TR-I-0258

# A Basic Introduction to Planning and Meta-Decision-Making with Uncertain Nondeterministic Actions using Second-Order Probabilities

二次確率のインントロダクション

John K. Myers

真龍主·星音

March 3,1992

*Abstract*

The first part of this paper presents a basic introduction to second-order probability theory: what it is, and what kinds of problems it is used to solve.

The remaining parts of this paper introduce a system of theories and implementations for *planning* and *meta-decision-making* with uncertain-outcome actions represented using *second-order probabilities.* A situation-based theory of representing states, situations, and nondeterministic actions is implemented by the ATMS-based B-SURE system, which supports uncertain-action planning. A second-order probability theory allows an agent to model a probable continuum of universes, only one of which is correct; each universe describes a set of possible worlds labeled with probabilities. This represents the difference between the *likelihood* of an outcome and the *confidence* with which that likelihood is believed. Confidence is shown to govern meta-decision making, particularly meta-decisions concerning the gathering of information to clarify outcomes' likelihoods. Second-order probabilities are defined over *partitions* and individual *events.* Event distributions are convenient but cannot be used for accurate union nor expectation-distribution conputation. Partition and event distributions are initialized using maximum-entropy methods which significantly do not require prior frequency information, and are updated using Bayesian methods. *Value of Information* equations are defined, which support meta-decisions. A simple example demonstrates meta-decision-making with the implemented system. No other system has used second-order partition probabilities for planning or meta-decision-making.

# Contents

# 1 Basic Introduction

This paper presents an introduction to the theory of second-order probability, and how it can be used in uncertain planning, decision-making, and meta-decision-making.

## 1.1 The Problems

In order to understand second-order probability well, it is useful to review the kinds of problems it was designed to solve.

### 1.1.1 Representing Uncertainty Explicitly: Degrees of Confidence

A friend asks you to play some games with betting on a coin flip, whether it lands "heads" or "tails". The coin could be weighted, so you flip it twice to test it out–it comes up heads once, and tails once. You conclude that the probability of heads is 0.5. However, this test does not seem to be very satisfying–you want to test the coin out some more. In fact, you might not want to play with the coin unless you are allowed to test it more. After 1000 flips, you find that the coin landed heads 500 times, and tails 500 times. Now you are quite confident that the probability of heads is 0.5, and you are ready to start betting with your friend.

What has changed here? In both the "before" case and the "after" case, *the probability estimate is the same.* However, the cases seem quite different, based on what you are willing to do. The difference can be characterized by the *confidence* in the probability estimate. How can this be represented explicitly?

### 1.1.2 Estimating Probabilities with Few Samples

Your friend asks you to play a game with a 6-sided die. The die definitely has six sides, but it could be weighted. You throw it twice to check it out, and get a two and a six. Remembering from basic mathematics that probability should be estimated by $\frac{k}{n}$, the number of observed occurrences divided by the number of observations, you conclude that this die will roll twos $\frac{1}{2}$ of the time and sixes $\frac{1}{2}$ of the time, and ones, threes, fours, and fives will never come up (zero probability).

After rolling the die for 1000 times, you can tell that it definitely is weighted, because the "one" has only come up 3 times. You conclude that the probability of getting a one is $\frac{3}{1000}$.

What is wrong with estimating probability in this way?

### 1.1.3 Deciding When To Stop

A scheduling program is overseeing the processing of a list of a large number of alternatives, which are ranked by quality estimates. There are too many alternatives and they take too long to process for the scheduler to comfortably process them all. In addition, the best alternative is probably somewhere at the top of the list, and it is a waste of time to process all of the remaining alternatives, since only the best one will be chosen. At any one time, it is easy to pick the best result out of all the alternatives that have been processed so far–this is a *decision* problem. It is not so easy to determine when to stop processing alternatives and choose the best one, or whether to keep going–this is a *meta-decision* problem. How can this problem be formulated, and what information is needed to solve it?

### 1.1.4 Uncertain Planning

Given actions that have uncertain outcomes with chances that are not known well, how can an agent plan what course of action to take? And, how can a computer recognize and predict the plans of that agent?

**How is Second-Order Probability Useful?** In summary: Second-order probability is useful for the following tasks:

- Representing uncertainty in a precise and accurate manner
- Correctly representing probabilities derived from information obtained from a low number of sampling experiments, or a low number of successful outcomes
- Planning with uncertain actions, and performing plan recognition on people planning with uncertain actions
- Decision-making with uncertain quantities
- Meta-decision-making about whether it is time to stop and make a decision now, or whether it is important to continue to gather information about probabilities

## 1.2 The Most Important Formula

The simplest and most important result of second-order probability theory has to do with first-order probability. Based on results shown using second-order confidences, maximum-entropy initializations, and average (effective) probability, the following results can be proven:

> Given a situation in which $n$ outcomes $A_1 ... A_n$ are definitely known a priori to be possible
> and nothing else is known about the situation
> and a total of $m$ trial experiments have already been performed
> and the $A_i$'th outcome has definitely been observed to occur $k_i$ times

(and other outcomes have definitely been observed to occur $(m-k_i)$ times)

then the best current estimate for the probability $p(A_i)$ of the $A_i$'th outcome occuring on the next trial is:

$$p(A_i) = \frac{k_i + 1}{m + n} \tag{1}$$

For instance, this shows that the current estimated probability of a two or a six in the previous dice problem is $\frac{2}{8}$ apiece, and the estimated probability of a one, three, four, or five is $\frac{1}{8}$ apiece. The estimated probability of rolling a one is $\frac{4}{1006} = 0.0040$ in the other problem, not 0.0030.

The estimation is optimal in the sense that it is the center of gravity of all probable estimations, given all the information that is available at the moment.

This result is significantly different from the customary $\frac{k}{m}$ estimate when $m$ is small or when $k$ is small (under 10). A person who bets using this estimator will make more money in the long run than a person who bets using the old estimator. Note that both estimators approach the same limit as $m$ approaches infinity.

