The Primitive ACTs of Conceptual Dependency

Roger C. Schank
Yale University
New Haven CT

How should a proposition appear in memory? It simply cannot be stored in a form that bears a strong relationship to the original form of the input in natural language. Since it is possible to say the same thing in any number of different ways, it is unreasonable to suppose that people are constantly checking to see if a proposition that they have stored one way in memory is the same as another they have stored somewhere else in a different form. What is far more likely is that people have developed a kind of canonical form. No two propositions in memory can be equivalent in this view, since the canonical form for their meaning forces them to be represented in the same way and thus stored and treated in the same way by memory.

A canonical representational scheme cannot have natural language words as its elements. Words can be ambiguous and a meaning representation system must be unambiguous. Furthermore, words can be synonymous and overlap in meaning. By disallowing synonymous canonical forms of propositions we have also disallowed the use of words in the representation.

The consequences of this are consistent with what we know about people. In memory experiments people consistently exhibit recognition confusions because of their remembrance of the "gist" of a sentence rather than the words that they actually saw or heard.

Conceptual Dependency has as its basic premise that the ACTION is the basis of any proposition that is not descriptive of a static piece of the world.

All propositions that describe events are made up of CONCEPTUALIZATIONS. A conceptualization is composed of an ACTOR, an ACTION, and a set of cases that are dependent on that ACTION.

The main insight of Conceptual Dependency theory is that there must be a canonical form for meaning representations. The types of conceptual roles that there are in this canonical form do not correspond to syntactic roles. Rather, words are broken down into their conceptual parts and then placed in a meaning representation with respect to their conceptual role rather than the syntactic role that was chosen for them as a result of some, perhaps random, lexical choices. For example, the word 'danger' has a nominal form, a verbal form (endanger), an adjectival form (dangerous), and an adverbial form (dangerously). The meaning of the concept underlying 'danger' is the same no matter what lexical form is selected to realize that concept. Basically, 'danger' means that something can possibly happen that will result in a negative effect in somebody's physical state. 'Danger' is represented conceptually as a conceptualization linked to a state description. Certain elements of the conceptualization are unknown, but can be filled in by the words surrounding 'danger'. The state description is partially known. It says that somebody's physical state may be negatively changed.

So we see that 'endanger' is no more an action than it is an actor. Rather, 'endanger' refers to a partial conceptual structure. The word 'endanger' also sets up an expectation that the syntactic subject of that verb will contain the item that will help fill in the missing action that was 'dangerous'. Thus, in the sentence "The bees endangered Bill," the 'bees' are the actors of a conceptualization that is causally related to the state description that the concept 'Bill' will undergo a negative change. The action is still unknown, but now can be inferred. The memory is consulted for actions that bees are known to perform that endanger humans. The simplest explanation is preferred and the best candidate for the action is 'sting'. The object of that action is 'Bill'. The conceptual representation of "The bee's endangered Bill" says that "the bees may do an action (possibly 'sting') which will cause Bill to become hurt."

We said before that words do not appear in conceptualizations. Following that point of view it is obvious that 'sting' cannot really satisfy our needs here. The idea of 'sting' is correct, but it must be broken down into its component parts. We define an action as something an actor can do to an object. We define an actor as an animate object, and an object as any concrete physical object. 'Sting' still works under these constraints, but couldn't we just have well used the word 'bite'? Although these two words are slightly different in their strict interpretations, to the average user, these words are synonyms. We are left with the problem of describing the concept underlying them in order to satisfy the requirements that we have set out above.

In order to handle similarities and overlaps in meaning as used by the man in the street we have developed the concept of the PRIMITIVE ACTION. Primitive actions are intended to be the building blocks out of which the meanings of verbs and abstract and complex nouns are constructed. The primitives are not category names but rather elements that can be used in many varied combinations to express the meaning of what underlies a given word.

Conceptual Dependency representations use only twelve primitive acts. These acts were arrived at by noticing structural similarities that existed when sentences were put into an actor-action-object framework. Using this set of acts, a larger set of states and a set of conceptual roles, it is possible to express a large amount, if not all, of the meanings expressible in natural language. We will now describe the most important of the primitive ACTs and

briefly mention the others.

ATRANS The transfer of an abstract relationship such as possession, ownership or control. ATRANS requires an actor, object and recipient.
ex.
give=ATRANS an object from the actor to the recipient.
take=ATRANS an object from someone to the actor.
buy=two ATRANS actions, each causing the other. An ATRANS of money from the first actor to the second actor. An ATRANS of an object from the second to the first.

PTRANS The transfer of physical location of an object. PTRANS requires an actor, object and direction.
ex.
go=PTRANS of an actor by an actor to a location.
fly=PTRANS of an actor to a location. Any ACT can have another act as its instrument. For 'fly' that instrument is PROPEL.

