Reasoning about Hand-Eye Coordination

Ernest Davis*
Courant Institute
New York, New York

Presented at the IJCAI-89 Workshop on Knowledge, Perception, and Planning

July 11, 1989

Abstract

This paper outlines a formal theory capable of supporting inferences about plans involving continuous coordination between sensors and effectors, such as "Follow the sun until reaching a river; follow the river until reaching a lake." In particular, it presents a framework in which such actions can be represented and their physical effects, physical preconditions and knowledge preconditions can be defined.

Robotic actions in a world that is not wholly known and controlled almost always involve the use of sensory information as a guide. (There are some remarkable exceptions: for example, [Peshkin and Sanderson, 87] shows how, in specialized environments, a robot can place an object in a desired configuration despite being ignorant of its original position.) If actions must be quickly responsive to external states whose exact time or circumstance cannot be predicted in advance, then perceptions must be integrated with effector actions in a close feedback loop. *Hand-eye coordination* is this quasi-continuous integration of perception with action. Many robotic systems have been built with some degree of hand-eye coordination *Monitoring* a sensor during action, which is one form of hand-eye coordination, has been used as a planning primitive in [Firby, 89], and has been considered as a primitive of motion planning algorithms [Lumelsky,87] [Cox, 88].

In order for an AI system controlling a robot with suitable sensors and effectors to apply hand-eye coordination to a broad range of problems, the system must have a theory for reasoning about plans involving hand-eye coordination, and for determining their effects and feasibility. Such reasoning requires an understanding of the physical properties of the effectors and sensors and of the relations between perception, knowledge, and action. In this paper, we study how existing theories of these domains can be applied to reasoning about hand-eye coordination. (I understand that a similar analysis is carried out in [Sandewall, 89], but as of the date of writing, I have not seen this paper.) In particular, we draw implicitly on the theories of knowledge, action, and planning by Moore [1980] and Morgenstern [1987]; on the analysis of continuous action in [McDermott, 82], [Davis, 84a], and [Davis,84b]; and on the theory of perception and knowledge in [Davis, 88]. The representations used below are mostly drawn from these previous paper. We trust that they are clear enough for the discursive purposes of this informal paper, and that the reader can fill in the technical details for himself [Davis, in prep.].

Of course, we do not expect to design a formal theory that can generate or explain the precise details of a complex perceptual feedback system, such as Donner's juggling system [Reference?].

*This research was supported by NSF grant #IRI-8801529.

Even the human designers of such systems often cannot justify the details of their design on the basis of general principles; a large measure of empirical fine-tuning is almost always involved. The most that we can ask from our theory is that it justify some basic commonsensical inferences; in particular, that it discriminate between plans that are obvious workable, and plans that are obviously ridiculous.

1 Examples

We will illustrate the kind of reasoning we wish to model, using variants of a single example:

> A traveler in the desert wishes to go to Lake Arthur. He knows that, a mile or two to his east, is a rapidly-flowing river that flows into the lake. The sun has just risen. He plans to walk toward the sun until reaching the river, then follow the river downstream to the lake.

Our theory should support the conclusion that this plan will work. By contrast, the theory should support the conclusion that none of the following plans are feasible:

1. The plan, "Go up into the air until you are high enough to see the lake. Go on a straight line to the lake."
2. The plan "Follow the great circle connecting you to the lake until reaching the lake."
3. It is an overcast night. The traveler constructs the plan, "Travel east to the river, then follow the river downstream."
   a. The traveler is disoriented at the start.
   b. The traveler knows that he is currently facing east, but he has no way to maintain a constant direction over a long trip.

The first plan fails because it is physically impossible. The remaining plans fail because they violate a *knowledge precondition*; they each require information that will not be available to the traveler. Our task in this paper is to provide a system for representing plans such as these, and criteria for evaluating their feasibility.

2 Monitored Actions

The major problem in developing a theory of plans such as those above is to find a representation and feasibility criteria for events or actions (we will use the terms interchangeably) such as "Go towards the sun until reaching the river," and "Follow the river downstream until reaching the lake." Sensory information enters into these actions in two ways: (1) in determining the end of the action (being at the river; being at the lake); (2) in guiding the course of the action during execution (moving toward the sun; following the river.) We will begin with (1), which is the simpler issue.