## 1.3 What Is Second-Order Probability Theory?

Second-order probability theory basically consists of five things:

1. A theory of what probability is, and how it is used;

2. A method of representing probability, based on a continuous function that represents the (second-order) probability that any one particular (first-order) probability is real;

3. A method of initializing the representation in (2) before performing any experiments, and of updating a representation based on the known results of experiments;

4. An interpretation of the mathematics in (2) and (3), that says that the first-order probability should be interpreted as an estimate of the *likelihood* that an action will have a particular outcome, the second-order probability should be interpreted as the *confidence* with which this estimate is believed,

5. A series of formulas that show how to use the mathematics in (2) and (3) for problems in search, planning, and decision-making.

These points will be briefly discussed, in order, in the following sections.

## 1.4 What Is Probability, and How is it Used?

Probability, and second-order probability, were both developed to deal primarily with uncertain outcomes that occur when a nondeterministic action is performed. A "nondeterministic" action is an action that does not have one fixed outcome that can be determined ahead of time, before the action is performed. Second-order probability theory is based on the following basic assumptions:

1. A nondeterministic action is in fact nondeterministic. There are several possible outcomes to the action, each of which in fact could happen.

2. A nondeterministic action has a fixed, constant, stationary numerical probability for each outcome, that corresponds to the actual likelihood that the outcome will occur if the action is executed.[1] This number exists and is unique.

   This number is known as the *real probability.* It is typically unknowable with complete accuracy. Note that it is not necessary for the action to be performed many times, or even to be performed at all, in order for a real probability to exist.

3. Based on all information known about the system, an ideal (unlimited) observer can determine *estimates* of the real probability. Such an estimate is known as a *believed probability* because it is an agent's opinion of the real probability, not the real probability itself. An ideal observer's opinion is called a *normative probability*, and it corresponds to what every agent ideally *should* believe about the probability of an action. Because the real probability exists in all cases, normative probability estimates should exist in all cases.

4. An actual human observer may or may not have an opinion about a real probability. Such a believed probability is called a *doxastic probability*, because it describes what the person *actually* believes, not what he or she *should* believe.

It is possible to weaken assumption 1 to a relativistic version, without apparent damage to the rest of the system:

> 1A. A nondeterministic action is nondeterministic to the performing/deciding agent. As far as this observing person can determine, there are several possible outcomes to the action, each of which could happen.

Since the system is concerned with believed probabilities, it apparently does not matter whether the action is "in fact" nondeterministic, or is only believed to be nondeterministic by the observer.[2]

[1] Nonstationary processes are not treated by this theory yet.

[2] Interesting possibilities occur when observers have access to different information, and one observer believes the system is nondeterministic while another observer believes it is

### 1.4.1 Other theories

It is interesting and useful to compare this theory of probability against other prevalent theories.

**The Frequency Definition.** One main definition of probability is the number that results from taking the limit of the relative frequency of observations versus the total number of experiments as the number of experiments approaches infinity [vM57, p.15,221] [Fre71, p.36]. This interpretation is well-grounded in reality.

The main problem with this definition is that the probability of an experiment that cannot be repeated a large number of times is undefined. Also, it is necessary to perform the experiment a "large number" of times before an accurate probability can be determined. Finally, a completely accurate probability cannot be determined without repeating the experiment an infinite number of times.

**The Subjective Definition.** A second main definition of probability is the percentage that a person would feel comfortable with when asked to bet money on the outcome of the experiment [Fre71, p.36].

This definition seems to be equivalent to our doxastic probabilities. However, it ignores the fact that doxastic probabilities depend on the way that the problem is defined, and do not have to be unique. The person may not feel comfortable betting at all. The main problem with this definition is that a subjective probability can vary arbitrarily from person to person, and even within the same person from moment to moment. There is no objective definition.

**The Set-Theoretic Definition.** A third main definition of probability is grounded in mathematics. Sets are divided up into subsets and elements corresponding with events; these are assigned numbers which are called "probability measures". A probability is a number that is assigned to a set. A set of three axioms allows derivation of useful equations for working with the numbers. This system only tells how to compute with probabilities; it does not say what they are, or where they come from. In particular, it does not determine how to derive the probability measures in the first place [Fre71, p. 39].

The main problem with this system is that probabilities cannot be determined in the system; it is necessary to reach outside of the system to initialize the probability measures. Probabilities are not grounded in the real world. In addition, there is no way to represent confidence.

## 1.5 How is Second-Order Probability Represented?

A second-order probability representation is quite simple to understand. For any one problem, there is a known set of $n$ possible outcomes $A_1..A_n$. Each

deterministic. These situations will not be covered in this paper.

outcome can, in theory, have any probability between 0.0 and 1.0. So this forms an $n$-dimensional problem space, where each axis represents a probability of an outcome, and each axis only varies from 0.0 to 1.0. For instance, if there are three possible outcomes, this forms a three-dimensional space. Any one point in the space corresponds to a definite labeling of the probabilities of each of the three outcomes. Of course, if we want to pay attention to the fact that all of the probabilities have to add up to one, then not all of the points in the space will be valid–most will be invalid, and the only usable points will lie on the surface of an equilateral triangle (in three-dimensional space).

Of course, a point can also be considered to be a vector. We will like some of these probability assignments better than other assignment points (vectors), and so we will assign a function to each point to grade how much we like it. This will be a vector density function, that maps each point (vector) in the space into a scalar, which represents confidence. For convenience, this number will also be a probability (i.e. the function is a continuous probability density function) that measures the likelihood or the confidence with which we believe this probability estimate could be true. Because this is a probability of a probability, the theory is called "second-order probability".

Explicitly representing an entire space like this (called the *partition probability distribution* function) would be too difficult to do inside present-day computers–it would take up too much space. Eventually it will be possible to work directly with equations, and not have to worry about numerical representations. However, at present, it is necessary to use a condensed representation. For this reason, the space is projected onto each of the $n$ axes, and then $n$ separate one-dimensional functions, called *event probability distributions*, are represented using arrays. This representation is exact for some problems, but loses information for other problems.

Thus, there are two representations for second-order probability, the partition distribution and the event distribution. It will be necessary to develop formulas for each type of representation.

The normal first-order probability of an outcome can be found by taking the average of the outcome's event distribution. However, it is possible to find other, more important information by using these distributions, that cannot be found using normal scalar probabilities.

