What's Necessary to Hide?: Modeling Action Verbs

James F. Allen
Computer Science Department
University of Rochester
Rochester, NY 14627

Abstract

This paper considers what types of knowledge one must possess in order to reason about actions. Rather than concentrating on how actions are performed, as is done in the problem-solving literature, it examines the set of conditions under which an action can be said to have occurred. In other words, if one is told that action A occurred, what can be inferred about the state of the world? In particular, if the representation can define such conditions, it must have good models of time, belief, and intention. This paper discusses these issues and suggests a formalism in which general actions and events can be defined. Throughout, the action of hiding a book from someone is used as a motivating example.

1. Introduction

This paper suggests a formulation of events and actions that seems powerful enough to define a wide range of event and action verbs in English. This problem is interesting for two reasons. The first is that such a model is necessary to express the meaning of many sentences. The second is to analyze the language production and comprehension processes themselves as purposeful action. This was suggested some time ago by Bruce [1975] and Schmidt [1975]. Detailed proposals have been implemented recently for some aspects of language production [Cohen, 1978] and comprehension [Allen, 1979]. As interest in these methods grows (e.g., see [Grosz, 1979; Brachman, 1979]), the inadequacy of existing action models becomes increasingly obvious.

The formalism for actions used in most natural language understanding systems is based on case grammar. Each action is represented by a set of assertions about the semantic roles the noun phrases play with respect to the verb. Such a formalism is a start, but does not explain how to represent what an action actually signifies. If one is told that a certain action occurred, what does one know about how the world changed (or didn't change!). This paper attempts to answer this question by outlining a temporal logic in which the occurrence of actions can be tied to descriptions of the world over time.

One possibility for such a mechanism is found in the work on problem-solving systems (e.g. [Fikes and Nilsson, 1971; Sacerdoti, 1975]), which suggests one common formulation of action. An action is a function from one world state to a succeeding world state and is described by a set of prerequisites and effects, or by decomposition into more primitive actions. While this model is extremely useful for modeling physical actions by a single actor, it does not cover a large class of actions describable in English. For instance, many actions seemingly describe non-activity (e.g. standing still), or acting in some non-specified manner to preserve a state (e.g. preventing your television set from being stolen). Furthermore, many action descriptions appear to be a composition of simpler actions that are simultaneously executed. For instance,

"Walking to the store while juggling three balls"

seems to be composed of the actions of

"walking to the store

and

"juggling three balls."

It is not clear how such an action could be defined from the two simpler actions if we view actions as functions from one state to another.

The approach suggested here models events simply as partial descriptions of the world over some time interval. Actions are then defined as a subclass of events that involve agents. Thus, it is simple to combine two actions into a new action. The new description simply consists of the two simpler descriptions holding over the same interval.

The notions of prerequisite, result, and methods of performing actions will not arise in this study. While they are important for reasoning about how to attain goals, they don't play an explicit role in defining when an action can be said to have occurred. To make this point clear, consider the simple action of turning on a light.

There are few physical activities that are a necessary part of performing this action. Depending on the context, vastly different patterns of behavior can be classified as the same action. For example, turning on a light usually involves flipping a light switch, but in some circumstances it may involve tightening the light bulb (in the basement), or hitting the wall (in an old house). Although we have knowledge about how the action can be performed, this does *not* define what the action is. The key defining characteristic of turning on the light seems to be that the agent is performing some activity which will cause the light, which is off when the action starts, to become on when the action ends. The importance of this observation is that we could recognize an observed pattern of activity as "turning on the light" even if we had never seen or thought about that pattern previously.

The model described here is in many ways similar to that of Jackendoff [1976]. He provides a classification of event verbs that includes verbs of change (GO verbs) and verbs that assert a state remaining constant over an interval of time (STAY verbs), and defines a representation of action verbs of both types by introducing the notion of agentive causality and permission. However, Jackendoff does not consider in detail how specific actions might be precisely defined with respect to a world model.

