Mind Is About Conditional Predictions

Rich Sutton

3/21/2000

Simplifying and generalizing, one thing seems clear to me about mental activity---that the purpose of much of it can be considered to be the making of predictions. By this I mean a fairly general notion of prediction, including conditional predictions and predictions of reward. And I mean this in a sufficiently strong and specific sense to make it non-vacuous.

For concreteness, assume the world is a Markov Decision Process (MDP), that is, that we have discrete time and clear actions, sensations, and reward on each time step. Then, obviously, among the interesting predictions to make are those of immediate rewards and state transitions, as in "If I am in this state and do this action, then what will the next state and reward be?" The notion of value function is also a prediction, as in "If I am in this state and follow this policy, what will my cumulative discounted future reward be?" Of course one could make many value-function predictions, one for each of many different policies.

Note that both kinds of prediction mentioned above are conditional, not just on the state, but on action selections. They are *hypothetical* predictions. One is hypothetical in that it is dependent on a single action, and the other is hypothetical in that it is dependent on a whole policy, a whole way of behaving. Action conditional predictions are of course useful for actually selecting actions, as in many reinforcement learning methods in which the action with the highest estimated value is preferentially chosen. More generally, it is commonsensical that much of our knowledge is beliefs about what would happen IF we chose to behave in certain ways. The knowledge about how long it takes to drive to work, for example, is knowledge about the world in interaction with a hypothetical purposive way in which we could behave.

Now for the key step, which is simply to generalize the above two clear kinds of conditional predictions to cover much more of what we normally think of as knowledge. For this we need a new idea, a new way of conditioning predictions that I call *conditioning on outcomes*. Here we wait until one of some clearly designated set of outcomes occurs and ask (or try to predict) something about which one it is. For example, we might try to predict how old we will be when we finish graduate school, or how much we will weigh at the end of the summer, or how long it will take to drive to work, or much you will have learned by the time you reach the end of this article. What will the dice show when they have stopped tumbling? What will the stock price be when I sell it? In all these cases the prediction is about what the state will be when some clearly identified event occurs. It is a little like when you make a bet and establish some clear conditions at which time the bet will be over and it will be clear who has won.

A general conditional prediction, then, is conditional on three things: 1) the state in which it is made, 2) the policy for behaving, and 3) the outcome that triggers the time at which the predicted event is to occur. Of course the policy need only be followed from the time the prediction is made until the outcome triggering event. Actions taken after the trigger are irrelevant. [This notion of conditional prediction has been previously explored as the models of temporally extended actions, also known as "options" ([Sutton, Precup, and Singh, 1999](#); Precup, thesis in preparation).

Let us return now to the claim with which I started, that much if not most mental activity is focused on such conditional predictions, on learning and computing them, on planning and reasoning with them. I would go so far as to propose that much if not most of our knowledge is represented in the form of such predictions, and that they are what philosophers refer to as "concepts". To properly argue these points would of course be a lengthy undertaking. For now let us just cover some high points, starting with some of the obvious advantages of conditional predictions for knowledge representation.

Foremost among these is just that predictions are grounded in the sense of having a clear, mechanically determinable meaning. The accuracy of any prediction can be determined just by running its policy from its state until an outcome occurs, then checking the prediction against the outcome. No human intervention is required to interpret the representation and establish the truth or falsness of any statement. The ability to compare predictions to actual events also make them suitable for beling learned automatically. The semantics of predictions also make it clear how they are to be used in automatic planning methods such as are commonly used with MDPs and SMDPs. In fact, the conditional predictions we have discussed here are of exactly the form needed for use in the Bellman equations at the heart of these methods.

A less obvious but just as important advantage of outcome-conditional predictions is that they can compactly express much that would otherwise be difficult and expensize to represent. This happens very often in commonsense knowledge; here we give a simple example. The knowledge we want to represent is that you can go to the street corner and a bus will come to take you home within an hour. What this means of course is that if it is now 12:00 then the bus might come at 12:10 and it might come at 12:20, etc., but it will definitely come by 1:00. Using outcome conditioning, the idea is easy to express: we either make the outcome reaching 1:00 and predict that the bus will have come by then, or we make the outcome the arrival of the bus and predict that at that time it will be 1:00 or earlier.

A natural but naive alternative way to try to represent this knowledge would be as a probability of the bus arriving in each time slot. Perhaps it has one-sixth chance of arriving in each 10-minute interval. This approach is unsatisfactory not just because it forces us to say more than we may know, but because it does not capture the important fact that the bus will come eventually. Formally, the problem here

is that the events of the bus coming at different times are not independent. If may have only a one-sixth chance of coming exactly at 1:00, but if it is already 12:55 then it is in fact certain to come at 1:00. The naive representation does not capture this fact that is actually absolutely important to using this knowledge. A more complicated representation could capture all these dependencies but would be just that -- more complicated. The outcome-conditional form represents the fact simply and represents just what is needed to reason with the knowledge this way. Of course, other circumstances may require the more detailed knowledge, and this is not precluded by the outcome-conditional form. This form just permits greater flexibility, in particular, the ability to omit these details while still being of an appropriate form for planning and learning.