In *Proceedings of the First International Conference on Artificial Intelligence Planning Systems* (AIPS), J. Hendler (ed.), Morgan Kaufmann Publishers, San Mateo, Cal., 1992.

Abstract Probabilistic Modeling of Action

Lonnie Chrisman
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213
chrisman@cs.cmu.edu

Abstract

Action models used in planning systems must necessarily be abstractions of reality. It is therefore natural to include estimates of ignorance and uncertainty as part of an action model. The standard approach of assigning a unique probability distribution over possible outcomes fares poorly in the presence of abstraction because many unmodeled variables are not governed by pure random chance. A constructive interpretation of probability based on abstracted worlds is developed and suggests modeling constraints on the outcome distribution of an action rather than just a single outcome distribution. A belief function representation of upper and lower probabilities is adopted, and a closed-form projection rule is introduced and shown to be correct.

1 Introduction

Models of actions used in A.I. systems for planning and reasoning must necessarily be abstractions of reality. The real world is too complex and too intricate to model perfectly in complete detail. Whenever we model the effects of actions, we are forced to omit detail which could, in some situations, significantly influence the resulting outcome. It is natural, therefore, to include estimates of ignorance and uncertainty as part of an action model. Probabilistic models in various forms provide tools for expressing this uncertainty.

The most natural form for probabilistic models comes from the framework of Bayesian probability, where a conditional probability on outcomes for an action is specified as $P(o|s)$, where o is the outcome of executing the action from situation s. A number of examples of recent A.I. planning research have employed models of this form ([Drummond and Bresina, 1990], [Kanazawa and Dean, 1989], [Hanks, 1990], [Chrisman and Simmons, 1991], [Christiansen and Goldberg, 1990]). Bayesian models provide a comfortable framework because the knowledge used (conditional probabilities) can be placed in correspondence with the causal structure of the system being modeled, and the probabilities can be easily interpreted. However, strict Bayesian models deal poorly with abstraction. This difficulty can be traced to the fundamental constructive interpretation of Bayesian probability [Shafer, 1981] that assumes unmodeled variables can always be compared to random chance processes. The next section demonstrates that this is a particularly bad assumption in a planning context. For example, unmodeled variables that influence outcome may depend upon previous actions an agent has taken, not upon random chance. Comparing action outcome to a random chance process is deceptive and can lead a planner to produce plans that work much differently in the real world than predicted. The correct way to address abstraction-related deficiencies in a Bayesian framework is to refine models to greater detail, an approach in direct conflict with the ubiquitous and often desirable demand to use abstract models.

This paper introduces an approach, designed to properly handle abstraction, for probabilistically modeling action. Section 3 introduces an alternative constructive interpretation of probability that adopts abstraction as fundamental. This suggests representing constraints on possible probability distributions over modeled variables rather than representing unique probability distributions. Section 4 introduces a model for action based on a belief function representation for lower probability. Finally, since projection is one of the most important uses of an action model, Section 5 presents a closed-form projection rule and two theorems demonstrating its correctness. The overall framework is a generalization of Bayesian probabilistic reasoning.

2 Abstraction Problems

Abstraction is the omission of detail. The presence of abstraction means that there will be some unmodeled factors that may influence the outcome of an action. Abstraction can arise in many different forms, all of which can be problematic to planning with strict Bayesian models which commit to complete specification of outcome probability distributions. The problems arise because many unmodeled variables are not governed by random chance. This section considers one specific form of abstraction, the omission of state variables, and presents an example to illustrate the problems.

I am considering the action of throwing a rock across a river, with the possible outcomes that the rock lands on the opposite shore or lands in the water. There are many factors that determine the outcome, including the rock's weight and coefficient of air friction, the wind velocity, my adrenaline level, the condition of my arm (warmed up, stretched out, injured, etc), my distance to the water's edge, and the quality of my footing. Many of these can be further refined to greater levels of detail, and many additional factors exist. Suppose, however, that some of this detail — for example, the quality of my footing — is omitted from my model for the throwing action. The Bayesian approach compensates for this missing information by adopting a probability distribution over the possible outcomes, reflecting in this case the expected frequencies that the rock will land in the water versus on the opposite shore. However, this distribution is not determined by the modeled variables alone. If I am wearing cleats on a grassy shore, the rock will be reasonably likely to clear the river, but when wearing flat bottom shoes on damp, moss-covered boulders, reaching the far side may be impossible. By committing to a single distribution, a Bayesian treatment does not admit ignorance (i.e., that there is any lack of knowledge about the actual outcome distribution). After accepting a Bayesian assessment that on any given throw consistent with the modeled situation, I will have a 15% chance of succeeding, I might happily produce a plan to repeatedly throw rocks until one lands on the far side. The model gives no indication that this may have no hope of ever succeeding. The problem here is that the quality of my footing is not governed by a random chance process, and modeling it as such is entirely misleading.