## 1.6 How is Second-Order Probability Initialized?

Second-order probability should be initialized using maximum entropy theory. Maximum entropy says that there are all these possible values that the probabilities could take on, but at the beginning there is no known reason to favor any one estimate over any other estimate. We are equally ignorant on all estimates. So, the best thing to do is to act as if they are equally likely. Of course, this is just an estimate; we do not believe they actually ARE equally likely, we just don't have any information to tell us which ones are more important. As a matter of fact, the mathematics tells us this–we have an extremely low confidence in all estimates.

In order to set up a problem like this, it is first important to sit down and decide which outcomes are possible, and which are impossible. Include all the possible ones, and leave out all the impossible ones. This is important information for the system–and, as a matter of fact, the mathematics presented here assumes that this is all the information that you have.

As we make more and more experiments, the probabilities adjust, and a hill grows somewhere in the middle of each event distribution. We become more and more certain about the probability estimates of each outcome as we gather more data. But, we can always be surprised. The distribution never goes *completely* to zero at any one point (except at p=1, the endpoint)–there is always a possibility that the current distribution estimate could be wrong, and the hill could shift over.

## 1.7 Interpretation: Likelihood vs. Confidence

Once a distinction has been drawn between the likelihood that something will happen, and the confidence with which we believe that likelihood, then certain problems can be seen to depend on likelihood, while others depend on confidence.

It is still a research issue to separate these out and to pin down how they can be used in general situations. Certain specific situations are presented in the paper.

## 1.8 The Formulas

The formulas are presented in the main body of the paper. It is not important to understand them unless you want to use them.

## 1.9 Discussion: Why is Second-Order Probability more important than First-Order Probability or Interval Probabilities?

Simple first-order probability has been used for a long time to represent uncertainty about estimates of an action's outcomes. However, in the past 25 years, people have grown dissatisfied with this representation, mostly because they feel that it does not represent all that they know about the estimate. In particular, sometimes people feel quite confident about a probability estimate, and sometimes they do not feel confident at all. Thus, one number is insufficient information to represent this phenomenon.

It is easy to give an example to illustrate this. If I flip a coin two times, and it comes up "heads" once and "tails" once, then I can say that I believe that the probability of heads is 0.50. However, I do not feel very confident about this–I would not want to bet a lot of money on whether the coin comes up heads 5 times out of 10. The coin could be biased, and I might lose. If, however, I flip the coin 1000 times, and 500 times it comes up "heads" while 500 times it comes up "tails", then I would feel quite confident about betting on the coin. I still think the probability of heads is 0.50, but the two situations are different. Something else is needed in order to represent confidence.

People have tried using interval probabilities to represent this lack of certainty or lack of confidence. An interval probability consists of two numbers, a lower bound and an upper bound, that serve to model and to restrict the probability. The real probability is assumed to be somewhere inside this interval.

However, although interval probabilities are easy to compute with, they are discontinuous and not accurate. If a probability is represented as being between 0.25 and 0.75, does that mean that it is perfectly alright for the probability to be 0.25001, but that it can't be 0.24999? Why such a large difference? And, is it really true that the probabilities 0.2501 and 0.50 are equally likely?

In the past, doxastic interval probabilities have been determined by asking people what a low guess is for the probability, and what a high guess is. However, people normally feel hesitant about providing exact numbers for these bounds. A good reason for this could be that people actually are using a smooth second-order function to represent the probabilities, perhaps in a neural network. When asked to determine the bounds for a probability, they might cut the curve off at about the 95% level on both sides. However, since the actual functions are smooth, it is hard to tell where to place the boundaries–the interval probability is only roughly approximating the information found in the second-order probability distribution.

The second-order probability representation subsumes those of normal first-order probabilities and also those of interval probabilities.

## 1.10 Discussion: Why Not Third-Order Probability?

Some people ask, "If second-order probability is so useful, why not third-order or fourth-order probability? Where does it stop?". The answer to this question is that, although third- and fourth-order probabilities are mathematically well-defined, in current problems there is no need for them–there is no interpretation of the mathematics that yields useful information.

If we consider this question, we would have to ask, "How might it be useful to represent multiple possible universes of possible confidences and likelihoods?". In a single-agent system, there is no need to represent this. Only if we move to multiple-agent systems might this become useful. If we are trying to represent the beliefs that agent A has about how likely it is that agent B believes that outcome C is likely to happen with a certain confidence, and how confident agent A is in believing these likelihoods, then it would be useful to represent such problems using third- and fourth-order probabilities. But such situations are really too complex for the present analysis, and must wait for future research. Third- and fourth-order probability systems are not discussed here for the simple reason that they are not needed yet.

## 2 Introduction

Realistic planning systems must deal with actions having nondeterministic outcomes. Often conditions are highly uncertain, and a large number of previous action trials allowing confident derivation of a set of frequency-based outcome probabilities is not available. In addition, limited-resource systems must decide when they must act, and when they can afford to gather more information. A planning system needs a well-grounded numeric model that can represent the uncertainty of a system directly and present precise answers as to when and how to act.

This is provided by a system of theories and implementations, which is introduced in this paper. A situation-based model of *uncertain action* provides a representation formalism. A model of *second-order probabilities* represents uncertainties explicitly and distinguishes *confidence* and *likelihood* measures. Confidence measures are used to make *meta-decisions* concerning the gathering of additional information on uncertain likelihoods. A system of *partition-* and *event-based equations* supports the probability model. A maximum-entropy method *initializes* unknown second-order probabilities without using frequencies and a Bayesian method *updates* them from observations. *Value of information* equations provide quantitative solutions to meta-decision problems. The implemented system is used to solve a simple example which illustrates meta-decision making.

## 3 Previous Works

Recently, many researchers have investigated planning [Seg88, Han90a, HH90, KD89, GI89, Wel90b, DB90, BDKL91],[3] plan inference [Mye91], or meta-decision making [HCH89, RW89] using first-order probabilities in a decision-theoretic framework. However, these do not use second-order probabilities.

