EVENT SHAPE DIAGRAMS

D. L. Waltz

Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, IL 61801 U.S.A.

ABSTRACT

"Event shape diagrams" are proposed as a representation for capturing the nuances of meaning of verbs that describe similar events. These diagrams represent timing, causal relationships between case roles, and typical value ranges for role fillers. Event shape diagrams are expressed in terms of primitive predicates and timing information that we believe could be computed by perceptual systems, and are intended to be a step toward the eventual connection of language systems to perceptual (vision, hearing, and touch) sensing systems. The diagrams are capable of representing modification of verbs by adverbs, can support judgements of the degree of plausibility of various interpretations of a sentence's meaning, and may be useful in figuring out the meaning of certain kinds of metaphors.

A. Introduction

A satisfactory representation scheme for natural language sentence meaning must offer distinguishably different meaning structures for sentences that people judge to differ in meaning, and at the same time it ought to represent similar sentences similarly. For the most part, natural language representation systems have been much more concerned with capturing the similarity of meaning of various verbs than with representing the nuances in meaning. For example, in Schank's Conceptual Dependency (CD) representation system [1], eat, eat up, overeat, nibble, gulp, wolf, drink, swig, swallow, inject, and smoke (e.g. a cigar) are all represented by structures centered around the "primitive" INGEST. After processing sentences involving these different verbs, the only trace of difference in the meaning representation would be that the objects of the verb would differ; thus, certain objects are liquids (e.g. milk, beer, coffee) while others are non-liquid (e.g. hot dogs, jello, broccoli), drugs (e.g. insulin, heroin) or smokable substances (e.g. tobacco, pot) and a system that knew the nature of various objects might, at least in some cases, be able to differentiate between drinking, eating, injecting, or smoking types of INGESTs.

There is much to be said for lumping these meanings together. Most importantly, similar inferences can be made about such events, for example that the agent's desire for the object may be lessened after INGESTing the object, that the object of INGEST is inside the agent after INGESTing, and so on. However, certain important distinctions are necessarily lost: a system that simply substitues INGEST for all the verbs above will be incapable of making certain kinds of predictions or inferences that a speaker would expect a person to make. For example, compare (1) "John nibbled at some food" with (2) "Mary wolfed some food". As human listeners or readers, we could be expected to infer from (1) that John was either did not feel very hungry or that he didn't like the food very much, and from (2) that Mary was probably very hungry or in a hurry. In the proper context, we could even be expected to understand from (1) that John was upset or depressed, and from (2) that Mary had an urgent task to carry out. While I have only used examples that involve INGEST, similar arguments can be made concerning all the CD primitives.

There are other problems with CD as well. As pointed out by Wilks [2], the set of CD primitives is incomplete -- there are no CD primitives for many actions (e.g., divide, construct, bend, join, fasten); and a very wide class of verbs is lumped under the STATE CHANGE primitive. For example, break, chip, crack, destroy, damage, and scratch would all be represented very similarly, as STATE CHANGES to a state which is negative with repect to the original state. Moreover, unless I have missed something important, the CD representations for get sick, be injured, go crazy, fall down, and be indicted could differ from each other and from the previous examples only in the degree of negativity of the STATE CHANGE.

In recent years, Schank et al. have shown little interest in repairing the shortcomings of CD, but have concentrated instead on developing larger memory structures, in particular scripts [3], and MOPS [4], each of which may contain a large number of CD structures. These bear on the problems above to some degree. For instance, if eat were potentially attached to a script, then

---

This work was supported in part by the Office of Naval Research under contract N00014-75-C-0612, and in part by the National Science Foundation under grant NSF IST 81-17238.

from the representation for "John ate a meal," a CD-based system could infer that both eating and drinking probably occurred. However, the representation of nuances in general requires a finer-grained representation, not scripts or other larger structures.

I do not want the preceding remarks to be taken as an attack on Schank et al. In fact, I have used these examples because CD is probably the most advanced system for representation of actions that has been developed by AI! (The only other candidates that spring rapidly to mind are the LNR representation schemes [5], Wilks' preference semantics [6], and in a sense, Lehnert's emotion modeling constructs [7].)