The next two sections of this paper will introduce the temporal logic and then define the framework for defining events and actions. To be as precise as possible, I have remained within the notation of the first order predicate calculus. Once the various concepts are precisely defined, the next necessary step in this work is to define a computationally feasible representation and inference process. Some of this work has already been done. For example, a computational model of the temporal logic can be found in Allen [1981]. Other areas are currently under investigation.

The final section demonstrates the generality of the approach by analyzing the action of hiding a book from someone. In this study, various other important conceptual entities such as belief, intention, and causality are briefly discussed. Finally, a definition of what it means to hide something is presented using these tools.

2. A Temporal Logic

Before we can characterize events and actions, we need to specify a temporal logic. The logic described here is based on temporal intervals. Events that appear to refer to a point in time (i.e., finishing a race) are considered to be implicitly referring to another event's beginning or ending. Thus the only time points we will see will be the endpoints of intervals.

The logic is a typed first order predicate calculus, in which the terms fall into the following three broad categories:

- terms of type *TIME-INTERVAL* denoting time intervals;
- terms of type *PROPERTY*, denoting descriptions that can hold or not hold during a particular time; and
- terms corresponding to objects in the domain.

There are a small number of predicates. One of the most important is *HOLDS*, which asserts that a property holds (i.e., is true) during a time interval. Thus

HOLDS(p,t)

is true only if property *p* holds during t. As a subsequent axiom will state, this is intended to mean that *p* holds at every subinterval of t as well.

There is no need to investigate the behavior of *HOLDS* fully here, but in Allen [forthcoming] various functional forms are defined that can be used within the scope of a *HOLDS* predicate that correspond to logical connectives and quantifiers outside the scope of the *HOLDS* predicate.

There is a basic set of mutually exclusive relations that can hold between temporal intervals. Each of these is represented by a predicate in the logic. The most important are:

DURING(t1,t2)--time interval t1 is fully contained within t2, although they may coincide on their endpoints.

BEFORE(t1,t2)--time interval t1 is before interval t2, and they do not overlap in any way;

OVERLAP(t1,t2)--interval t1 starts before t2, and they overlap;

MEETS(t1,t2)--interval t1 is before interval t2, but there is no interval between them, i.e., t1 ends where t2 starts.

Given these predicates, there is a set of axioms defining their interrelations. For example, there are axioms dealing with the transitivity of the temporal relationships. Also, there is the axiom mentioned previously when the *HOLDS* predicate was introduced: namely

(A.1) *HOLDS(p,t) & DURING(t1,t) => HOLDS(p,t1)*

This gives us enough tools to define the notion of action in the next section.

3. Events and Actions

In order to define the role that events and actions play in the logic, the logical form of sentences asserting that an event has occurred must be discussed. Once events have been defined, actions will be defined in terms of them. One suggestion for the logical form is to define for each class of events a property such that the property *HOLDS* only if the event occurred. This can be discarded immediately as axiom (A.1) is inappropriate for events. If an event occurred over some time interval T, it does not mean that the event also occurred over all subintervals of T. So we introduce a new type of object in the logic, namely events, and a new predicate *OCCUR*. By representing events as objects in the logic, we have avoided the difficulties described in Davidson [1967].

Simply giving the logical form of an event is only a small part of the analysis. We must also define for each event the set of conditions that constitute its occurrence. As mentioned in the introduction, there seems to be no restriction on what kind of conditions can be used to define an event except that they must partially describe the world over some time interval.

For example, the event "the ball moving from x to y" could be modeled by a predicate *MOVE* with four arguments: the object, the source, the goal location, and the move event itself. Thus,

MOVE(Ball, x, y, m)

asserts that m is an event consisting of the ball moving from x to y. We assert that this event occurred over time t by adding the assertion

OCCUR(m,t).

With these details out of the way, we can now define necessary and sufficient conditions for the event's occurrence. For this simple class of move events, we need an axiom such as:

(forall object,source,goal,t,e)
  MOVE(object,source,goal,e) & OCCUR(e,t)
    <=> (exists t1,t2)
      OVERLAPS(t1,t) & OVERLAPS(t,t2) &
      BEFORE(t1,t2) &
      HOLDS(at(object,source),t1) &
      HOLDS(at(object,goal),t2)

A simple class of events consists of those that occur only if some property remains constant over a particular interval (cf. Jackendoff's STAY verbs). For example, we may assert in English

"The ball was in the room during T."