There has also been much activity concerning propagation of belief values in probabilistic networks [Pea88][Wel90a]; however, current designs cannot represent nonmonotonic nor repeated actions. Some researchers have investigated timelines, choice, and nondeterministic action representation [McD82, HM91, RG91a, RG91b, Sho89], or uncertain beliefs in time [Han87, KD89, DK88, DK91a]. The theory of probability has been investigated extensively by Halpern and Fagin [Hal89, Hal90, FH89, HF90b], Kyburg [Kyb91b, Kyb91a], Bacchus [Bac90b], and others. Good recent summaries on meta-decisions using the value of information can be found in [Dea91] and [RW91]. Spiegelhalter [Spe86], Heckerman and Jimison [HJ87], and later Pearl [Pea88, p.360-372] discuss confidence based on conditioning events, which is mathematically similar to discrete second-order probabilities, but requires known probabilities. Pearl defines a philosophy of second-order probability

[3]See however Feldman and Sproull [FS77] for an excellent early paper on decision-theoretic planning.

[Pea89][Pea88, p.358-359] different from ours, but then rejects it as unnecessary (ibid,p.372), ignoring meta-decisions. Spiegelhalter [Spe86], Fung and Chong [FC86], and Cheeseman [Che85] point out that confidence can be represented by a second-order (event) distribution and that confidence increases with more samples. Kyburg [Kyb89] explores second-order probabilities but rejects them because they are not needed for decisions, ignoring their role in meta-decisions and representing confidence. Raiffa shows how to derive doxastic second-order event distributions [Rai68, pp. 161-168] and uses these event distributions for meta-decisions, but does not use normative partition distributions.

No known previous researchers have used second-order probabilities for planning, have distinguished second-order partition probabilities from event probabilities, have presented a system of equations for partition probabilities, nor have used partition probabilities to quantify the value of obtaining perfect confidence as used in meta-decisions.

# 4 UNDA Theory

Planning and reasoning with Uncertain NonDeterministic Actions (UNDAs) requires a model of action representation. The model is based on situation theory [BP83]. Action types (plan schemata) are instantiated in situation instances. An uncertain action nondeterministically transitions to one of a known set of possible outcome situations, with a specified believed (second-order) probability. Performing a particular action type in a particular situation type grants license for using a particular probability (similar to [Pea88, p.13]).

## 4.1 The B-SURE system

The UNDA theory is implemented in a system called B-SURE (Believed Situation and UNDA Representation Environment) [Mye92]. The B-SURE system is based on an ATMS [dK86] that represents valued state, situation, uncertain transition, and action types and instances in multiple possible temporal action worlds with nonmonotonic assertions and deletions, in a manner extended from the theory presented by Morris and Nado [MN86]. The use of situation and state types allows reasoning to be performed ahead of time, and simultaneously over multiple worlds; the system is not forced to wait until states and situations are instantiated, and then reason separately in each possible world. B-SURE supports inferencing and planning.

## 4.2 History Mechanism

An interactive planning/execution system requires the use of a history mechanism to represent the current progress of the agent. This is provided by changing a transition assumption into a premise when it is known that an agent has started performing an action or an action has finished with a given outcome. An additional "past" flag is set on previous situations and actions to distinguish actions that are currently being executed (e.g., high-level actions) from those that have finished execution.

### 4.3 Planning Considerations

This paper's methods can be applied in most planning formalisms (cf. [THD90]). The current system uses a simple case-based expansion method [Ham89]. Significant issues include: (1) Plans form reactive *trees* of *contingency plans*; (2) Alternate *sets* of scored goals are searched for; (3) Planning is neither purely predictive [FN71] nor reactive [Sch87] but *interactive*; (4) Nondeterministic decompositions become problematical; (5) Practical systems require implementing *recognition demons* that can infer and certify when a given outcome has occurred [Mye91].

## 5 A Theory of Probability

There are two basic kinds of probability: *real*, and *believed* probability.

Assume nondeterminism actually exists. *Real probability* represents the actual likelihood that a particular outcome follows from the execution of a particular nondeterministic action in the world. A real probability associated with a possible outcome is an innate property of a situation/action pair. If the action type is repeatable and is repeated a large number of times in instances of the same situation type, the observed frequency of an outcome's occurrence will converge on the real probability.[4] It is not necessary for an action to be repeatable or even to be actually performed in order to have real probabilities associated with it. The real probability of an outcome is a unique constant. Executing an action repeatedly in the real world can be modeled by repeatedly drawing a colored ball from the *same* urn with replacement. Real probabilities correspond to Barwise and Perry's "real situations" [BP83, pp.49,57-60]. It does not make sense to talk of putting real probabilities into a computer.

A *believed probability* is a model of a real probability that is represented by an agent.[5] Believed probabilities correspond to Barwise and Perry's "abstract situations" [BP83]. They do not have to be unique nor constant.

Believed probabilities can be subdivided into *normative probabilities* (what an agent ideally *should* believe), and *doxastic probabilities* (what an agent actually *does* believe). [6] Normative probabilities are derived from a mathematical

[4]This paper only deals with stationary processes, defined as actions in which the real probabilities associated with outcomes are constants and do not change.

[5]The term "subjective probability" has been used to describe at least believed, normative, and doxastic probability, and hence these other terms are proposed.

[6]Normative probabilities are contrasted with *prescriptive* probabilities by Baron [Bar88]: normative probabilities may be too difficult or slow for an agent to use in practice. and should be replaced with fast heuristics.

model[7], and must be *valid* ($0 \leq p_i \leq 1$, $\Sigma_i p_i = 1$). Doxastic probabilities are personal opinions and are not constrained to be valid.

Believed probabilities may be represented by a universe of possible worlds and probabilistic transitions between them; *first-order* probabilities are defined over these transitions. Instances of the same situation/action pair should have the same transition probabilities in the same universe. Since the real universe is unknown, however, the agent should consider a continuum of possible universes. This continuum has a *second-order* probability distribution representing the believed chances that a given believed universe accurately represents the real universe.[8]

Slightly more formally, if the believed probability of an outcome is treated to be not a constant $p$ but rather a random variable p with a range [0,1], then a (believed) continuous probability distribution pr(p=$p$) can be associated with this variable [9] [Pap84, p.85]. This is called a *second-order probability distribution*, because it represents the believed probability that the real probability is a particular constant. For clarity, such functions will be represented by $q(p) \stackrel{def}{=}$ pr(p=$p$). This does *not* represent the probability of the event that the agent will come to believe p=$p$ [Pea88, p.359]; the agent already *simultaneously* believes p=$p_1$ and p=$p_2$ for a continuum of $p$'s.