B. Other Related Background

Rieger's "common sense algorithm" (CSA) work [8] attempted to enumerate causal relationships between states, actions, and tendencies (such as gravity) for use in representing the operation of physical mechanisms. While very promising in certain ways, CSA diagrams have never been well-integrated with CD, and have not been able to represent timing, quantitative state variable values, concurrency or hierarchical relationships in a satisfactory way. Furthermore, the construction of CSA diagrams is still an art. More recently Forbus [9] has developed QPT ("qualitative process theory"), a very promising, interesting, and natural body of programmable methods for reasoning about the behavior of physical objects and mechanisms. While Forbus has not to my knowledge attempted to use his work for representing the meaning of languages, the possibilities for doing so are very promising, and I intend to explore them. The spirit of QPT is, I believe, similar to, and compatible with, event shape diagrams.

Adverbs (e.g., quickly, softly, hard, suddenly) have seldom been mentioned in AI papers on natural language understanding. When they have been mentioned [10] they have been viewed as difficult or impossible to deal with. We have developed mechanisms for dealing with adverbs [11] and have recently improved upon these mechanisms. Many adverbs (including the ones above) can be represented very naturally in event shape diagrams, though certain manner adverbs (e.g., viciously, kindly, bluntly) still seem difficult.

Recent work by Lehnert on summarizing narratives [12], by Allen on a "temporal logic" [13], by Abelson on the relationships between events, actions, plans, and emotions [14], and by Talmy on the relationships between grammar and perception [15] have all played a part in the development of our "event shape diagrams," summarized below.

C. Event Shape Diagrams

In their simplest forms, event shape diagrams have a time line, a scale, and values on the scale at one or more points.*

Diagrams can be used to represent concurrent processes, causation, and other temporal relations by aligning two or more diagrams, as illustrated in Fig. 1.

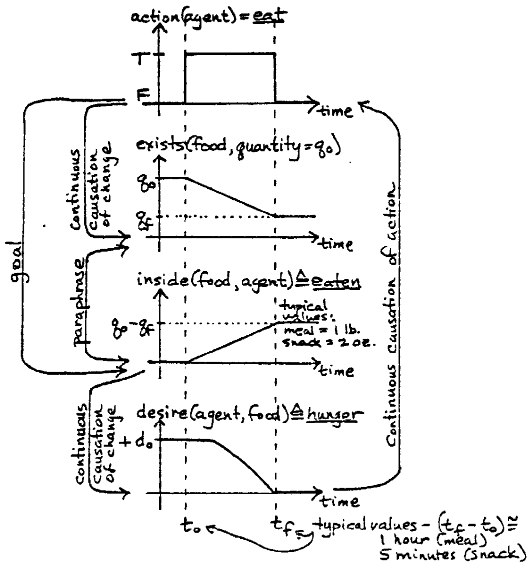

Fig. 1. Event Shape Diagrams for "Eat".

Fig. 1 shows the representation for "eat." Note that four simple diagrams are aligned, and that each has different kinds of scales, and different event shapes. The top scale corresponds to the CD primitive INGEST. Causal relations hold between the events described in each simple diagram. Some of the names for the causal relations are adopted from Rieger's CSA work. The action INGEST stops in this default case where "desire to eat" goes to zero. "Desire to eat" sums up in one measure coercion, habit, and other factors as well as hunger. Typical values for amounts of food, time required to eat, and so on are also associated with the diagram, to be used as default values.

More levels of detail can be added if needed. For instance, the action diagram can be expanded so that eating involves many recurrences of putting food in one's mouth, biting, chewing, and swallowing, and the diagram for the amount of food inside the agent can reflect a series of stepwise changes as each mouthful is ingested. In the other direction, eat should point to diagrams representing the normal cycle of eating. In the direction of greater detail, I believe that diagrams should eventually "bottom out" in primitive predicates such as contact, surround, be near, and support,

*While diagrams are shown here for ease of understanding, data structures must of course be constructed to represent these diagrams. The data structures borrow heavily from Allen [13] for representing before, after, while, etc.

which could be computed for scenes by a vision system; this goal is similar to that of Miller and Johnson-Laird [16].

Many adverbial modifiers can be represented neatly: "eat quickly" shrinks the value of $t_f - t_o$ with respect to typical values; "eat a lot" increases the values of $q_o - q_f$ above typical values. Similarly "eat only half of one's meal," "eat very slowly," "eat one bite," etc. can be neatly represented.

