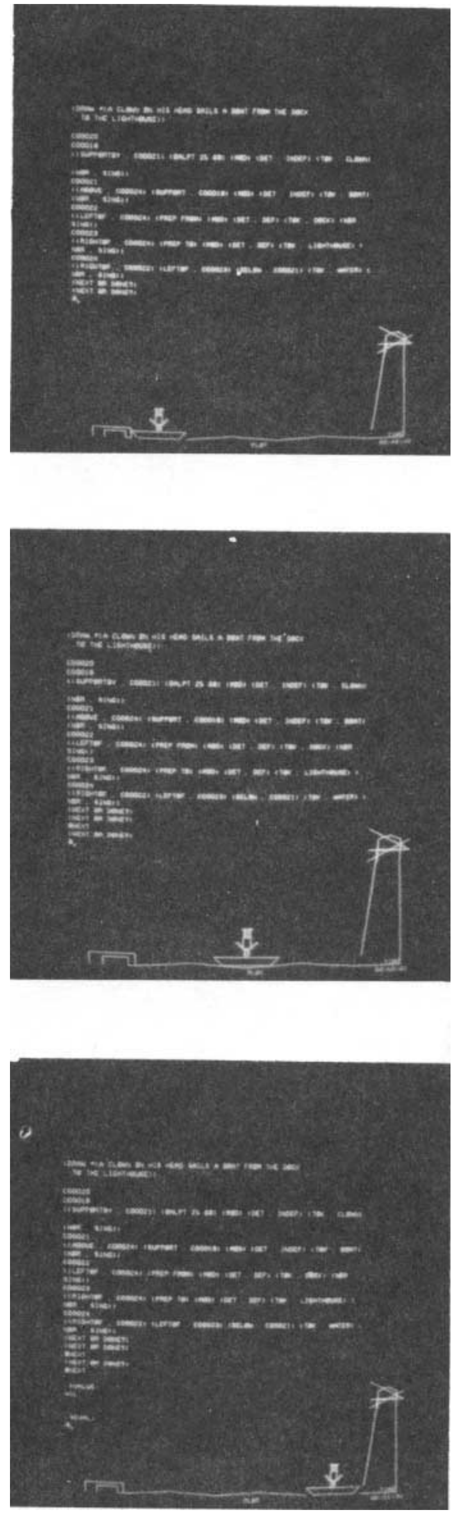

Figure 2. A Motion Verb