A second-order probability can be modeled by assuming that nature has drawn an urn representing the universe from an infinite cave of urns, and is drawing colored balls from that same one urn repeatedly with replacement to represent (possible-world) outcomes of a repeated given action. The agent has a believed opinion about the distribution of urns in the cave; however, the agent does not know *which* urn was chosen by nature to be the *actual* urn.

In another conceptualization, state (determining a possible world) is a fundamental observable quantity analogous to position. First-order probability (likelihood) governs the changes between states (a universe of possible worlds and transitions), and, like velocity, can only be measured by differences. Second-order probability (confidence) governs the changes between probability models (which believed universe is most accurate) and is thus analogous to acceleration.[10]

Note that second-order probabilities subsume both certainly known first-order (constant) probabilities and also interval probabilities.

[7] It is not claimed that the Bayesian model presented later is the only normative model. Other models, for instance the Dempster-Shafer formalism [Sha79], are possible.

[8] The 1st/2nd-order distinction becomes important when representing *confidence* and when committing to decisions involving *repeated actions*.

[9] The theory supports discrete distributions (e.g., [Pea88, p.366-372]). However, the later Initialization section shows that these normally assume more known information than is justified and are hence usually misguided.

[10] As in physics, higher-order terms are well-defined but are useless in most practical matters. Third- and fourth-order probabilities will be required for solving plan inference problems involving agents' opinions of other agents' opinions of confidences and likelihoods.

# 6 Decisions vs. Meta-Decisions

The *decision* problem is the task of choosing between a known set of known but uncertain *alternatives* each comprising a tree of actions with a set of ultimate outcome situations, their values, and the likelihoods of reaching them. Given an evaluation strategy such as maximum expected value[11], it is straightforward to evaluate a decision problem and choose the most desired alternative [HM83]. Where, then, is the difficulty? The problem is that the current believed decision problem may not represent the best model of the real decision problem–and if more information is added, the model *may change* to a *new* decision problem. This can happen in the following ways: (1) a new alternative is added to the set; (2) an old alternative is deleted; (3) an action's expected outcome set is modified; (4) an alternative's anticipated value changes; (5) an alternative's expected likelihood changes. These modifications are products of *information* derived from *information-gathering actions* which include: (a) expanding a situation by mentally instantiating a new hypothetical action instance in the situation, thus making it more feasible; (b) performing other mental reasoning to increase confidence in values, likelihoods, and alternative and outcome sets; and (c) actually performing actions, including test actions. Deciding whether to decide and act with the current decision problem, or whether to attempt to change the current problem by performing an information-gathering action, is a *meta-decision problem.*[12] Meta-decision problems involving likelihood estimates (5) are properly answered using *confidence* measures. If the agent has high confidence in its believed estimates of likelihoods, with known values, alternative sets and outcome sets, then it believes that *learning more information will probably not change the decision problem in a significant manner.* With low confidence, *learning may significantly change the believed decision problem model*, and is therefore valuable. *Second-order probability distributions* represent confidence and should be used to solve meta-decision problems involving clarification of likelihood estimates.

# 7 A Calculus of Normative 2nd-Order Probabilities

## 7.1 Definition

A sample space $S$ is partitioned into $n$ known elemental outcome events $[A_1, \ldots, A_n]$ composing the partition $\mathcal{A}$. The corresponding probability vector

[11]Other significant strategies include minimum risk, maximum possible gain, maximum thrill, and maximum learning. Maximum utility subsumes these and is subtle. Therefore only the maximum value evaluation strategy is supported in this paper.

[12]Other meta-decision alternatives besides *observing/interacting* include *do nothing*, *wait* for a possible change, *waffle* by trying to take two or more alternatives, *consult* experts, *relegate* the problem to a human superior, *delegate* an inferior or sibling agent to make the decision, *react* randomly, *respond* in a habitual/reflexive manner, or *transcend* the problem.

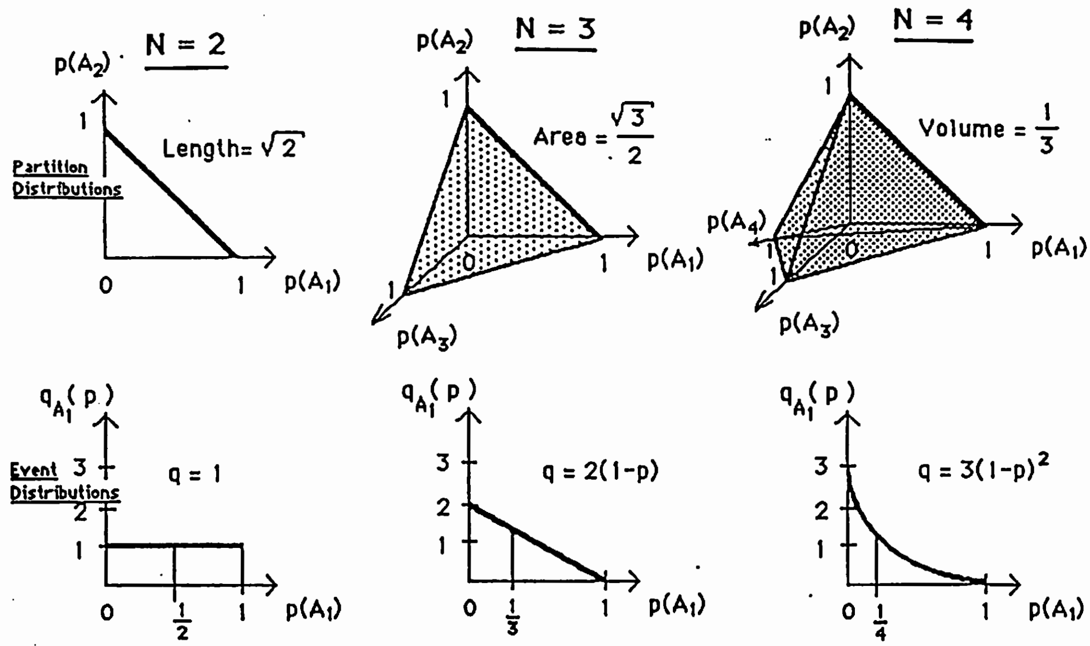


Figure 1: Initial 2nd-Order Partition and Event Probability Distributions.