The point of time from which events are viewed can also be clearly represented. Past tense (e.g. "we ate 3 hamburgers") puts "now" on the time line to the right of the action, while future tense puts "now" to the left of the action, and present progressive (e.g. "we are eating") puts "now" between t_o and t_f.

AI systems have dealt with models of human belief, expectation, and attitude in only quite simple situations (e.g. [17]). While it is premature to make grandiose claims, the examples I have worked on so far have not presented great difficulties for our event shape diagram formulations. For example, Fig. 2 shows the representation of the apparently rather hard sentence, "I was surprised that John ate so much."

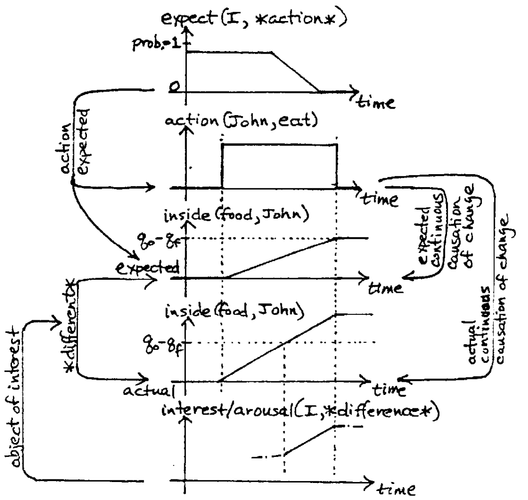

Fig. 2. Representation of "I was surprised that John ate so much."

The structure in Fig. 2 uses the portion of the preceding meaning for eat that is selected by the pattern eat + <quantity>. "Be surprised by" has slots for an agent, an expected event or action, and an actual event or action, which must differ. In this case, we know that John actually ate more than he was expected to eat, so we can fill in his actual behavior in some detail. If the sentence were instead "I was surprised at how much John ate," it is also possible that John ate less than he was expected to, and the representation could not be constructed until the ambiguity was resolved. A possible representation could simply note that the actual and expected amounts were different. The "interest arousal" scale shows up as part of the meaning of many verbs, such as <u>like</u>, <u>enjoy</u>, <u>hate</u>, <u>pay attention to</u>, <u>desire</u>, <u>fear</u>, and so on.

D. <u>Metaphor with Event Shape Diagrams</u>

Metaphors can be used to transfer complex combinations of information from one well-known domain to another less well known or completely unfamiliar one. Understanding metaphorical language first requires noting that the language <u>is</u> metaphorical, that is that it couldn't be literal descriptive text. This in turn requires an internal model of what is ordinary, expected, or possible that a system can use to judge the plausibility of novel language -- this sort of information is represented in event shape diagrams by the attached typical values for various scales (see Fig. 1). Next, links must be established between the base domain of the metaphor and the target (novel) domain that the metaphor is being used to describe. The result can become the basis for learning about a new domain (by transferring knowledge from the base domain selectively) or it may simply be that a metaphor allows one to express in a few words many notions about a target domain that would otherwise require a much lengthier exposition (see Ortony [18]). Consider for example:

(S1) John ate up the compliments.

or

(S2) Robbie's legs ate up the space between him and Susie.

Assuming that these sentences represented novel uses of the words "ate up" for our system, we might want the system to infer that in the first sentence John desired the compliments, eagerly "ingested" them with his mind, thereby making them internal and being given by them, and that in the second sentence, the distance between Robbie and Susie was being reduced to zero, just as an amount of food is reduced to zero when it is "eaten up".

We have been developing methods for making the correct inferences in examples like these (see DeJong and Waltz [19] for details). The methods depend on matching of the meanings structures, similar to the schema matching of Winston [20]. Object meanings are taken to be literal, and are used to suggest candidate verbs for matching. In (S1) "compliments" suggests "tell" or "hear" as the "true verbs," whereas in (S2) "space" suggests verbs with underlying PTRANS meaning as the "true verb." For (S1), then, we would match "eat up" with "tell" and "hear," select "hear" as a better match, and suggest that the "residue" of eat up (e.g. the portion of its meaning not matched) is intended to be transferred to "hear." Thus, we could infer that John desired to hear the compliments, and felt pressure as a result.