Unmodeled influences can be particularly bothersome for a planner when the agent's own decisions influence unmodeled variables. In this case, the outcome distribution really depends upon decisions made in the past, and different choices at those past decision points can result in different outcome distributions for the current action. My choice this morning to wear flat bottom shoes now affects the probability of getting a rock across the river. Just as some state variables will remain unmodeled, influences between actions will also go unmodeled. As a planner explores alternative projection paths, the distribution on effects of the same action in what appears to be identical situations (given the modeled variables) may be different. Pretending to know the distribution simply leads to the construction of plans that don't work as expected when actually executed in the real world.

3 Abstraction-Based Interpretation

The problems with Bayesian action models stem from implicit assumptions arising from the Bayesian interpretation of probability. Glenn Shafer [Shafer, 1981] has eloquently argued for a constructive view of probability. In the constructive view, the meaning of probability in any given formalism "amounts to comparing one's evidence to a scale of canonical examples, and a constructive theory of probability judgement must supply ... the scale of canonical examples." Random chance provides the canonical example for the Bayesian interpretation. If a variable does behave according to random chance, it is said to be aleatory. But as the previous section demonstrated, when abstraction is invoked in planning contexts, many variables might not be aleatory. In this section a different constructive view of probability is arrived at by accepting the abstracted grand aleatory world as the canonical example. Essentially, abstraction is assumed as a fundamental starting point for the development of this view, rather than random chance alone as in the Bayesian case.

Action models or state descriptions include certain (modeled) variables and necessarily omit others (unmodeled variables). A Bayesian model assumes that a fixed, stationary probability distribution exists over the set of modeled variables. Implicitly, this amounts to assuming that all unmodeled variables are aleatory, so that the modeled distribution is obtained from $P(mv) = \sum_{uv} P(mv|uv)P(uv)$, where mv are modeled values and uv are unmodeled values. To drop this assumption and obtain a new constructive interpretation for probability, assume that there is some (possibly infinite) set of unmodeled (non-aleatory) variables, such that if these were to be concatenated with our modeled variables, a fixed, stationary distribution would result. This entire collection of variables describes a grand aleatory world. Essentially, if we did have a complete model with no abstraction, then we would be able to model action outcomes with a probability distribution. When we consider only the modeled variables, assigning values to some of these variables does not imply that a unique probability distribution exists over the unassigned modeled variables because some unmodeled variables are non-aleatory. By considering all possible assignments to unmodeled variables, what we do get is a set of constraints defining a (possibly infinite) set of probability distributions over the unassigned modeled variables. This set of distributions properly describes the belief afforded to an abstract description.

In the example from the previous section, foothold quality is a non-aleatory unmodeled variable. Even though it (and other variables) are unmodeled, what we may know is that even in the best of all situations, we cannot possibly have more than a 30% probability of actually getting a rock across the river. As far as we know, the actual probability might be 30%, but it may also be 18%, 5% or even 0% (totally impossible). This can simply be expressed as $P(oppositeShore) \leq 0.3$. The use of constraints on probability distributions, rather than precisely specified distributions, provides one method for probabilistically modeling action outcomes, despite the problems introduced by abstraction. Because foothold quality cannot be compared to random chance, it is wrong to assume a single probability of success in this example. To handle examples such as this, strict Bayesian modeling would require the model itself to be refined all the way to a grand aleatory world — a process which could require infinite detail and contradicts the necessity (and our desire) to use abstract models.

Sets of probability functions are commonly referred to as lower probability distributions and can be traced back to [Good, 1962]. They have been studied mathematically by many people, but I have not seen the above constructive interpretation given previously. Usually, lower probability work assumes that an underlying distribution exists; conversely, the above interpretation does not assume a distribution to exist over the modeled space. Like some mathematical treatments of non-measurable sets, [Halpern and Fagin, 1990] do not assume that an underlying distribution exists, but they do not provide the scale of canonical examples necessary to qualify the formalism as a full constructive interpretation of probability.

4 Modeling Action

The previous section argued that abstract probabilistic models should utilize sets of distributions rather than always committing to exactly specified distributions. To apply this idea, it necessary to select a mechanism for representing sets of distributions. Since the number of such sets is uncountably infinite, the choice of representation scheme will necessarily limit the possible distributions that can be represented. [Fagin and Halpern, 1989], [Wasserman, 1990], and [Kyburg, 1987] have all advanced the use of belief function representations for conveniently representing convex[1] lower probability functions, and [Lemmer and Kyburg, 1991] proves that most of the important convex sets can be represented by belief functions. While it is easy to construct non-convex belief sets, the usefulness for planning or decision making of mechanisms which are general enough to represent non-convex sets of probabilities is dubious, so a belief function representation is adopted here. This framework is most similar to [Fagin and Halpern, 1989].