$\vec{p} = (p_{A_1}, \ldots, p_{A_n})$, where $p_{A_i} \equiv p_i \equiv p(A_i)$, forms an n-dimensional space $\vec{\mathrm{p}}$. A point in this space $\vec{\mathrm{p}} = \vec{p}$ constitutes a unique assignment of (first-order) probabilities to the elemental outcome events. Normative probabilities have the property that the only valid probability assignments are contained in an $(n-1)$-dimensional hypertetrahedron defined by the formulas $0 \leq p_i \leq 1$, $\Sigma_i p_i = 1$. These are shown for the first three nondeterministic-outcome dimensions in figure 1. For instance, for $n = 3$, all valid assignments comprise a 2-dimensional equilateral triangle (plus its edges and verticies).[13] Doxastic probabilities are not constrained to be valid. Thus they may take on probability assignment points anywhere in the n-dimensional first-quadrant unit hypercube.

A second-order probability function of a partition $\mathrm{q}(\mathcal{A})$ is then defined as a function $\mathrm{q}_{\mathcal{A}}(\vec{p})$ defined over the vector space $\vec{\mathrm{p}}$ that maps a vector $\vec{p}$ into a scalar $q$. For normative probabilities, $q$ is always equal to 0 in invalid regions of the space. This function has the quality that the integral of q over the entire space is equal to 1. This second-order *partition probability distribution* function $\mathrm{q}_{\mathcal{A}}(\vec{p})$ is the fundamental unit of computation in this calculus.

[13]The (hyper-)edges have the interesting property that they represent situations where one of the outcomes has a probability of 0.0. The verticies represent situations where all but one of the outcomes has a probability of 0.0. It is important that the user know the number of possible outcomes at the beginning of the problem; the mathematics assumes that the event that one of the outcomes has a probability of 0.0 is possible but not more significantly likely than anything else. Problems where the number of possible outcomes is uncertain or unknown are not covered in this paper.

## 7.2 Projection

The second-order *event probability distribution* $q_{A_i}(p_i)$ for any one elemental outcome $A_i$ in the partition is determined by taking the projection of the $q(\vec{p})$ partition probability surface onto that outcome's p axis:

$$q_{A_i}(\mathrm{p}_i) = \int_0^1 \cdots \int_0^1 q_{\mathcal{A}}(p_1, \ldots, p_{i-1}, \mathrm{p}_i, p_{i+1} \ldots, p_n)\; dp_1 \ldots dp_{i-1}\, dp_{i+1} \ldots dp_n \qquad (2)$$

The integral of $q_{A_i}(p_i)$ over $p_i$ is also equal to 1. Note: Doxastic event distributions can be determined from experts [Rai68, p.161-168]. However, this paper points out that (1) normative, not doxastic, distributions should be used when possible; (2) partition, not event, distributions should be used when possible.

## 7.3 Equivalent Probability

The (first-order) *equivalent* or *expected probability* of an event $A_i$ is the weighted mean of its distribution [Pap84, p.85]:

$$p(A_i) = \int_0^1 q_{A_i}(p_i)\; p_i \;\; dp_i \qquad (3)$$

This probability can be used in making decisions.

## 7.4 Negation

The normative second-order event probability of the complement $\neg A_i$ of an event $A_i$ is formed by reversing the event distribution:

$$q_{\neg A_i}(p) = q_{A_i}(1-p) \qquad (4)$$

*Corollary:* The normative second-order event probability distributions of the outcomes of a two-outcome action are the reverse of each other.

## 7.5 Union

Taking the union $A_{i+j} = A_i \vee A_j$ of disjoint events $A_i$ and $A_j$ in $\mathcal{A}$ forms a new partition $\mathcal{A}'$ that has $(n-1)$ elements in it. This forms a new partition probability distribution $q_{new}(\mathcal{A}')$ found by collapsing the $p_i$ and $p_j$ axes into a single $p_{i\vee j}$ axis. The q mass density is integrated accordingly. Without loss of generality let $i = n-1$ and $j = n$. Then:

$$q_{new\mathcal{A}}(\ldots p_{A_{n-2}},\; p_{A_{i\vee j}}) = \int_0^{p_{A_{i\vee j}}} \sqrt{2}\, q_{\mathcal{A}}(\ldots p_{A_{n-2}},\; (s),\; (p_{A_{i\vee j}} - s))\; ds \qquad (5)$$

The union of more than two disjoint events can be found by collapsing the appropriate set of axes and integrating over the corresponding hypervolume.

The second-order event probability of the union can be determined directly from the old partition distribution function by collapsing the axes and projecting the rest of the space at the same time:

$$q_{A_i \vee_j}(p) \;=\; \int_0^1 \cdots \int_0^1 \int_0^{p_{A_i \vee_j}} q(p_{A_1},, \;\; (s = p_{A_i}),((p_{A_i \vee_j} - s) \;=\; p_{A_j}), \;\; ,p_{A_n}) \;\; ds dp_1 \, dp_{k \neq i,j} \, dp_n \tag{6}$$

In general, *there is not enough information to determine the second-order event distribution of the union from the event distributions of $A_i$ and $A_j$.* This is equivalent to the inverse-projection problem, and is the main reason why partition distributions should be used instead of event distributions.

## 7.6 Conditionals

A partition $\mathcal{B}$ with $m$ elemental events, that is conditioned without loss of generality on $A_n$ with a corresponding conditional partition distribution $q_{\mathcal{B}}(\mathcal{B}|A_n)$, forms a refinement of the partition $\mathcal{A}$ into a new partition $\mathcal{A}'$ with the $(n+m-1)$ elements $[A_1, ..., A_{n-1}, \; A_nB_1, ..., A_nB_m]$. The partition distribution function of $\mathcal{A}'$ is found by splitting the $p_{A_n}$ axis into $m$ new axes, and weighting the previous $q_{\mathcal{A}}$ mass by the $q_{\mathcal{B}}$ proportion:

$$q_{\mathcal{A}'}(p_{A_1}, ..., p_{A_{n-1}}, \; p_{A_nB_1}, ..., p_{A_nB_m}) \;=\; q_{\mathcal{A}}(p_{A_1}, ..., p_{A_n}) \, q_{\mathcal{B}}(p_{B_1}, ..., p_{B_m}) \tag{7}$$

The event probability distribution of a new event $A_iB_j$ is found by combining the event distributions of the component events $q_{A_i}(p_i)$ and $q_{B_j|A_i}(p_j)$:

$$q_{A_iB_j}(p_{ij}) \;=\; \int_{p_{ij}}^1 \frac{1}{|s|} \, q_{A_i}(s) \;\; q_{B_j|A_i}(\frac{p_{ij}}{s}) \;\; ds \tag{8}$$

with the usual adjustments if $p_{ij} = 0$.

