Matt Forbes
February 2011
CS 400 - AI Indep. Study
Chapter 13 Discussion

Uncertainty

In real-world systems, decisions can't be made so simply as those in closed, hypothetical environments. As soon as an agent ventures out of the safety of discreteness and full observability, things get a bit more complex. In this new scenario, the results of its actions can't be determined until they have actually been executed, introducing risk. Even actions as basic as movement can have unforeseen consequences in an uncertain environment. Plans of action can not be generated by walking through a search tree, because the results of one action have probabilistic repercussions..

Presented in this chapter, are two modes of probability representation: prior probability and conditional probability. Definitions of such are directly related to their names; prior probability denotes the agent's belief that some event is true with absolutely no other knowledge of the environment. On the other hand, conditional probability is the belief that an event is true given some set of previous knowledge.

An atomic event is the description of the world accounting for all propositions/conditions that can be reasoned about. A comprehensive table listing all permutations of atomic events is called the full join probability distribution. As the number of variables in the world grows, this table becomes infeasible to represent literally. This is okay, because using techniques described in this chapter, we can break down probabilities to much simpler expressions using the notion of independence.

Two variables are independent if their values do not influence each other. For example, the state of the weather is independent of my preference of ice cream flavors. Existence of absolute independence between variables is rare, but can reduce the size of the full join probability distribution. More interesting, is the idea of conditional independence. When two variables are independent when the value of a third is known, we can treat them as independent. An example of this might be in the following situation of driving a car (and crashing):

> The variables have(whiplash) and have(trafficTicket) are not necessarily independent, because if I have a traffic ticket, I could have been in an accident where I received whip lash. If I *know* I was in an accident, then these two variables are now independent. Whether or not it was my fault (and receieved a ticket) does not influence whether or not my neck was injured (whiplash).

Bayes' Rule

Bayes' rule is a pretty simple algebraic manipulation of the probability product rule. This is the rule:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

By itself, Bayes' rule is useful when we have three of the four probabilities but want the fourth. We can simplify small examples with this rule alone, but it won't scale well. In fact the information we need to know to compute this is on the order of 2^n where n is the number of variables.

In the first part of this write-up, conditional independence is mentioned. A formal definition that expresses this conditional independence is:

$$P(a \wedge b|C) = P(a|C)P(b|C)$$

From this definition and Bayes' rule, many interesting expressions can be derived. One of which that I find interesting is when we have a set of n conditionally independent variables given a cause C:

$$P(C, x_1, x_2, \ldots, x_n) = P(C) \prod_{i=1}^{n} P(x_i|C)$$

This equation severely breaks down the information we need to know in order to compute this probability. Although in general, not every single one of the variables in our system will be conditionally independent, supposedly making this assumption can yield good results.

As noted above, it can be useful to make the false assumption that the variables are conditionally independent so they can be used in this equation, and for that reason the technique is called the Naive Bayes' model.