E. *Conclusion*

This work is just beginning. It seems promising, but already it should be clear that event shape diagrams are *not* a completely general representation, but that they are especially useful for representing multiple concurrent processes and functional/causal relationships. Research will continue on the use of event shape diagrams for representation as well as on their range of coverage.

REFERENCES

[1] Schank, R.C. "The Primitive ACTs of Conceptual Dependency." In *Theoretical Issues in Natural Language Processing*. R. Schank, B. Nash-Webber (eds.). Arlington, VA: ACL, 1975.

[2] Wilks, Y.A. "Primitives and Words." In *Theoretical Issues in Natural Language Processing*. R. Schank, B. Nash-Webber (eds.). Arlington, VA: ACL, 1975, 38-41.

[3] Schank, R.C. and R. Abelson. *Scripts, Plans, Goals, and Understanding*. Hillsdale, NJ: Lawrence Erlbaum Assoc., 1977.

[4] Schank, R.C. "Reminding and Memory Organization: An Introduction to MOPs." Research Report 170, Computer Science Dept., Yale University, December 1979.

[5] Rumelhart, D.E. and D.A. Norman. "Active Semantic Networks as a Model of Human Memory." In *Proc. IJCAI-73*. Stanford University, Stanford, CA, August 1973, pp. 450-547.

[6] Wilks, Y.A. "A Preferential, Pattern-Seeking, Semantics for Natural Language Inference." *Artificial Intelligence* 6:1 (1975) 53-74.

[7] Lehnert, W.G., M.G. Dyer, P.N. Johnson, C-J. Yang, and S. Harley. "BORIS -- An Experiment in In-Depth Understanding of Narratives." Technical Report 188, Computer Science Dept., Yale University, January 1981.

[8] Rieger, C. "The Commonsense Algorithm as a Basis for Computer Models of Human Memory, Inferene, Belief and Contextual Language Comprehension." In *Theoretical Issues in Natural Language Processing*. R. Schank, B. Nash-Webber (eds.). Arlington, VA: ACL, 1975, 180-195.

[9] Forbus, K. "Qualitative Process Theory." AI Memo No. 664, MIT AI Lab, Cambridge, MA, February 1982.

[10] Cercone, N. "A Note on Representing Adjectives and Adverbs." In *Proc. IJCAI-77*. MIT, Cambridge, MA, August 1977, pp. 139-40.

[11] Waltz, D.L. "Toward a Detailed Model of Processing for Language Describing the Physical World." In *Proc. IJCAI-81*. Vancouver, B.C., Canada, August 1981, pp. 1-5.

[12] Lehnert, W.G., J.B. Black and B.J. Reiser. "Summarizing Narratives." In *Proc. IJCAI-81*. Vancouver, B.C., Canada, August 1981, pp. 184-189.

[13] Allen, J.F. "An Interval-Based Representation of Temporal Knowledge." In *Proc. IJCAI-81*. Vancouver, B.C., Canada, August 1981, pp. 221-226.

[14] Abelson, R. Personal communication. 1981.

[15] Talmy, 1978 Talmy, L. "The Relation of Grammar to Cognition - A Synopsis." In *TINLAP-2: Theoretical Issues in Natural Language Processing-2*. D.L. Waltz (ed.). Association for Computing Machinery and The Association for Computational Linguistics, 1978, pp. 14-24.

[16] Miller, G.A. and P. Johnson-Laird. *Language and Perception*. Cambridge, MA: Harvard University Press, 1976.

[17] Cohen, C.R. and P.R. Perrault. "It's for Your Own Good: A Note on Inaccurate Reference." In *Elements of Discourse Understanding*. Joshi, Sag, Webber (eds.). Cambridge, England: Cambridge Univ. Press, 1981, 217-230.

[18] Ortony, A. "Why Metaphors are Necessary and Not Just Nice." *Educational Theory* 25 (1975) 45-53.

[19] DeJong, G. and D.L. Waltz. "Understanding Novel Language. *Intl. J. Computers and Mathem.*, special issue, to appear (1982).

[20] Winston, P.H. "Learning New Principles from Precedents and Exercises: The Details." MIT AI Memo No. 632, Cambridge, MA, May 1981.