REASONING ABOUT CONTROL. THE INVESTIGATION OF AN EVIDENTIAL APPROACH *

Leonard P. Wesley

Artificial Intelligence Center
SRI International
Menlo Park, CA 94025

ABSTRACT

The complexity of the domains in which expert systems are expected to operate requires that they be capable of "reasoning*[1] about their actions It has been argued that expert systems must reason from *evidential* information i.e., *uncertain, incomplete,* and occasionally *inaccurate* information [LOW82a|. As a consequence, a model for reasoning about control must deal with several problems being able to organize "control-related" evidential information that is generically distinct and from disparate sources, to overcome minor errors in the evidential information needed to reach a decision, and to explain the actions taken by the system These are a few of some formidable control issues and problems that remain largely unsolved [BAR82] Thus, we report on an investigation into how these issues and problems can be addressed when the problem of reasoning about system control is viewed as an evidential process.

INTRODUCTION

Expert systems that operate in complex domains are continually confronted \M(h the problem of deciding what to do next. Furthermore, such systems must "reason" about their alternatives and choose an action on the basis of *uncertain, incomplete,* and *inaccurate* information, called *evidential* information (LOW82a). For instance, the decision to take a particular action can be influenced by the expected outcome of taking that action. However, situations may arise in which *uncertainty* exists about the consequences of taking any action, particularly when uncontrollable or unpredictable events may intervene. Resource limitations, for example, might not permit gathering all relevant facts, thus forcing decisions to be made with *incomplete* information. Finally, it must be anticipated that the information supplied to the system may be *inaccurate* because, among other reasons, the sources of the information are imperfect We, as well as the systems we build, must be capable of choosing an action on the basis of evidential information. Therefore, the on-going research reported here is concerned with some problems that must be dealt with before such capabilities can be realized.

Some Control Problems

Typically the information needed to choose between alternatives is obtained from multiple sources and varies in both the type and the unit of measurement For example, costs and goals/subgoals are two distinct types of information. One source might talk about costs in terms of CPU cycles, and another distinct source might talk about costs m terms of the degradation of the system's ability to complete its task A problem of interest is how can generically distinct types of information be combined to obtain a consensus of opinions, from disparate sources, about the appropriate action to take?

A second problem is that the information that is required to choose between alternatives may contain minor errors. No system can always measure accurately the costs of taking an action or determine precisely the goals/subgoals that should be satisfied. Yet we require that expert systems be robust enough to make effective decisions despite such errors What mechanisms can be employed to correct for minor errors?

It is import ant that expert systems be able to explain their actions. Sometimes decisions are based on information that tends to support those propositions in favor of choosing a particular action, tends to refute those propositions in competition with the action taken, or both (i e , conflicting information). At other times an action is taken because the system is partially ignorant about some aspect of the information required to reach any decision. In short, to explain its decisions, a system must be able to distinguish among evidence that tends to support, tends to refute, and neither supports nor refutes the propositions of interest What is an adequate representation of the total evidential information, as it bears on the propositions of interest, that will allow systems to provide more meaningful and accurate explanations of their actions?

ACTIONS, CONTROL-FEATURE SPACES, AND A CONTROL ENVIRONMENT

We consider an action to be the invocation of a parameterized process, such as a knowledge source (KS), where a KS is a procedure or sensor that makes observations about the environment or the invocation of processes for obtaining additional information that is required to choose an action.

Selecting the appropriate action depends on the observation of features in an environment. Just as spectral attributes can be considered features of objects, so can goals/subgoals be features of actions. If objects can be partially discerned from information about what spectral features are observed, then the appropriate actions can be partially

*This research was supported, in part, by the Defense Advanced Research Projects Agency (DAHPA) under contract number N00014-81-C-0115. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of SRI International, DARPA, or the U.S. Government.

**The author is currently enrolled in the department of Computer and Information Science (COINS), University of Massachusetts, Amherst, MA 01003.

discerned from information about which goals/subgoals need to be satisfied.

Examples of control-feature spaces are goals/subgoals, costs associated with invoking various KSs, the reliability of KSs, and so on. A control-feature space, therefore, is any observable or quantifiable aspect of an environment and system that can possibly help to discern the appropriate action. The set of all control-feature spaces of potential interest constitutes a "control environment". In general, a control environment includes any aspect of a system and its environment about which information may be obtained to assist in making a decision.

