Sampling Methods for Action Selection in Influence Diagrams

Luis E. Ortiz
Computer Science Department
Brown University
Box 1910
Providence, RI 02912 USA
leo@cs.brown.edu

Leslie Pack Kaelbling
Artificial Intelligence Laboratory
Massachusetts Institute of Technology
545 Technology Square
Cambridge, MA 02139 USA
lpk@ai.mit.edu

Abstract

Sampling has become an important strategy for inference in belief networks. It can also be applied to the problem of selecting actions in influence diagrams. In this paper, we present methods with probabilistic guarantees of selecting a near-optimal action. We establish bounds on the number of samples required for the traditional method of estimating the utilities of the actions, then go on to extend the traditional method based on ideas from sequential analysis, generating a method requiring fewer samples. Finally, we exploit the intuition that equally good value estimates for each action are not required, to develop a heuristic method that achieves major reductions in required sample size. The heuristic method is validated empirically.

Introduction

The problem of decision-making involves the selection of an *optimal strategy*. A strategy determines how we should act based on observations or available information about the variables of the system relevant to the decision problem. Posed in the framework of decision theory, an optimal strategy is one that maximizes our utility. The utility defines our notion of value associated with the execution of actions and the states of the system. The states result from the combination of the state of the individual variables in the system. In the case of decision-making under uncertainty, we are uncertain about both the state of the system and the result of the actions we take. We express this uncertainty as probabilities. Therefore, in this context an *optimal strategy* is one that *maximizes our expected utility*.

In this paper our main interest is in decision problems under uncertainty formulated as *influence diagrams (ID)*. An *influence diagram* is a graphical model that provides a compact representation of (1) the probability distribution governing the states, (2) the structural strategy model representing how we make decisions, and (3) a utility model defining our notion of value associated with actions and states. We study the problem of selecting an optimal strategy in an influence diagram, concentrating on the case in which there is only one decision to be made. This is because we can decompose the problem of multiple decisions into many subproblems involving single decisions (i.e., by using the technique presented by Charnes & Shenoy (1999)). We note that we can apply methods developed to solve IDs of this kind to obtain methods to solve finite-horizon Markov decision processes (MDPs) and partially observable Markov decision processes (POMDPs) expressed as dynamic Bayesian networks (DBNs) (i.e., by modifying the technique presented by Kearns, Mansour, & Ng (1999)).

The problem of strategy selection involves the subproblem of selecting an *optimal action*, from the set of action choices available for that decision, *for each possible observation* available at the time of making the decision. Therefore, we want to select the action that maximizes the expected utility for each observation. One way to do action selection is to compute, exactly or approximately, the probabilities of the sub-states of the system directly relevant to our utility in order to evaluate the expected utility or *value* of each action. A sub-state is formed from the state of a subset of variables in the system. We believe this approach fails to take advantage of an important intuition: it only matters which action is best. Therefore, the problem of action selection is primarily one of comparing the values of the actions. We combine this with the intuition that actions that are close to optimal are also good. In this paper, we present methods for action selection in IDs that take advantage of these intuitions to make major gains in efficiency.

Notation

Before we present the definition of the ID model, we introduce some notation used throughout the paper. We denote one-dimensional random variables by capital letters and denote multi-dimensional random variables by bold capital letters. For instance, we denote a multi-dimensional random variable by $\boldsymbol{X}$ and denote all its components by $(X_1, \ldots, X_n)$ where X_i is the i^{th} one-dimensional random variable. We use small letters to denote assignments to random variables. For instance, $\boldsymbol{X} = \boldsymbol{x}$ means that for each component X_i of $\boldsymbol{X}$, $X_i = x_i$. We also denote by capital letters the nodes in a graph. We denote by $Pa(Y)$ the parents of node Y in a directed graph.

We now introduce notation that will become useful during the description of the methods presented in this paper. For any function h with variables $\boldsymbol{X}$ and $\boldsymbol{Z}$, the expression

$$h(\boldsymbol{X}, \boldsymbol{Z})|_{\boldsymbol{Z}=\boldsymbol{z}}$$