"The ball remained in the room during T."

While these appear to be logically equivalent, they may have very different consequences in a conversation. This formalism supports this difference. The former sentence asserts a proposition, and hence is of the form

HOLDS(in(Ball,Room),T)

while the latter sentence describes an event, and hence is of the form

REMAIN-IN(Ball,Room,e) & OCCURS(e,T).

We may capture the logical equivalence of the two with the axiom:

(forall b,r,e,t)
REMAIN-IN(b,r,e) & OCCUR(e,t)
<=> HOLDS(in(b,r),t).

The problem remains as to how the differences between these logically equivalent formulas arise in context. One possible difference is that the second may lead the reader to believe that it easily might not have been the case.

Actions are events that involve an agent in one of two ways. The agent may cause the event or may allow the event (cf. [Jackendoff, 1976]). Corresponding to these two types of agency, there are two predicates, *ACAUSE* and *ALLOW*, that take an agent, an event, and an action as arguments. Thus the assertion corresponding to

"John moved B from S to G"

is

MOVE(B,G,S,e1) & ACAUSE(John,e1,a1) & OCCUR(a1,t)

The axiomatization for *ACAUSE* and *ALLOW* is tricky, but Jackendoff provides a reasonable starting set. In this paper, I shall only consider agency by causation further. The most important axiom about causality is

(A.2) *(forall a,e,act,t)*
ACAUSE(a,e,act) & OCCUR(act,t)
=> OCCUR(e,t)

For our purposes, one of the most important facts about the *ACAUSE* relation is that it suggests the possibility of intentionality on the part of the agent. This will be discussed in the next section.

Note that in this formalism composition of events and actions is trivial. For example, we can define an action composition function *together* which produces an action or event that consists of two actions or events occuring simultaneously as follows:

(A.3) *(forall a,b,t)*
OCCURS(together(a,b),t) <=>
OCCURS(a,t) & OCCURS(b,t)

4. What's Necessary to Hide?

The remainder of this paper applies the above formalism to the analysis of the action of hiding a book from someone. Along the way, we shall need to introduce some new representational tools for the notions of belief, intention, and causality.

The definition of hiding a book should be independent of any method by which the action was performed, for, depending on the context, the actor could hide a book in many different ways. For instance, the actor could

- put the book behind a desk,
- stand between the book and the other agent while they are in the same room, or
- call a friend Y and get her or him to do one of the above.

Furthermore, the actor might hide the book by simply not doing something s/he intended to do. For example, assume Sam is planning to go to lunch with Carole after picking Carole up at Carole's office. If, on the way out of Sam's office, Sam decides not to take his coat because he doesn't want Carole to see it, then Sam has hidden the coat from Carole. Of course, it is crucial here that Sam believed that he normally would have taken the coat. Sam couldn't have hidden his coat by forgetting to bring it.

This example brings up a few key points that may not be noticed from the first three examples. First, Sam must have intended to hide the coat. Without this intention (i.e., in the forgetting case), no such action occurs. Second, Sam must have believed that it was likely that Carole would see the coat in the future course of events. Finally, Sam must have acted in such a way that he then believed that Carole would not see the coat in the future course of events. Of course, in this case, the action Sam performed was "not bringing the coat," which would normally not be considered an action unless it was intentionally not done.

I claim that these three conditions provide a reasonably accurate definition of what it means to hide something. They certainly cover the four examples presented above. As stated previously, however, the definition is rather unsatisfactory, as many extremely difficult concepts, such as belief and intention, were thrown about casually.

There is much recent work on models of belief (e.g., [Cohen, 1978; Moore, 1979; Perlis, 1981; Haas, 1981]). I have little to add to these efforts, so the reader may assume his or her favorite model. I will assume that belief is a modal operator and is described by a set of axioms along the lines of Hintikka [1962]. The one important thing to notice, though, is that there are two relevant time indices to each belief; namely, the time over which the belief is held, and the time over which the proposition that is believed holds. For example, I might believe today that it rained last weekend. This point will be crucial in modeling the action of hiding. To introduce some notation, let

"A believes (during Tb) that p holds (during Tp)"

be expressed as

HOLDS(believes(A,holds(p,Tp)),Tb).