PROPEL The application of a physical force to an object. If movement takes place as a result of a PROPEL, a PTRANS has taken place too if it can be determined that that PROPEL was intended by the actor. Direction, object and actor are required.
ex.
push=PROPEL an object in a direction by an actor.
pull=PROPEL an object in a direction (towards actor) by an actor.
kick=PROPEL an object in a direction by an actor using the instrument MOVE object foot.

MTRANS The transfer of information between animals or within an animal. For the purpose of dealing with the words people use to discuss memory, we partition memory into three pieces: the CP (the conscious processor9. Something being in CP means that it is being thought of at that moment. The IM (the intermediate memory) where current context is kept and contextually relevant information is stored while in use. The LTM (long term memory) where informaton is stored. The various sense organs serve as the origins in an MTRANS. Recipient, object and actor are required, although here these cases refer to the above mental locations as recipients and mental information as possible objects.
ex.
remember=MTRANS information from LTM to IM
forget=MTRANS cannot be accomplished on some information from LTM
see=MTRANS information to IM from the eye
tell=an actor MTRANS information to another actor
read=MTRANS information from a book to IM by means of ATTEND eye to the book.

MBUILD The construction within an animal of new informaton from old information. MBUILDing takes place within IM and receives its inputs from the CP and from IM. It transforms them into a new idea and places that idea in the CP.
ex.
describe=MBUILD a future action or a new state.
imagine=MBUILD a new idea that is hypothetical
consider=MBUILD from a mentioned input into an unspecified output
answer=MBUILD a specific output to somebody else's input that was just received in the CP.

INGEST The taking in of an object by an animal into the inner workings of that animal. INGEST requires an actor, object, and direction.
ex..left margin 16
eat=INGEST a solid into the body through the mouth
breathe=INGEST air into the body through the mouth
shoot up=INGEST a narcotic into the body through a vein by PROPEL narcotic through needle.
smoke=INGEST smoke from a cigarette into the mouth by INGEST air through the cigarette.

All conceptualizations require conceptualizations that are instruments for them. The remainder of the primitive ACTs are primarily the ACTs of instrumental conceptualizatons. They are:

GRASP To grasp an object.

ATTEND To focus a sense organ to an object.

SPEAK To make a noise.

MOVE To move a bodypart.

EXPEL To push something out of the body.

We have listed above only one sense of each verb that we have chosen to describe. The power of the primitive ACTs is that they can point up similarities between words (such as the use of ATRANS for 'give', 'take', and 'buy' while also highlighting their differences. In addition they serve as a vehicle for disambiguating ambiguous words. Thus, 'smoke' as in "the barn is smoking" would not have INGEST in it but would use a state description involving

'smoke' and 'heat'. Similarly, 'take' as in "take an aspirin" would be INGEST rather than ATRANS. A program that makes these disambiguations and assigns the correct conceptual structure to an input sentence was written by Riesbeck (1974). The program uses knowledge about how words fit together in English and some general knowledge about the physical world to make its distinctions. Upon encountering 'take' for example, the program asks if there is an object around which is medication. It also must find out if the sentential object is really an action, as in "take a beating." If it is we have neither ATRANS or INGEST but a causal structure. This series of tests helps to determine the correct conceptual sense for a word.

The primitive ACTs are useful for organizing the inference process. Rather than stating that if you see something, then you know it and if you hear something then you know it and if you read something then you know it and so on, we simply state that whenever an MTRANS exists, a likely inference is that the MTRANSed information is in the mental location LTM (our representation for 'know'). This is a tremendous savings of time and space. For example, any number of forces can cause different kinds of state changes of damage, location and the like. By listing the criteria that must be met for these inferences once under PROPEL, rather than under each individual verb that might call them up, we organize our knoweldge of facts and events in a very useful way.

(Organizing the inferences under the ACTs also serves to define the ACTs themselves. The theoretical decision for what constitutes a primitive ACT is based on whether the proposed new ACT carries with it a set of inferences not already accounted for by an ACT that is already in use. Similarly, a primitive ACT is dropped when we find that its inferences are already handled by another ACT.)

We have recently added a new primitive ACT called PLAN. PLAN is intended to account for an individual's ability to decide on a step by step course of action that leads to a goal. PLAN takes as input a goal and a set of possible plans that are known to be ways to achieve that goal. PLAN decides among them and produces as output a sequential chain of conceptualizations that the actor intends to act out in order to achieve the desired goal. PLAN takes MBUILD as its instrument where the items MBUILDed are the particular decisions that make up the selection of a given plan.

The general form of PLAN is with an object that takes as input a goal and the plans associated with that goal and produces as output the chain of acts that the actor intends to perform to achieve the goal.