REASONING FROM EVIDENTIAL INFORMATION

Lowrance and Garvey have characterized uncertain, incomplete, and inaccurate information as evidential information that must be used by expert systems to reason about their environment [LOW82a].[1] Just as the information needed to interpret our environment is evidential in nature, so is the information needed to decide what to do next. That is, information about such things as goals, consequences of actions, the power of KSs, and how much is known about the environment is not easily expressed in terms of Boolean logic or probabilities. Rather, such information is evidence that tends to confirm or refute hypotheses about what control feature values have been observed, and these hypotheses in turn directly or indirectly imply what actions should be taken.

For our approach, the information required to reason about what to do next is obtained through a set of control knowledge sources (CKSs).[2] Similar to KSs, CKSs provide partially processed, "control-related," evidential information that is based upon their observations of the control environment.

A Formal View

Let A, the set of mutually exclusive and exhaustive actions that a system can possibly take, be defined as

$$A \quad = \quad \{a_1, a_2, \quad \ldots \quad , a_n\}$$

Let $F_1, F_2, \quad \ldots \quad , F_m$ constitute the control-feature spaces of interest.[3] Associated with each F_i, $1 \le i \le m$, is a set $\mathcal{F}_i$ of possible feature values of F_i, where

$$\mathcal{F}_i \quad = \quad \{ f_i \mid f_i \quad \text{is a possible feature value of} \quad F_i \}.$$

For each $f_i \in \mathcal{F}_i$, it is possible to identify a subset of actions for which each action is possibly the best one to take when f_i is observed. For instance, if an analysis of the situation indicates that the goal of getting from point A to point B as fast as possible should be satisfied, then flying or driving, as opposed to walking, is possibly the best action. Alternatively, if the goal of getting from point A to point B by relatively economical means should be satisfied, then walking is possibly the best action.

We can now construct a frame of action, Θ_a.

$$\Theta_a = \{ (a_j, f_1, f_2 \ldots f_m) \mid a_j \quad \text{is possibly the best action to take when} \quad f_1, f_2, \ldots, \quad and \quad f_m \quad \text{are observed} \},$$

$$\Theta_a \quad \subseteq \quad A \times \mathcal{F}_1 \times \mathcal{F}_2 \times \quad \ldots \quad \times \mathcal{F}_m.$$

Each $\mathcal{F}_i$ induces an equivalence relation over Θ_a that partitions Θ_a into $|\mathcal{F}_i|$ equivalence classes. Each $f_i \in \mathcal{F}_i$, therefore, is effectively an equivalence class whose members are those elements of Θ_a that exhibit feature value f_i. Therefore, $f_i \subseteq \Theta_a$, and as a consequence

$$\Theta_a \quad = \quad \bigcup_{f_i \in \mathcal{F}_i} f_i$$

In reality, a CKS cannot always convey its observations in terms of the individual f_i's. A CKS must be capable of communicating in terms of disjunctions of possibilities (e.g., through the proposition $f_1' \vee f_1''$ where $f_1', f_1'' \in \mathcal{F}_1$). Similarly, at times a CKS might need to express total ignorance about some observation (e.g., through a proposition corresponding to Θ_a). Consequently, a model for reasoning must incorporate those propositions with which CKSs wish to express their observations.

A system cannot always obtain the exact value for a subset of features, nor will it always know what subset of features is required to choose between alternative actions. At best, a system will only be able to induce and refine a partial ordering over a subset of alternatives to the extent that it is capable of realizing what control-feature spaces are required to discern the appropriate action and is able to reason from imperfect measurements of those features.

BODIES OF EVIDENCE

A CKS conveys its observations by providing a mass vector, which it derives from a body of evidence that tends to confirm or refute a subset of the propositions of interest. In an evidential approach, every CKS has a unit of "mass" that it may distribute, on the basis of its beliefs about what it has observed, among the various propositions. Given its observations, a CKS might believe that a subset of the propositions is partially or completely true. Such beliefs can be conveyed by attributing a proportionate amount of its unit mass directly to the truthfulness of those propositions. Conversely, a CKS can attribute a portion of its mass directly to the negation of a subset of propositions if it believes that they are partially or completely false.