The notion of intention is much less understood than the notion of belief. However, let us approximate the statement

> "A intends (during Ti) that action a happen (during Ta)"

by

> "A believes (during Ti) that a happen (during Ta)"

and

> "A wants (during Ti) that a happen (during Ta)"

This is obviously not a philosophically adequate definition (e.g., see [Searle, 1980]), but seems sufficient for our present purposes. The notion of wanting indicates that the actor finds the action desirable given the alternatives. This notion appears impossible to axiomatize as wants do not appear to be rational (e.g. Hare [1971]). However, by adding the belief that the action will occur into the notion of intention, we ensure that intentions must be at least as consistent as beliefs.

Actions may be performed intentionally or unintentionally. For example, consider the action of breaking a window. Inferring intentionality from observed action is a crucial ability needed in order to communicate and cooperate with other agents. While it is difficult to express a logical connection between action and intention, one can identify pragmatic or plausible inferences that can be used in a computational model (see [Allen, 1979]).

With these tools, we can attempt a more precise definition of hiding. The time intervals that will be required are:

Th--the time of the hiding event;

Ts--the time that Y is expected to see the book;

Tb1--the time when X believes Y will see the book during Ts, which must be *BEFORE* Th;

Tb3--the time when X believes Y will not see the book during Ts, which must be *BEFORE* or *DURING* Th and *AFTER* Tb1.

We will now define the predicate

HIDE(agent,observer,object,act)

which asserts that act is an action of hiding. Since it describes an action, we have the simple axiom capturing agency:

(forall agent,observer,object,act
HIDE(agent,observer,object,act)
=> (Exists e ACAUSE(agent,e,act)))

Let us also introduce an event predicate

SEE(agent,object,e)

which asserts that *e* is an event consisting of *agent* seeing the *object*.

Now we can define HIDE as follows:

(forall ag,obs,o,a,Th,
HIDE(ag,obs,o,a) & OCCUR(a,Th)
=> (Exists Ts,Tb1,Tb3,e)
1) HOLDS(intends(ag,occur(a,Th)),Th)
2) HOLDS(believes(ag,occur(e,Ts)),Tb1)
3) HOLDS(believes(ag,¬occur(e,Ts)),Tb3)
where
4) SEE(obs,o,e)

and the intervals Th, Ts, Tb1, Tb3 are related as discussed above. Condition (4) defines e as a seeing event, and might also need to be within ag's beliefs.

This definition is lacking part of our analysis; namely that there is no mention that the agent's beliefs changed because of something s/he did. We can assert that the agent believes (between Tb1 and Tb3) he or she will do an action (between Tb1 and Th) as follows:

(exists a1,e1,Tb2
5) ACAUSE(ag,e1,a1)
6) HOLDS(believes(ag,OCCUR(a1,Ta1)),Tb2)
where *Tb1 < Tb2 < Tb3 and*
Tb1 < Ta1 < Th

But this has not captured the notion that belief (6) caused the change in belief from (2) to (3). Since (6) and (3) are true, asserting a logical implication from (6) to (3) would have no force. It is essential that the belief (6) be a key element in the reasoning that leads to belief (3).

To capture this we must introduce a notion of causality. This notion differs from *ACAUSE* in many ways (e.g. see [Taylor, 1966]), but for us the major difference is that, unlike *ACAUSE*, it suggests no relation to intentionality. While *ACAUSE* relates an agent to an event, *CAUSE* relates events to events. The events in question here would be coming to the belief (6), which *CAUSES* coming to the belief (3).

One can see that much of what it means to hide is captured by the above. In particular, the following can be extracted directly from the definition:

- if you hide something, you intended to hide it, and thus can be held responsible for the action's consequences;
- one cannot hide something if it were not possible that it could be seen, or if it were certain that it would be seen anyway;
- one cannot hide something simply by changing one's mind about whether it will be seen.