A belief function can be viewed as a type of data structure, much like a list is a type of data structure. As such, it is a useful mechanism both for representing lower probability, as it is used here, and for representing evidential support, as in the Dempster-Shafer theory [Shafer, 1976]. However, the semantics of what is represented differs in these two cases and should not be confused. For example, Dempster's rule of combination gives perfectly reasonable results when it is used to combine evidence, but it does not make sense to apply it to a lower probability interpretation such as the one advanced in this paper. The reader wishing to understand the difference between the representation of evidence and lower probability should read [Halpern and Fagin, 1990] and [Shafer, 1976]. For the purposes of the current discussion, the reader should just be aware that the Dempster-Shafer theory of evidence is not being utilized here, even though the same syntactic machinery is employed to represent belief in both cases.

A belief about the state of the world is modeled using a "marginal" belief function. This belief function simply specifies constraints on the possible probability distributions, thus characterizing the agent's beliefs about the state of the world. Let Ω_{pre} delimit a mutually exclusive and exhaustive propositional space of possible situations before an action's execution, called a *prestate frame of discernment*, and let Ω_{post} delimit a *poststate* (after the action's execution) *frame of discernment*. A prestate belief is represented internally by the mass-assignment $m_{pre} : 2^{\Omega_{pre}} \rightarrow [0,1]$ such that

$$\sum_{B \subseteq \Omega_{pre}} m_{pre}(B) = 1, \qquad m_{pre}(\emptyset) = 0$$

Usually only a few sets, called focal elements, will have non-zero mass-assignments. The mass-assignment specifies a set of constraints on the possible probability distributions describing the current state. For a set $B \subseteq \Omega_{pre}$, $m_{pre}(B)$ is the amount of probability mass that is constrained to the situations in B, but which may be freely allocated in any way between the

[1] A lower probability distribution is convex if any weighted average of any two consistent probability distributions is also consistent with the lower probability distribution.

situations in B. To simplify the notation, we will informally drop the subscripts and write $m(B)$ when it is clear that the set B refers to a prestate (B="before"). Belief and plausibility functions are defined as

$$Bel(B) = \sum_{C \subseteq B} m(C) \quad (1)$$

$$Pls(B) = \sum_{C \cap B \neq \emptyset} m(C) \quad (2)$$

The belief $Bel(B)$ is the total amount of probability mass that is trapped within B, and the plausibility $Pls(B)$ is the total amount of probability mass that is allowed in B. Each of the three functions m, Bel, and Pls uniquely determine the other two. For example, $Bel(B) = 1 - Pls(\bar{B})$ and m can be obtained from Bel using [Shafer, 1976]

$$m(B) = \sum_{C \subseteq B} (-1)^{|B-C|} Bel(C) \quad (3)$$

A probability assignment $P : \Omega_{pre} \rightarrow [0,1]$ is said to be *consistent* with Bel when $Bel(B) \leq \sum_{b \in B} P(b)$ for all $B \subseteq \Omega_{pre}$. The set of all consistent probability assignments is denoted by $\mathcal{P}$. It is well known [Shafer, 1976] that $Bel(B)$ and $Pls(B)$ are lower and upper bounds respectively on the probability of B:

$$Bel(B) = \inf_{P \in \mathcal{P}} \sum_{b \in B} P(b)$$

$$Pls(B) = \sup_{P \in \mathcal{P}} \sum_{b \in B} P(b)$$

The relationship between a belief function and a particular consistent Bayesian distribution can be given by specifying how the mass assignment $m(B)$ for each set B is reallocated amongst the basic elements of B. This mapping, called an *allocation mapping function*, simply demonstrates how a given consistent probability distribution satisfies the constraints specified by a belief function on the possible distributions. The allocation mapping function is characterized by the following lemma. The proof appears in the Appendix.

Lemma 1 *A probability distribution P is consistent with belief function Bel if and only if there exists an allocation mapping function $f : 2^{\Omega} \times \Omega \rightarrow [0,1]$ such that the following conditions hold:*

1. $\sum_{B \subseteq \Omega} f(B,b) = P(b) \qquad \forall b \in \Omega$
2. $\sum_{b \in \Omega} f(B,b) = \sum_{b \in B} f(B,b) = m(B) \qquad \forall B \subseteq \Omega$
3. $f(B,b) \geq 0$, *for all* $B \subseteq \Omega$, $b \in \Omega$
4. $f(B,b) = 0$ *if* $b \notin B$

The effects of actions are modeled using basic conditional belief functions, $Bel_{post|b} : 2^{\Omega_{post}} \rightarrow [0,1]$. To simplify notation, we write $Bel(A|b)$ when it is clear that A refers to a poststate (A="after"). This is the belief assigned to the poststate situation A given that the prestate situation is exactly b, where $b \in \Omega_{pre}$. $Bel(A|b)$ is represented internally by $m(A|b)$ such that

$$Bel(A|b) = \sum_{C \subseteq A} m(C|b) \quad (4)$$

Belief functions are generalizations of probability distributions. When only single element sets are assigned mass (i.e., $m_{pre}(B) = 0$ when B has more than one element), then Bel_{pre} represents an exact probability distribution and is called a Bayesian belief function. $Bel_{post|b}$ can similarly represent an exact conditional probability distribution. At the other extreme, when absolutely no information about the prestate is available, the vacuous belief function, where $m_{pre}(\Omega_{pre}) = 1$, can be used to represent a state of total ignorance. In between, the various assignments represent constraints on the possible probability distributions.