Since we are dealing with continuous events, we will assume that the time line has the structure of the real numbers. Following Shoham [1988], we define an event type E to be *liquid* if it is arbitrarily divisible across subintervals and combinable across adjoining intervals; that is, if E occurs during an interval I, it occurs during any subinterval of I; if it occurs in adjoining intervals I and J, then it occurs in join(I, J).

$$\text{liquid}(E) \Leftrightarrow$$
$$\forall_I (\forall_{S \in I} \exists_{I1}\, S \in I1 \wedge \text{occurs}(I1, E)) \Rightarrow \text{occurs}(I, E)$$

For example, "Go toward the sun," or "Follow the river" are liquid event types.

Given any liquid action type E and temporal state Q, we define the composite action "monitor(Q, E)" as the performance of E until Q become true.

$$\text{occurs}([SB, SE], \text{monitor}(Q, E)) \Leftrightarrow$$
$$[\ \text{occurs}([SB, SE], E) \wedge [\forall_S\, SB \leq S < SE \Rightarrow \neg\text{true\_in}(S, Q)] \wedge$$
$$[\forall_{S2 > SE} \exists_{S1}\, SE \leq S1 < S2 \wedge \text{true\_in}(S1, Q)]]$$

We can thus represent the action "Go toward the sun until reaching the river" in the form,

$$\text{monitor}(\text{distance}(\text{me},\text{river}) < 5 \cdot \text{foot}, \text{go\_towards}(\text{sun}))$$

Note that the definition above assumes that perception is continuous, so that the action will terminate as soon as the termination condition becomes true. In the case of an analog perceptor, such as thermometer or a voltage meter, this is an accurate model; in the case of a discrete perceptor, it is an approximation of a rapid iteration of discrete perceptions.

The action "monitor(Q, E)" is physically feasible in situation S if E is feasible in S and remains feasible as long as Q is false. Thus, the action "Go towards the sun until reaching the river," is feasible in S, if executing the action "Go towards the sun" will bring the robot to the river before it brings him to any other obstacle.

$$\text{true\_in}(S1, \text{phys\_poss}(\text{monitor}(Q, E))) \Leftrightarrow$$
$$[\text{true\_in}(S1, \text{phys\_poss}(E)) \wedge$$
$$\forall_{S2}\ \text{occurs}([S1, S2], E) \Rightarrow [\text{phys\_poss}(S2, E) \vee [\ \exists_{S3}\, S1 \leq S3 \leq S2 \wedge \text{true\_in}(S3, Q)]]]$$

The knowledge conditions for the action "monitor(Q, E)" are satisfied for a robot A if A knows that throughout its execution, (i) the knowledge preconditions for E will be satisfied; and (ii) A will know whether Q is true.

$$\text{true\_in}(S1, \text{kp\_satisfied}(A, \text{monitor}(Q, E))) \Leftrightarrow$$
$$\text{true\_in}(S1, \text{know}(A, \forall_{S2}\ \text{occur}([S1, S2], \text{monitor}(Q, E)) \Rightarrow$$
$$\forall_{S \in [S1, S2]}\ \text{true\_in}(S, \text{know\_whether}(A, \text{true\_in}(S, Q))) \wedge$$
$$\text{true\_in}(S, \text{kp\_satisfied}(A, E))))$$

3 Guided Actions

Our next problem is to represent and characterized actions guided by perceptions, such as, "Go toward the sun", or "Follow the river." We will discuss two types of approaches to this problem. The first representation is specific to each different type of task, but applies to all robots carrying out that task. The second representation must be tailored to each robot, but can be used for a wide variety of tasks to be performed by that robot.