The union operation and the conditional operation are duals of each other.

## 7.7 Expectation Distribution

Associate a value vector $\vec{V}$ of known constants $V_1, ..., V_n$ with the respective outcome events $A_1, ..., A_n$. Then the expected value of a given point $\vec{p}$ in the probability space is the dot product $E(\vec{p}) \;=\; (\vec{p} \cdot \vec{V})$, a function of $\vec{p}$. E is a random variable, and the confidence that this particular $E$ is correct is the partition distribution $q_{\mathcal{A}}(\vec{p})$. The probability distribution $q_{E(\mathcal{A})}(E)$, representing the likelihood that any particular scalar $E$ is the actual expectation of $\mathcal{A}$, is found by integrating over the probability points where $E$ holds:

$$q_{E(\mathcal{A})}(E) \;=\; \int_{\vec{p}} q_{\mathcal{A}}(\vec{p}) \;\; \Delta(E - \vec{p} \cdot \vec{V}) \, d\vec{p} \tag{9}$$

where $\Delta$ is the unit selector function $\Delta(0) = 1, \; \Delta(x \neq 0) = 0$. This is a second-order distribution because E is based on first-order probabilities. Expectation is not a constant, but is a variable based on which universe of possible-worlds

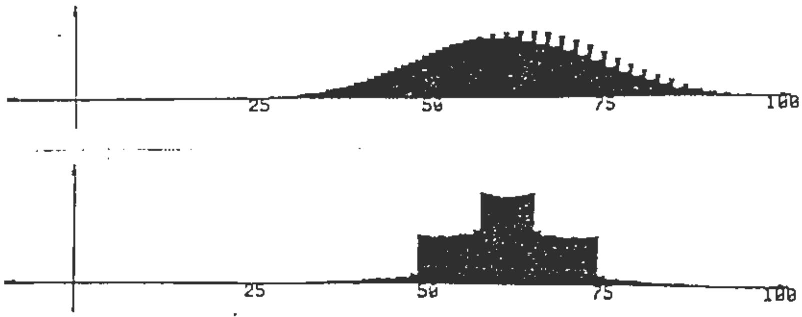


Figure 2: Actual (Partition) and Estimated (Event) Expectation Distributions for an Action with n=4

and probabilistic transitions is true. The scalar *equivalent expectation* $E(\mathcal{A})$ is of course the weighted mean of this distribution, $\int_{-\infty}^{\infty} E\ q_{E(\mathcal{A})}(E)\ dE$.

In general, *there is insufficient information to determine an accurate expectation distribution from the n event distributions* $q_{A_i}(p)$. However, a rough heuristic approximation can be found by taking the average of an approximation from each axis:

$$\hat{q}_{E(\mathcal{A})}(E) \;=\; \frac{1}{n} \sum_{i=1}^{n} q_{A_i}\left(\frac{E - V_{bi}}{V_i - V_{bi}}\right) \tag{10}$$

where $V_{bi} = \frac{1}{(n-1)} \sum_{j \neq i} V_j$ is the $i$'th base value found by averaging the others. Note that $q_{A_i}(p) = 0$ if $p < 0$ or $p > 1$. This approximation is exact if $n = 2$. See figure 2 for an example.

Comparision of (*decision* between) two action alternatives $\mathcal{A}$ and $\mathcal{B}$ under a maximum-expected-value strategy uses the equivalent expectations (assuming independence):

$$E(\mathcal{A}) \geq E(\mathcal{B}) \tag{11}$$

which is a condensation of $p(\,E(\vec{p}_{\mathcal{A}}) \geq E(\vec{p}_{\mathcal{B}})) \geq 0.5$. Using this equation and the conditional equations (7) or (8) allows decision trees to be evaluated via "averaging out and folding back" [Rai68].

## 7.8 Maximum and Greater-than-or-equal

The following equations are generally useful, and could be used for comparing decisions in cases involving strategies other than maximizing expected value. Given two independent random variables x and y with corresponding probability densities $p_x(x)$ and $p_y(y)$, the maximum distribution is:

$$p_{max(x,y)}(z) \;=\; \int_{-\infty}^{z} p_x(s)\,p_y(z) \;+\; p_x(z)\,p_y(s)\ ds \;-\; \int_{z^-}^{z} p_x(s)\,p_y(s)\ ds \tag{12}$$

(The last term is zero, and thus may be ignored, in systems having no impulses in their probability distributions.)

The scalar probability that $x \geq y$ is:

$$p_{x \geq y} = \int_{0}^{\infty} \int_{-\infty}^{\infty} p_x(z+s)\,p_y(s)\ ds\,dz \tag{13}$$

Of course, if Y is a constant, this reduces to

$$p_{x \geq Y} = \int_{Y}^{\infty} p_x(s)\, ds \tag{14}$$

## 7.9 Other Required Formulas

**Sum.**

$$p_{x+y}(z) = \int_{-\infty}^{\infty} p_x(z-s)\, p_y(s)\, ds \tag{15}$$

**Difference.**

$$p_{x-y}(z) = \int_{-\infty}^{\infty} p_x(z+s)\, p_y(s)\, ds \tag{16}$$

# 8 Origins of Normative 2nd-Order Probabilities

## 8.1 Initialization

Normative second-order probabilities are initialized using maximum entropy theory applied to all known information. If only $n$, the number of elemental outcomes, is known, and no other information,[14] then any one valid probability vector is equally likely, and the partition probability function is flat and is equal to the inverse of the valid surface hyperarea:[15]

$$q_{\mathcal{A}}(\vec{p}) = \frac{(n-1)!}{\sqrt{n}} \quad (\text{when } \Sigma p_i = 1,\ p_i \geq 0;\ \ q_{\mathcal{A}}(\vec{p}) = 0 \text{ otherwise}) \tag{17}$$

The event probability distribution is the same for each outcome and also is a function of $n$:

$$q_{A_i}(p_i) = (n-1)\,(1-p_i)^{n-2} \tag{18}$$

Examples are shown in figure 1. Most people are familiar with the flat distribution for $n = 2$. Note $q_{A_i}(0) = (n-1)$ and the expected $p(A_i) = \frac{1}{n}$.