An action is modeled by associating one basic conditional belief function with each B_i, where $\{B_1, B_2, ..., B_m\}$ form a set partition on Ω_{pre}. An action model is given as

$$\begin{array}{ll} \forall b \in B_1, & Bel_{pre|b}(\cdot) = Bel_1(\cdot) \\ \forall b \in B_2, & Bel_{pre|b}(\cdot) = Bel_2(\cdot) \\ \vdots & \\ \forall b \in B_m, & Bel_{pre|b}(\cdot) = Bel_m(\cdot) \end{array}$$

Each line of the action model represents one context-dependent outcome, where B_i is the precondition specifying the context. A finer partition granularity corresponds to greater knowledge about which distinctions influence the outcome, and finer distinctions usually lead to more precise outcome specifications. Each belief function, Bel_i, is represented internally by explicitly specifying the corresponding mass assignments. As with a prestate description, we say a conditional probability assignment, $P(a|b)$, is consistent with an action model $Bel(A|b)$ when $Bel(A|b) \leq \sum_{a \in A} P(a|b)$ for all $A \subseteq \Omega_{post}$ and $b \in \Omega_{pre}$. The set of all consistent distributions is denoted by $\mathcal{P}_{post|b}$.

As a simple example of modeling rock throwing, take the possible prestates to be $\Omega_{pre} = \{light, heavy\}$. While this leaves out many important state variables, it also abstracts in another way by representing a quantitative variable (weight) qualitatively. The precise weight of the rock is important for determining the probability of success, but this precision is abstracted away by adopting only the qualitative distinctions of *light* and *heavy*. If it were the case that of all the light rocks I typically pick up, the actual weight obeyed some fixed distribution, then a single probability distribution would also be sufficient for expressing the outcome probabilities in the abstracted model. This is unlikely to be the case, however, because the distribution of weight depends upon such things as the choice of decision procedure that I used in selecting which rock to pick up. Therefore, as was also the case with omitted variables, a single distribution may not suffice for summarizing the outcome distribution of an action when a quantitative variable is abstracted into a qualitative one. Therefore, a belief function is used to express the abstract action model. Take the possible poststates to be $\Omega_{post} = \{oppositeShore, inWater\}$. An action description might look like:

$$\begin{array}{ll} \texttt{if } b \in \{light\}: & m(\{oppositeShore\}|b) = .15 \\ & m(\{inWater\}|b) = .7 \\ & m(\Omega_{post}|b) = .15 \\ \texttt{if } b \in \{heavy\}: & m(\{inWater\}|b) = .8 \\ & m(\Omega_{post}|b) = .2 \end{array}$$

The action model specifies that the probability of getting a light rock to the other side is between 15% and 30%, while the probability of throwing a heavy rock that far is less than 20% but may be impossible. Compare this to a Bayesian model that predicts that exactly 15% of the heavy rocks will land on the far side. Not only does the Bayesian model assert more precision than is actually warranted, it also gives no indication that the outcome may actually be impossible — information that could be critical to a planner.

5 Projection

Projection is the process of predicting the effects of an action's execution given an action model and prestate description. Projection is one of the primary uses for an action model, and is utilized directly by some planners ([Drummond and Bresina, 1990], [Hanks, 1990]). In this section, a closed-form projection operation for belief function representations is introduced, and two theorems prove its correctness. The method is a generalization of Bayesian projection, performed in a Bayesian framework using Jeffrey's rule:

$$P(a) = \sum_{b \in \Omega_{pre}} P(a|b)P(b) \qquad (5)$$

Given an action model as in the previous section, $Bel(A|b)$, and a belief about the prestate, $Bel(B)$, the projection operation derives a belief function, $Bel(A|Bel_{pre})$, that represents the belief about the poststate situation resulting from the action's execution. We will also make use of one additional conditional belief function, $Bel(A|B)$, given by

$$Bel(A|B) = \min_{b \in B} Bel(A|b) = \min_{b \in B} \sum_{C \subseteq A} m(C|b)$$

$Bel(A|B)$ is almost the same as Halpern and Fagin's conditional belief functions; however, there are several subtle, but important differences too detailed to cover here. $Bel(A|B)$ is not considered to be the action model — it is only a function that is computed from the given model $m(A|b)$. Notice that there are four different belief functions involved here. In each of the four cases, mass-assignments $m(A|b)$, $m(B)$, $m(A|Bel_{pre})$, and $m(A|B)$ represent belief functions in memory (see (3) and (4)).