MBUILD is now redefined as having an object that takes multiple inputs dealing with the current facts and the logical consequences and beliefs associated with those facts and produces a new thought which may or may not be an intended future action.

```
          ACT1,ACT2,ACT3...
actor <=> PLAN <-
          GOAL
           ACTS and PLANBOXES

            CONCEPTUALIZATION
actor <=> MBUILD <-
            FACT
            FACT
            BELIEF
            RESULTS OF FACTS
```

The difference in these two ACTs with respect to their realizations in ENGLISH is as follows:

John decided to go to the Bahamas.
MBUILD new thought.

John developed a plan to go to the Bahamas.
PLAN unknown set of ACTs.
GOAL that is input to ACT
PLAN is known

John knows how to get the money.
PLAN with known goal.

John figured out who the murderer is.
MBUILD new thought

John did that to get Mary to like him.
PLAN

John intended to hurt Mary.
PLAN (ambiguous if no ACT was done.)

John intends to hurt Mary.
Ambiguous (not known if he has figured out how)

The differences with respect to inference are:

MBUILD We can infer that the facts were there to make the decision and question the applicability of those facts.

PLAN We can infer that the steps to a goal were thought about before an ACT was taken.

We MBUILD goals. We PLAN ways to get those goals to become reality.

The idea of primitive ACTs has met with a general agreement that such a thing is in principle reasonable on the one hand, and a general uneasiness with the set that Conceptual Dependency proposes on the other. Do we really need such a small set? Must we break down all words into these primitives each and every time? How do you arrive at the correct set? What about words like 'drive' or 'dance' or 'hunt'?

The answer to these questions is not simple. We never sought to create an extremely small set. We just happended on that set in the course of attempting to find out what the entities such as 'understand3' and 'believe1' that we were creating were. The set that we came up with has so far been adequate to handle a myriad of domains that we have tried. We expect that over time a different set than we are currently using will emerge. However, our expectation is that the order of magnitude is correct. It should be possible to handle most worlds with a very small set.

The question of whether we must break down everything we see in terms of primitive ACTs seems to be the one that excites the most controversy. Wouldn't it be simpler to use English words where we can and break down into primitives only when we have to? Yes, it would be simpler to do that if we had a data base that had in it "John drives a red car" and we were asked "Who drives a red car?" Researchers who work with highly restricted data bases probably have no need for primitive ACTs. But real understanding systems must be capable of handling qustions such as "Who was that I saw moving down the street on those fancy wheels??" or even just "Who parked that car over there?" Life is rarely as simple as naive question answering systems would like it to be. Real understanding implies knowing what the meaning of everything you hear is, before you attempt to process it in memory. The truth is that it just isn't any harder to map input sentences into conceptualizations involving the primitive ACTs than to do a syntactic parse. (See the paper by Riesbeck in this volume.)

But what about words such as 'drive' and 'dance'? The answer is that these words (and many others) have no simple meaning in Conceptual Dependency terms. They are what we used to call SDO, meaning they were names of prescribed sequences of conceptualizations, any subset of which might not be meant when the word was being used. These action sequences can be handled using the primitive ACTs. The problem is in describing just what 'drive' does mean. The answer is that it has a basic primitive associated with it (PTRANS by PROPEL here) but it also has what we now (see Schank and Abelson [1975]) call a script. This script carries along with it all the mundane information that is known about 'drive'.

In creating the primitive ACTs we made no claim that these were the last primitive entities that we would be creating. We are now working on a set of primitives plans and planboxes (described by me in another paper in this volume). These higher level abstract entities serve to organize information about people's goals and plans. They work on top of the conceptualizations that are received as input. Knowledge organizing primitives are every bit as important as event organizing primitives.

It is important to realize why we are creating these abstract entities in memory. Basically, the reason for creating them is organization. A primitive ACT in memory is useful for organizing inferences around. Whenever that ACT appears in a representation, we know that a certain body of knowledge applicable to that ACT is going to be relevant. Primitive or abstract constructionns in a memory are there precisely because a given body of knowledge is frequently accessed as a unit whenever that abstract entity is accessed.

Because we do not remember the original inputs, the abstract entities that we create to represent them must serve a dual role. They must both accurately represent the meaning of the input as well as call up relevant bodies of knowledge that will be applicable as data upon which to base inferences. This dual responsibility of the conceptual representation means that the abstract elements within that representation must be selected with respect to how they affect the other elements of the representation as well as whether they serve to organize information in an efficient manner. Since the primitive ACTs must be more carefully chosen, there are serious questions as to what can constitute an acceptable and workable set.

These considerations do not hold in the situation where abstract entities are created and the input information is retained. Thus, in situations where conceptual representations are to be evaluated for goals that are implicitly expressed within them, we are simply evaluating the input conceptualizations. They will not be replaced by whatever new constructions are invented. Thus, the selection of these higher level abstractions depends mainly on whether we can find a use for them.

So, we are saying that there are many possible kinds of primitives. Primitive actions underlie the words of a language. Operating on top of those primitives, and therfore affected by them, are a set of primitive knowledge organization pieces. Getting at these is what understanding understanding is all about.