In addition, there are many other possibilities related to the temporal order of events. For instance, you can't hide something by performing an action after the hiding is supposed to be done.

Conclusion

I have introduced a representation for events and actions that is based on an interval-based temporal logic. This model is sufficiently powerful to describe events and actions that involve change, as well as those that involve maintaining a state. In addition, the model readily allows the composition and modification of events and actions.

In order to demonstrate the power of the model, the action of hiding was examined in detail. This forced the introduction of the notions of belief, intention, and causality. While this paper does not suggest any breakthroughs in representing these three concepts, it does suggest how they should interact with the notions of time, event, and action.

At present, this action model is being extended so that reasoning about performing actions can be modeled. This work is along the lines described in [Goldman, 1970].

Acknowledgements

The author wishes to thank Jerry Feldman, Alan Frisch, Margery Lucas, and Dan Russell for many enlightening comments on previous versions of this paper.

This research was supported in part by the National Science Foundation under Grant No. IST-80-12418, and in part by the Office of Naval Research under Grant No. N00014-80-C-0197.

References

Allen, J.F., "A General View of Action and Time," TR, Dept. Computer Science, U. Rochester, forthcoming.

Allen, J.F., "A Plan-Based Approach to Speech Act Recognition," Ph.D. thesis, Dept. Computer Science, U. Toronto, 1979.

Allen, J.F., "Maintaining Knowledge about Temporal Intervals," TR86, Dept. Computer Science, U. Rochester, January 1981.

Brachman, R.J., "Taxonomy, Descriptions, and Individuals in Natural Language Understanding," in *Proc.*, 17th Annual Meeting of the Assoc'n. for Computational Linguistics, 33-37, UCSD, La Jolla, CA, August 1979.

Bruce, B., "Belief Systems and Language Understanding," Report 2973, Bolt, Beranek & Newman, Inc., 1975.

Cohen, P.R., "On Knowing What to Say: Planning Speech Acts," TR 118, Dept. Computer Science, U. Toronto, 1978.

Davidson, D., "The Logical Form of Action Sentences," in N. Rescher (Ed). *The Logic of Decision and Action.* Pittsburgh, PA: U. Pittsburgh Press, 1967.

Fikes, R.E. and N.J. Nilsson, "STRIPS: A New Approach to the Application of Theorem Proving to Problem Solving," *Artificial Intelligence 2*, 189-205, 1971.

Goldman, A. *A Theory of Human Action.* New Jersey: Princeton U. Press, 1970.

Grosz, B.J., "Utterance and Objective: Issues in Natural Language Communication," in *Proc.*, 6th IJCAI, 1067-1076, Tokyo, August 1979.

Haas, A., "Sententialism and the Logic of Belief and Action," Ph.D. thesis, Dept. Computer Science, U. Rochester, expected 1981.

Hare, R.M. "Wanting: Some Pitfalls," in Binkley, Bronaugh, and Morras (Eds). *Agent, Action, and Reason.* Toronto: U. Toronto Press, 1971.

Hintikka, J. *Knowledge and Belief.* Ithaca, NY: Cornell U. Press, 1962.

Jackendoff, R., "Toward an Explanatory Semantic Representation," *Linguistic Inquiry 7*, 1, 89-150, Winter 1976.

Moore, R.C., "Reasoning about Knowledge and Action," Ph.D. thesis, MIT, February 1979.

Perlis D., "Language, Computation, and Reality," Ph.D. thesis, Dept. Computer Science, U. Rochester, 1981.

Sacerdoti, E.D. *A Structure for Plans and Behavior.* New York: Elsevier North-Holland, Inc., 1977.

Schank, R. and R. Abelson. *Scripts, Plans, Goals, and Understanding.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1977.

Schmidt, C.F., "Understanding Human Action," in *Proc.*, Theoretical Issues in Natural Language Processing, Cambridge, MA, 1975.

Searle, J.R., "The Intentionality of Intention and Action," *Cognitive Science 4*, 1, 1980.

Taylor, R. *Action and Purpose.* New Jersey: Prentice Hall, 1966.

Wilensky, R., "Understanding Goal-Based Stories," Ph.D. thesis, Yale U., 1978.