Lemma 2 *If $P_{post|b}$ is consistent with $Bel_{post|b}$, then for all $b \in B$*

$$Bel(A|B) \leq P(A|b) = \sum_{a \in A} P(a|b)$$

Proof:

$$\begin{aligned} Bel(A|B) = \min_{b' \in B} Bel(A|b') \leq & \quad Bel(A|b) \ \forall b \in B \\ \leq & \quad P(A|b) \end{aligned}$$

□

The projection, represented by $m_{post|Bel_{pre}} : 2^{\Omega_{post}} \rightarrow [0,1]$, characterizes the belief about the state after the action is executed and is given by:

$$m(A|Bel_{pre}) = \sum_{B \subseteq \Omega_{pre}} m(A|B)m(B) \qquad (6)$$

The corresponding belief function can equivalently be written as

$$\begin{aligned} Bel(A|Bel_{pre}) &= \sum_{C \subseteq A} m(C|Bel_{pre}) \\ &= \sum_{B \subseteq \Omega_{pre}} Bel(A|B)m(B) \qquad (7) \end{aligned}$$

This post-state belief is really a marginal belief function, but Bel_{pre} is included in the notation to avoid confusion. Note the similarity between (6), (7), and the Bayesian projection rule (5). Just as the action model and prestate belief functions can be interpreted as representing constraints over the possible probability distributions, $Bel(A|Bel_{pre})$ also represents distribution constraints. We say a projected probability distribution, P, is consistent with a projected belief function, $Bel(A|Bel_{pre})$ if $Bel(A|Bel_{pre}) \leq \sum_{a \in A} P(a)$ for all $A \subseteq \Omega_{post}$. The set of all consistent projected probability distributions is denoted $\mathcal{P}_{post|Bel_{pre}}$. The following theorem establishes the completeness of the projection procedure.

Theorem 1 *Let P represent a Bayesian projection obtained by (5) from a Bayesian action model $P_{post|b}$ and a Bayesian prestate belief P_{pre}, both of which are consistent with $Bel_{post|b}$ and Bel_{pre} respectively. Then P is consistent with the projected belief function $Bel(A|Bel_{pre})$. Formally*

$$\left\{ \begin{array}{lll} P & : & P(a) = \sum_{b \in \Omega_{pre}} P(a|b)P(b), \\ & & P_{post|b} \in \mathcal{P}_{post|b}, P_{pre} \in \mathcal{P}_{pre} \end{array} \right\} \subseteq \mathcal{P}_{post|Bel_{pre}}$$

Proof: Assume P_{pre} and $P_{post|b}$ are consistent with Bel_{pre} and $Bel_{post|b}$ respectively. Applying Equation (7) and Lemmas 1(2), 2, 1(4), and 1(1) (in that order) yields:

$$\begin{aligned} Bel(A|Bel_{pre}) &= \sum_{B \subseteq \Omega_{pre}} Bel(A|B)m(B) \\ &= \sum_{B \subseteq \Omega_{pre}} Bel(A|B) \sum_{b \in B} f(B,b) \\ &\leq \sum_{B \subseteq \Omega_{pre}} \sum_{b \in B} P(A|b)f(B,b) \\ &= \sum_{B \subseteq \Omega_{pre}} \sum_{b \in \Omega_{pre}} P(A|b)f(B,b) \\ &= \sum_{b \in \Omega_{pre}} P(A|b)P(b) = P(A) \end{aligned}$$

Therefore, by definition, P is consistent with $Bel(A|Bel_B)$. □

This theorem verifies that the projection will not jump to conclusions that are not justified by the available information since valid poststate distributions will not be left out. However, this by itself is not a particularly strong guarantee. For example, the vacuous projection which always produces a poststate belief of total ignorance (i.e., $m(\Omega_{post}|Bel_{pre}) = 1$) also has this property. It is therefore of interest to explore whether the converse to Theorem 1 holds, i.e., whether $\mathcal{P}_{post|Bel_{pre}} \subseteq \{P : ...\}$. This would imply that $Bel(A|Bel_{pre})$ is the true convex projection. Unfortunately the converse does not hold. In fact, the true lower probability projection is not necessarily representable by a belief function, even if the prestate and the action model are represented by belief functions. This is demonstrated in the following example:

I am considering throwing a rock that is currently in my hand across the river. If it is a heavy rock, the odds are better than even that it'll land in the water. There is also at least a 20% chance that it lands on the opposite shore, and at least a 10% probability that I'll drop it without launching it at all. On the other hand, if it turns out to be light weight, the odds are better than even that it will land on the opposite shore, with absolutely no chance of accidentally dropping it and at least a 10% chance that it lands in the water. I haven't the slightest idea whether the rock I am about to throw is light or heavy, but I do know it is one of the two.