In the first representation, we use a different primitive action function for each different type of action that can be carried out using perceptual guidance. For instance, the primitive "go_towards(O)" represents the action of going towards object O. Its occurrence is defined by the rule that the direction of motion is always toward the object O. (Or rather, in the case of "going towards the sun," the direction of motion is the horizontal projection of the direction toward the sun.)

$$\text{occurs}(I,\text{go\_towards}(O)) \Leftrightarrow \forall_{S \in I}\ \text{true\_in}(S,\text{parallel}(\text{velocity}(\text{me}), \text{place}(O) - \text{place}(\text{me})))$$

The action is physically feasible if, at every moment, there is a physically feasible motion whose velocity satisfies the above constraint "parallel(velocity(me), place(O) − place(me))".

The action "go_towards(O)" is epistemically feasible if the angle between the orientation of the robot and the direction from the robot to O is known.

$$\text{true\_in}(S,\text{kp\_satisfied}(A,\text{go\_towards}(O))) \Leftrightarrow$$
$$\text{true\_in}(S,\text{know\_val}(A,\text{angle}(\text{place}(O) - \text{place}(A), \text{orientation}(A))))$$

Similarly, we may define the function "follow(P, D, B)", meaning the action of following path P at distance $\leq D$, keeping P on side B (right or left). A geometrical definition of this motion, analogous to the definition of "go_towards" but much more complicated, can be given. The follow action is physically feasible if there is a feasible motion satisfying the definition. It is epistemically feasible if there is a continuous function f(T) from the time line to the path such that (i) f(T) is always within D of the robot; and (ii) at time T, the robot knows that f(T), described relative to his own reference frame, is on the path.

The main disadvantage of this approach is that it gives no means for reasoning about actions that are not specifically in the library: either minor variations such as "Follow the path at distance exactly D," or entirely new actions such as "Stay between object O and fixed point P." Moreover, knowledge preconditions must be defined *ad hoc* for each action type; there is no general theory of knowledge preconditions.

The second approach treats every robot separately, but gives a uniform treatment of actions. The actions of a given robot are characterized in terms of a fixed set of parameters that the robot can control directly. For instance, we might characterize a mobile robot with inertial guidance as directly controlling its two-dimensional velocity. We might characterize a mobile robot without inertial guidance as directly controlling its rotational velocity and its linear speed. This characterization need not be too literal; for instance, it would be reasonable to characterize a robot arm as directly controlling the position and orientation of a hand, rather than as directly controlling the angle of the joints involved. However, if there is a significant different between what a robot tries to do and the results of its attempts, then the characterization should deal with the attempts and characterize the results in terms of a partial constraint relating it to the attempt. In the case of the robot discussed in example 3.b, which tends to lose its sense of direction, we would characterize its behavior in terms of its attempted speed and angular velocity, and then give a rule stating that the real speed and angular velocity lies within some range of the attempt.

An action is then a binding of the controllable parameters to the values of some fluents. The action occurs if the parameters take on the value of the fluents at each instant. The action is epistemically feasible if, at every instant, the values of all the fluents are known to the robot.

For instance, if the robot has a directly controllable velocity $\vec{v}$, then we could represent the action "go_toward(O)" in terms of the constraint $\vec{v}(t)=\alpha(t) \cdot$ direction(place(O) − place(me)), where $\alpha(t)$ is a positive scalar function. The knowledge precondition of this action is that the robot should know the value of "direction(place(O) − place(me))". This characterization of knowledge preconditions directly follows the rule in [Moore,80] and [Morgenstern, 87] that a primitive action description is epistemically feasible just if the value of its arguments is known.

One limitation of this approach is that different robots must be characterized in different terms; for example, the representation of "following a path" is entirely different from one robot to the next. Another limitation is that the continuous characterization of control used here may not be the most natural in all situations. Consider, for example, a robot tracking an object along a one-dimensional track. The robot's objective is always to stay within 2 feet of the object. The object moves at a maximum at 1 foot per second. The robot directly controls its velocity, v. It therefore applies the following control strategy:

At one-second intervals, do:
If the object is more than 1 foot to the right, set $v = 1$.
If the object is more than 1 foot to the left, set $v = -1$.
If the object is within a foot, set $v = 0$.