Bayesian distributions and mass vectors

A mass distribution can be viewed as a generalized Bayesian distribution of belief over a set of propositions. A Bayesian distribution assigns a unit of belief over a set of mutually exclusive and exhaustive propositions, as designated by the mapping m.

$$m : \Theta_a \mapsto [0,1], \quad \text{where} \sum_{p \in \Theta_a} m(p) = 1.$$

---

[1] Lowrance's and Garvey's development of the concept of evidential reasoning is based, in part, on Shafer's [SHA76] extension of Dempster's [DEM67, DEM68] work on upper and lower probabilities.

[2] A CKS differs from a KS only in the scope of the environment and types of features it is expected to perceive.

[3] Because we are limited in our capacity to reason with large bodies of information, a premise of this research is that we deal with this dilemma by employing some methods for producing a manageable set of alternative actions and control-related information. We suggest that methods analogous to this one must be employed in expert systems.

The probability of any proposition, for instance $B \subseteq \Theta_a$, is the sum of the belief attributed to propositions that imply B or one minus the sum of the belief attributed to propositions that imply not B (i.e., $\neg B$).

$$\text{for all} \quad B \subseteq \Theta_a, \quad Prob(B) = \sum_{p \in B} m(p).$$

It follows that

$$Prob(B) = 1 - Prob(\neg B)$$

However, a mass distribution need not attribute belief to mutually exclusive and exhaustive propositions. That is, belief may be distributed as designated by the mapping M.

$$M : 2^{\Theta_a} \mapsto [0,1], \quad \text{where} \quad M(\emptyset) = 0 \quad \text{and},$$

$$\sum_{p \subseteq \Theta_a} M(p) \quad = \quad 1.$$

The sum of the mass attributed to propositions that imply B plus the sum of the mass attributed to propositions that imply $\neg B$ need not equal one, because some mass may have been assigned to propositions that imply neither (e.g., Θ_a).

Each body of evidence, therefore, induces an interval, called an "evidential interval", within which belief about a proposition must lie. An evidential interval is a subinterval of the real interval [0,1]. The lower and upper bounds of the evidential interval will be called support (Spt) and plausibility (Pls), respectively. The Spt represents the total mass that tends to support a proposition:

$$Spt(B) \quad = \quad \sum_{p \subseteq B} M(p).$$

The Pls represents the degree to which the mass fails to refute the proposition

$$Pls(B) \quad = \quad 1 - Spt(\neg B) \quad = \quad 1 - \sum_{p \subseteq \neg B} M(p).$$

The interpretations of some evidential intervals are summarized below:

Completely true $[1,1]$;

Completely false $[0,0]$;

Completely ignorant $[0,1]$;

Tends to support $[Spt, 1]$, $0 < Spt < 1$;

Tends to refute $[0, Pls]$, $0 < Pls < 1$;

Tends to support and refute $[Spt, Pls]$, $0 < Spt \leq Pls < 1$.

AN EVIDENTIAL APPROACH

The application of our approach begins by invoking a subset of the available CKSs in order to make observations of the system's control environment. Each CKS conveys its observations by providing a body of evidence, in the form of a mass vector, that expresses its belief in propositions such as

- *THE COST OF INVOKING KS1 WILL BE 100 CPU SECONDS,*
- *OUR GOAL SHOULD BE TO IDENTIFY OBJECTS CLOSEST TO US FIRST.*

An integral part of our approach is the use of Dempster's rule to combine bodies of evidence [DEM67, DEM68]. Given two arbitrarily complex mass vectors, say M_1 and M_2, Dempster's rule produces a third mass vector, M_3, that reflects a consensus of the opinions expressed in M_1 and M_2. That is,

$$M_3 \quad = \quad M_1 \oplus M_2.$$

Shafer [SHA76] and Lowrance [LOW82b] can be consulted to understand how multiple bodies of evidence can be combined. Furthermore, Dempster's rule is both commutative and associative. Therefore, the order and grouping of combinations is immaterial. This allows results to be obtained with whatever degree of parallelism the host hardware can support.