If we set $\Omega_{pre} = \{light, heavy\}$, $\Omega_{post} = \{inWater, oppositeShore, dropped\}$, the prestate belief is given by $m(\Omega_{pre}) = 1$, and the action is modeled by:

$$\begin{array}{ll} \texttt{if } b \in \{heavy\}: & m(\{inWater\}|b) = .5 \\ & m(\{oppositeShore\}|b) = .2 \\ & m(\{dropped\}|b) = .1 \\ & m(\Omega_{post}|b) = .2 \\ \texttt{if } b \in \{light\}: & m(\{inWater\}|b) = .1 \\ & m(\{oppositeShore\}|b) = .5 \\ & m(\{inWater, oppositeShore\}|b) = .4 \end{array}$$

Applying the projection rule yields the following poststate mass-assignment:

$$\begin{array}{l} m(\{inWater\}|Bel_{pre}) = .1 \\ m(\{oppositeShore\}|Bel_{pre}) = .2 \\ m(\{inWater, oppositeShore\}|Bel_{pre}) = .4 \\ m(\{oppositeShore, dropped\}|Bel_{pre}) = .1 \\ m(\Omega_{post}|Bel_{pre}) = .2 \end{array}$$

This assignment coincides with intuition since the resulting Bel and Pls correctly bound the possible probability distributions. For example, the solution indicates that there is a 10% belief and 70% plausibility that the rock lands in the water. In a grand aleatory world for this situation, it may be the case that there is a 40% probability that the rock is light, a 30% chance that a light rock lands in the water, and a 60% chance that a heavy rock lands in the water. Notice that these are consistent with the problem. If this is the case, then the actual probability of landing in the water is 48%, which is clearly between the projected bounds of 10% and 70%. However, despite the intuitive appeal, the probability distribution

$$\begin{array}{l} P(inWater) = .1 \\ P(dropped) = .3 \\ P(oppositeShore) = .6 \end{array}$$

is also consistent with the projected belief function, yet it cannot be generated by any consistent Bayesian prestate-action model pair. In fact, the true poststate of this example cannot be exactly represented with a belief function.

One method for handling the above example exactly would be to adopt a representation scheme that is even more general than belief functions — for example, a scheme capable of representing certain non-convex sets of distributions. At this point, however, it is unclear how the additional generality could be of any use in the context of planning and decision making. Therefore, we keep the belief function based projection, even though a small amount of information is lost. Given that the representation is limited to belief functions, the following theorem shows that the projection rule is the best projection operation possible.

Theorem 2 *Let $A \subseteq \Omega_{post}$ be any set in Ω_{post}. There exists a consistent Bayesian prestate–action model pair, P_{pre} and $P_{post|b}$, such that*

$$P(A) = \sum_{b \in \Omega_{pre}} P_{post|b}(A|b) P_{pre}(b) = Bel(A|Bel_{pre})$$

Proof: Introduce the ordering $b_1, b_2, ..., b_n$ for all $b_i \in \Omega_{pre}$ such that $Bel(A|b_1) \leq Bel(A|b_2) \leq ... \leq Bel(A|b_n)$. Let $B_i = \{b_{i+1}, b_{i+2}, ..., b_n\}$, $B_n = \emptyset$. Then rewriting the sum in (7):

$$\begin{aligned} Bel(A|Bel_{pre}) &= \sum_{i=1}^{n} \sum_{B \subseteq B_i} Bel(A|\{b_i\} \cup B) m(\{b_i\} \cup B) \\ &= \sum_{i=1}^{n} Bel(A|b_i) \sum_{B \subseteq B_i} m(\{b_i\} \cup B) \\ &= \sum_{i=1}^{n} P(A|b_i) P(b_i) = P(A) \end{aligned}$$

where $P(A|b_i)$ was selected as equal to $Bel(A|b_i)$, and $P(b_i)$ was chosen by reallocating each mass $m(B)$ to the smallest element $b_i \in B$ as determined by the introduced ordering. □

The theorem implies that if the projected poststate belief function is tightened in any way, then valid Bayesian poststate distributions will be left out. Theorem 1 implies that if it were weakened in any way, impossible distributions would be let in. Therefore, the projection rule is the best one can hope for if one is not willing to use more complex representations. Geometrically, the projection rule finds the convex hull of the possible poststate probability distributions.

6 Levels of Abstraction

It is usually possible to model a system at many different levels of detail. As more detail is left unmodeled, it is said that the model is at a greater level of abstraction. Increasing the level of abstraction of an action model decreases the precision at which predictions into the future can be made using the model.