It is easily seen that this strategy will always keep the robot within two feet of the object. However, it does not conform to our system of description, since the velocity is not set continuously, but only at intervals. If we try to turn it directly into continuous form, replacing, "At one second intervals" with "At every instant", we end up with the differential equation

$$\dot{x} = \begin{cases} 1 & \text{if } x(t) \leq o(t) - 1 \\ -1 & \text{if } x(t) \geq o(t) + 1 \\ 0 & \text{otherwise} \end{cases}$$

which has no solution. We can turn it to a valid differential equation by making $\dot{x}$ a continuous function of x. For instance, we could use the differential equation $\dot{x} = o(t) - x$.

It is not clear whether the continuous or the discrete formulation is better overall. The continuous version is certainly easier to fit into a logical axiomatization; the discrete version may be generally closer to the actual truth. Which will be easier to implement in a reasoning system is an open question.

4 Outline of an Analysis

We can now see that one possible verification of our example plan, "Follow the sun until reaching the river; then follow the river downstream, staying within 10 feet of the river, until reaching the lake," would involve the following steps:

Physical correctness:

- Since the sun is in the east for several hours after dawn, following the sun will result in an eastward motion for several hours.
- Since the only obstacle to the robot is the river, the motion "Go towards the sun" can be continued until the river is reached.
- Since the river is a mile to the east, eastward motion for an hour will bring the robot to the river.

Therefore, the step, "Go towards the sun until reaching the river," will bring the robot to the river. The physical analysis of the second step is similar.

Epistemic adequacy:

- Since the sun can be seen during the time in question, the robot knows the direction from itself to the sun. The action "Go towards the sun," is therefore always epistemically feasible.

- The robot can always see whether or not it is at the river. Therefore, the action "Go towards the sun until reaching the river," is epistemically adequate.
- When the robot is at the river, it will be able to see which way it flows. The robot can therefore decide whether following the river downstream involves following it on the left or on the right.
- As long as the robot stays within 10 feet of the river, it can see enough of the river to be able to follow it.

We will provide a full formal account of this inference in a forthcoming paper [Davis, in prep.]

5 References

[Cox, 88] James Cox, *Online Motion Planning,* NYU Ph.D. Thesis, Comp. Sci. Dept., 1988.

[Davis, 84a] Ernest Davis, "An Ontology of Physical Actions," NYU Comp. Sci. Dept., Technical Report #123, 1984.

[Davis, 84b] Ernest Davis, "A High-Level Real-Time Programming Language," NYU Comp. Sci. Dept., Technical Report #145, 1984.

[Davis, 88] Ernest Davis, "Inferring Ignorance from the Locality of Visual Perception," Proc. AAAI, 1988, pp. 786-790.

[Davis, in prep.] Ernest Davis, "Reasoning about Continuous Perception and Action," in preparation.

[Firby, 89] James Firby, *Adaptive Execution in Complex Dynamic Worlds,* Yale University CS Dept. Research Report #672, 1989.

[Lumelsky, 87] V.J. Lumelsky, "Dynamic path planning for a planar articulated robot arm moving amidst unknown obstacles," *Automatica*, vol. 23, no. 5, 1987. pp. 551-570.

[McDermott, 82] Drew V. McDermott, "A Temporal Logic for Reasoning about Processes and Plans," *Cognitive Science*, 1982.

[Moore, 80] Robert Moore, *Reasoning about Knowledge and Action,* SRI Technical Note #191, 1980.

[Morgenstern, 87] Leora Morgenstern, "Knowledge Preconditions for Actions and Plans," Proc. IJCAI 1987, pp. 867-874.

[Peshkin and Sanderson, 87] M.A. Peshkin and A.C. Sanderson, "Planning Robotic Manipulation Strategies for Sliding Objects," Proc. IEEE Int'l. Conf. on Robotics and Automation, 1987.

[Sandewall, 89] Eric Sandewall, "Filter Preferential Entailment for the Logic of Action in Almost Continuous Worlds," IJCAI-89.

[Shoham, 88] Yoav Shoham, *Reasoning about Change,* MIT Press, 1988.