Dempster's rule accomplishes two functions. The first is to obtain a consensus about what actions each CKS believes is possibly the best to take. For instance, one CKS might provide evidence that pertains to the costs associated with invoking various KSs, whereas a second CKS may provide evidence that pertains to the reliability of invoking various KSs. Given the evidence they each provide, the task is to determine what actions both CKSs agree are the most appropriate to take.

$$\text{For all } p_1, p_2, p_3 \subseteq \Theta_a, \; M_3(p_3) = (1-K)^{-1} \sum_{p_1 \cap p_2 = p_3} M_1(p_1)M_2(p_2),$$

$$\text{where} \quad K = \sum_{p_1 \cap p_2 = \emptyset} M_1(p_1)M_2(p_2) < 1.$$

If both bodies of evidence are completely consistent, then there is at least one action that they both agree is possibly the best to take, and it can be said they are expressing completely compatible opinions, $K = 0$. Conversely, if the evidence they provide is not completely consistent, then their opinions are not completely compatible, $0 < K \leq 1$. Dempster's rule simultaneously determines which actions, if any, both CKSs agree are appropriate and provides a measure, K, of the compatibility between the two bodies of evidence.

The second function of Dempster's rule is to correct for minor errors that may occur in bodies of evidence. The assumption here is that the likelihood of distinct CKSs introducing the same type of error in their bodies of evidence is negligible. Therefore, any such errors can be overwhelmed by a sufficient amount of redundant and generally correct evidence. If a subset of CKSs makes gross errors, such bad information should be discarded if it is detected.

Next the result of applying Dempster's rule must be extrapolated from those propositions on which it bears directly to the remaining dependent propositions.[4] This extrapolation process is performed by an inference engine[5] that carries out its task by adjusting the bounds on

[4] Lowrance has developed a model, called dependency graphs, that are a formal representation of logical dependencies between propositions [LOW82b].

[5] The inference engine used in this approach is similar to that described in Lowrance's thesis [LOW82b].

the evidential interval that is associated with every dependent proposition.

After the extrapolation process terminates a partial ordering over the set of alternative actions is reflected in the evidential intervals associated with each a E 0_a. Selecting the appropriate action involves evaluating these evidential intervals. Although a complete decision theory for performing such an evaluation is not yet available, it is possible to choose actions on the basis of several simple criterion. For example, determining the best action is obvious for those propositions that correspond to alternatives that have evidential intervals that do not overlap For those propositions with overlapping evidential intervals, further evaluation is called for There are many utility- vs. cost-based theories that can be used to select an action on the basis of beliefs that are constrained by an evidential interval. However, a simple scheme might be to use the median of the evidential interval.

SUMMARY

This paper has described some important problems confronting expert systems that operate in complex domains pooling generically distinct evidential information, correcting for minor errors in evidential information, and realizing an adequate representation of the total evidence that tends to support, tends to refute, and neither supports nor refutes the propositions of interest. We have also described how an evidential approach to reasoning about control can address some of these problems.® Dempsters rule is a mechanism for obtaining a consensus about the appropriate actions to take and correcting minor errors: a system can better explain its actions because a confidence interval allows it to distinguish between supporting, refuting, and neutral evidential information.

ACKNOWLEDGMENTS

1 would like to thank Drs. John D. Lowrance and Thomas D. Garvey for their untiring guidance and extremely helpful suggestions

REFERENCES

|BAR82] Barnett, J A., "Some Issues of Control in Expert Systems,*[1] Proc International Conference on Cybernetics and Society, October 1082, pp. 1-5.

|DEM 67| Dempster, A P., "Upper and lower probabilities induced by a multivalued mapping," Annals of Mathematical Statistics, 38, 1967. pp. 32.V339.

[DEM08] Dempster, A.P., "A generalization of Bayesian inference," Journal of the Royal Statistical Society, Series B, 30, 1968, pp. 205-217.

[LOW82a| Lowrance, J.D. and T.D. Garvey, "Evidential Reasoning: A Developing Concept," Proc. International Conference on Cybernetics and Society, October 1982, pp. 6-9.

[LOW82b] Lowrance, J.D., "Dependency-Graph Models of Evidential Support," Ph.D. dissertation, Department of Computer and Information Science, University of Massachusetts, Amherst, MA, September 1982.

[SIIA76| Shafer, G.. *A Mathematical Theory of Evidence,* Princeton University Press, Princeton, New Jersey, 197G.

|WES8:i) Wesley, L P , "Reasoning About Control: An Evidential Approach," SRI International technical report in preparation.

[6]A more complete discussion of an evidential-based approach to addressing these and other problems can be found in |WES83).