The effects of abstracting too much can be easily spotted in a lower probability representation where ignorance is explicitly represented. For example, if an abstract model omits the most important variables in a given problem, after projecting through only a few consecutive actions, the agent may be left with the vacuous poststate belief (i.e., complete ignorance). In general, the distance into the future at which an agent can usefully make predictions (even statistical predictions) is limited by the level of abstraction of its models.

There is a distinct difference between how the lower probability approach handles the case of overly abstract models from how the Bayesian approach treats this case. If a Bayesian model is projected into the future to a point where unmodeled, non-aleatory variables significantly influence the final poststate, the Bayesian model still makes a precise statistical claim about the final poststate distribution without any explicit indication that the stated distribution lacks accuracy. If the actual outcome frequencies are measured, they are likely to be quite different from the predicted distribution (one example of this discrepancy actually occurring appears in [Christiansen and Goldberg, 1990]). On the other hand, the lower probability approach gives very loose bounds in this case, effectively admitting that it doesn't know precisely what to expect. It is important to realize that the actual accuracy is equivalent in the two cases, being determined not by the formalism, but by the level of abstraction in the models. The primary difference here between the formalisms is whether or not the level of precision is explicitly indicated. This also suggests that the extra information encoded by a belief function may be extremely valuable to any system that automatically adjusts its own level of abstraction during reasoning.

7 Conclusion

Current techniques for using probabilistic models for planning are weak and applicable only to fairly small or well behaved problems. Despite the development of Bayesian networks [Pearl, 1988] for concisely representing probability distributions, the ability to attack probabilistic planning problems with extremely large state spaces, the type typically of interest to A.I. researchers, will require the ability to properly handle abstract probabilistic models. The constructive framework presented here addresses these concerns. Conditional independence statements can still be asserted and used in the lower probability framework, for example, to address concerns related to the frame problem (cf. [Wellman, 1990]). However, the framework may also offer significant advantages for inference networks, since it may potentially allow "almost statistically independent" influences to be abstracted away, a feature which could be used to greatly reduce the number of predecessors of a node when it is more important to obtain a quick answer than it is to obtain a precise answer. This is one of the many interesting areas for future research.

One of my own research interests is closed-loop planning, where the planner considers, at planning time, whether or not to commit resources to obtain additional information at execution time [Chrisman and Simmons, 1991]. Not observing a state variable turns out to be a lot like abstracting a model. Since the unobserved state variable can usually be affected by the agent's own behavior, unobserved variables are not aleatory. The current framework arose out of an effort to deal with unobserved variables efficiently without having to continue reasoning about unobserved details as the strict Bayesian tools demand. Along with the projection rule, a closed-form conditioning rule, a generalization of Bayes' rule, has also been developed with similar associated theorems, thus allowing observations to be incorporated into the agent's belief about the state of the world (see also [Fagin and Halpern, 1989], [Wasserman, 1990]).

The most important area for future research is the development of effective probabilistic planning algorithms that use these abstract action models. Given the projection rule presented here, it should be fairly easy to adapt temporal projection algorithms ([Drummond and Bresina, 1990], [Hanks, 1990]) to the new representations. Additionally, the introduction of the extra representational power may contribute important capabilities that could significantly extend the state of the art of temporal projection algorithms. For example, Hank's algorithm bundles outcomes in order to obtain small projection graphs. The generalized probability representation may additionally allow action choices to be bundled as well. The result would be a form of least-commitment probabilistic temporal projection.

Models used by planners will always be abstractions of reality. Probability provides an important tool for representing uncertainty, but because many unmodeled influences are not governed by random chance, simply assuming a unique probability distribution over the possible action outcomes may work poorly. The new framework advocates representing constraints on outcome probability distributions rather than just single probability distributions. The extra representational power may be valuable for capturing ill-behaved and poorly understood actions while avoiding the misleading appearance of having more knowledge about the projected effects of an action than is actually the case.

Acknowledgements

I thank Sebastian Thrun for taking the time to examine my proofs (any remaining mistakes are my own), and to Reid Simmons, Sebastian Thrun, Rich Goodwin, and anonymous reviewers for comments on earlier drafts of the paper. The research was sponsored by NASA under contract number NAGW-1175. The views and conclusions contained in this paper are my own and should not be interpreted as representing the official policies, either expressed or implied, of NASA or the U.S. government.

References

[Chrisman and Simmons, 1991] Lonnie Chrisman and Reid Simmons. Sensible planning: Focusing perceptual attention. In *Ninth National Conference on Artificial Intelligence*, July 1991.

[Christiansen and Goldberg, 1990] Alan D. Christiansen and Kenneth Y. Goldberg. Robotic manipulation planning with stochastic actions. In *DARPA Workshop on Innovative Approaches to Planning, Scheduling, and Control*, San Diego, CA, November 1990.

[Drummond and Bresina, 1990] Mark Drummond and John Bresina. Anytime synthetic projection: Maximizing the probability of goal satisfaction. In *Eight National Conference on Artificial Intelligence*, July 1990.

[Fagin and Halpern, 1989] Ronald Fagin and Joseph Y. Halpern. A new approach to updating beliefs. Technical Report RJ 7222, IBM, May 1989.

[Good, 1962] I. J. Good. Subjective probability as the measure of a non-measurable set. In Nagel, Suppes, and Tarski, editors, *Logic, Methodology, and Philosophy of Science*, pages 319–329. Stanford University Press, 1962.

[Halpern and Fagin, 1990] Joseph Y. Halpern and Ronald Fagin. Two views of belief: Belief as generalized probability and belief as evidence. In *Eighth National Conference on Artificial Intelligence*, July 1990. A long version appears as IBM-TR-RJ-7221, 12/18/89.

[Hanks, 1990] Steve Hanks. Practical temporal projection. In *Eighth National Conference on Artificial Intelligence*, July 1990.

[Kanazawa and Dean, 1989] Keiji Kanazawa and Thomas Dean. A model for projection and action. In *Eleventh International Joint Conference on Artificial Intelligence*, August 1989.

[Kyburg, 1987] Henry E. Kyburg, Jr. Bayesian and non-bayesian evidential updating. *Artificial Intelligence*, 31:271–293, 1987.

[Lemmer and Kyburg, 1991] John F. Lemmer and Henry E. Kyburg, Jr. Conditions for the existence of belief functions corresponding to intervals of belief. In *Ninth National Conference on Artificial Intelligence*, July 1991.

[Pearl, 1988] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann, San Mateo, CA, 1988.

[Shafer, 1976] Glenn Shafer. *A Mathematical Theory of Evidence.* Princeton University Press, 1976.

[Shafer, 1981] Glenn Shafer. Constructive probability. *Synthese*, 48:1–60, 1981.

[Wasserman, 1990] Larry A. Wasserman. Prior envelopes based on belief functions. *Annals of Statistics*, 18:454–464, 1990.

[Wellman, 1990] Michael P. Wellman. The STRIPS assumption for planning under uncertainty. In *Proceedings Eighth National Conference on Artificial Intelligence*, July 1990.

Appendix: Proof of Lemma 1

Proof: First, assume the conditions hold for some f. Then

$$\begin{aligned} Bel(B) &= \sum_{C\subseteq B} m(C) = \sum_{C\subseteq B}\sum_{b\in C} f(C,b) \\ &= \sum_{b\in B}\sum_{C:b\in C, C\subseteq B} f(C,b) \\ &\leq \sum_{b\in B}\sum_{C\subseteq \Omega} f(C,b) = \sum_{b\in B} P(b) \end{aligned}$$

Therefore, P is consistent with Bel.

For the other direction, assume P is consistent with Bel. Pick a set $B \subseteq \Omega$ and an element $b \in B$ and change the belief function by transferring mass in the amount of $f(B,b)$ from B to $\{b\}$, where

$$f(B,b) = \min m(B), \min_{C\subseteq\Omega}\{P(C) - Bel(C) : B \not\subseteq C, b \in C\}$$

Repeatedly apply the transfer for every pair of $B \subseteq \Omega$ and $b \in B$, at each step modifying Bel by the shift in mass and using the modified belief to evaluate $f(B,b)$ for successive pairs. Take $f(B,b) = 0$ for all $b \notin B$. Notice from (1) that the only sets C for which the transfer will alter $Bel(C)$ are the sets where $B \not\subseteq C$ and $b \in C$. For these sets, the minimization ensures that the new $Bel(C)$ continues to be less than or equal to $P(C)$. Therefore, P will continue to be consistent with Bel after each mass transfer.

Next, it is shown that all the mass will eventually be transferred to single element subsets, therefore exactly specifying the probability distribution P. After all the mass has been transferred, suppose there is a set with more than one element where $m(B) > 0$. Then for each $b \in B$, there is a set $C_b : B \not\subseteq C_b \wedge b \in C_b$ such that $Bel(C_b) = P(C_b)$. But these sets $\{C_b | b \in B\}$ totally account for all the mass in $\hat{C} = \bigcup_{b\in B} C_b$. However, $m(B)$ also contributes to the belief of $\hat{C}$, so that if $m(B) > 0$, then the mass in $\hat{C}$ is overaccounted for, i.e., $Bel(\hat{C}) > P(\hat{C})$. This contradicts the fact that P is consistent; therefore, $m(B)$ must be zero for all B with more than one element. The resulting Bel is a Bayesian belief function, and the total amount of mass assigned to $m(\{b\})$ is $\sum_{B\subseteq\Omega} f(B,b)$, thus proving condition 1. Since all the mass is transferred out of B to the elements $b \in B$, condition 2 is proved. The third and fourth conditions follow directly from the specification of $f(B,b)$ above. □