## 8.2 Updating

If the action modeled by $\mathcal{A}$ is performed $m$ times and the outcomes $A_1, .., A_n$ are certainly observed to have occurred $k_1, .., k_n$ times respectively from sequential observation, $\sum k_i = m$, then the believed partition probability distribution $q_{old\mathcal{A}}()$ can be normatively updated to a more accurate model $q_{new\mathcal{A}}()$ as follows (derived from [Pap84, p.86]):

$$q_{new\mathcal{A}}(\vec{p}) = \frac{p_1^{(k_1)} ... p_n^{(k_n)}\ \ q_{old\mathcal{A}}(\vec{p})}{\left(\int_{\vec{p}} p_1^{(k_1)} ... p_n^{(k_n)}\ q_{old\mathcal{A}}(\vec{p})\, d\vec{p}\right)} \tag{19}$$

[14] If other information is known, appropriate initializations can be specified. However, the mathematical reasoning becomes rather complex, and is beyond the scope of this paper (see e.g. [Pap84, p.535-544]).

[15] The proof is available from the author.

Note that the denominator is a constant. If equation (17) is used for initialization, this expression is related to a beta density in $n$ dimensions. This type of updating method is central to so-called "Bayesian statistics". [Jus84]

A similar updating equation holds for event distributions using valid normative probabilities [Pap84, p.86]:

$$q_{newA_i}(p_i) = \frac{p_i^{(k_i)}(1-p_i)^{(m-k_i)} \; q_{oldA_i}(p_i)}{\left(\int_0^1 p_i^{(k_i)}(1-p_i)^{(m-k_i)} \; q_{oldA_i}(p_i)\, dp_i\right)} \tag{20}$$

Updating is significant because only the original distribution and the success counts need to be stored to compute the current distribution as required.

If equation (17) (or, equivalently, (18)) is used for initialization, then the results of (20) can be expressed in a closed form (using (18) and [Pap84, p.87]) as:

$$q_{newA_i}(p_i) = \frac{(m+n-1)!}{k_i!\,((m+n-2)-k_i)!}\, p_i^{(k_i)}(1-p_i)^{((m+n-2)-k_i)} \tag{21}$$

and the corresponding expected probability (derived from [Pap84, p.88]) is:

$$p(A_i) = \frac{k_i + 1}{m + n} \tag{22}$$

not $\frac{k_i}{m}$, as is customarily taught.

# 9 The Updating Function for Uncertainty Distributions

It is important to be careful with the distributions that are used as input to the updating function, as the function can only sharpen *uncertain* distributions. In particular, if a known doxastic first-order probability is represented as a unit spike, then the updating function will not change that probability–it is already known with certainty, and it can't change! Similarly, if an interval probability is represented using a $q$ distribution as having a flat or exponential distribution between the max and the min, and 0 elsewhere, then the updating function will only change the distribution inside the interval–all of the probabilities outside of the interval are certainly known not to be possible! In general, this will probably not be the desired behavior. It can be fixed by ensuring that the $q(p)$ distribution is not 1 nor 0 for any $p$ unless that $p$ is known certainly to take on or not take on that probability. This is ensured by the normative initialization function.

## 9.1 Value of Perfect Information

The general value of information is the sum of the possible gains times the probabilities that those gains occur [How66, Dea91]. Given that the agent is maximizing expected value, i.e. using eq. (11) to rank preferences, and that the effective expected value of the current best alternative is $E(a_{max})$, then the Expected Value of the Perfect Information that reveals the actual first-order probability $\vec{p}^*$ of alternative action $a_{alt}$ is:

$$EVPI(\vec{p}^*_{a_{alt}}) = \int_{E(a_{max})}^{\infty} (gain(\text{info}))\ q(E_{a_{alt}})\, dE_{a_{alt}} - cost(\text{info}) \qquad (23)$$

where $(gain())$ is normally $(E_{a_{alt}} - E(a_{max}))$ (although in some situations it is a constant, if being correct is all that counts), and $q(E)$ is taken from eq. (9) for partition distributions or estimated from eq. (10) for sets of event distributions.

## 9.2 Note

This equation concerns reducing uncertain confidence to perfect confidence and is one of the main results of this paper. In effect, it computes and uses $q(E_{a_{alt}} \geq E(a_{max}))$, the possibility that the expectation of the alternative *could be* greater than the current expectation, which is a second-order probability result. Note that if first-order probabilities are used, $E(a_{alt})$ is a constant, not a random variable, and $E(a_{alt}) \geq E(a_{max})$ is either true or not. Since the event distributions are coupled, it is insufficient to examine the events separately.

## 9.3 Value of Testing

The general expected value of information given one sample execution of an action is

$$EVI(sample) = \sum_{i=1}^{n} gain(q_{new\mathcal{A}}(\vec{p}) \mid A_i\text{observed})\ \ p(A_i) \qquad (24)$$

using equations (3), and (19) with $m$=1. The degree of change due to updating depends on the *confidence* of the current distribution.[16] This expression simplifies in cases.

## 9.4 Implementation note

Event distributions can only give approximately correct results. However, they are straightforward to implement, and allow a system to take advantage of the theory. Partition distributions are definitely preferrable if possible. Full use of partition distributions seems to require a symbolic integration package. The implemented system currently uses event distributions quantized at 0.05 intervals in $p$. Obviously quantization causes loss of precision [Wel90a]; this level seems acceptable for current applications.

[16]This equation apparently offers a closed-form optimal solution to the "two-armed bandit" problem [Har91] for $k$ arms with $n_k$ known outcomes apiece.

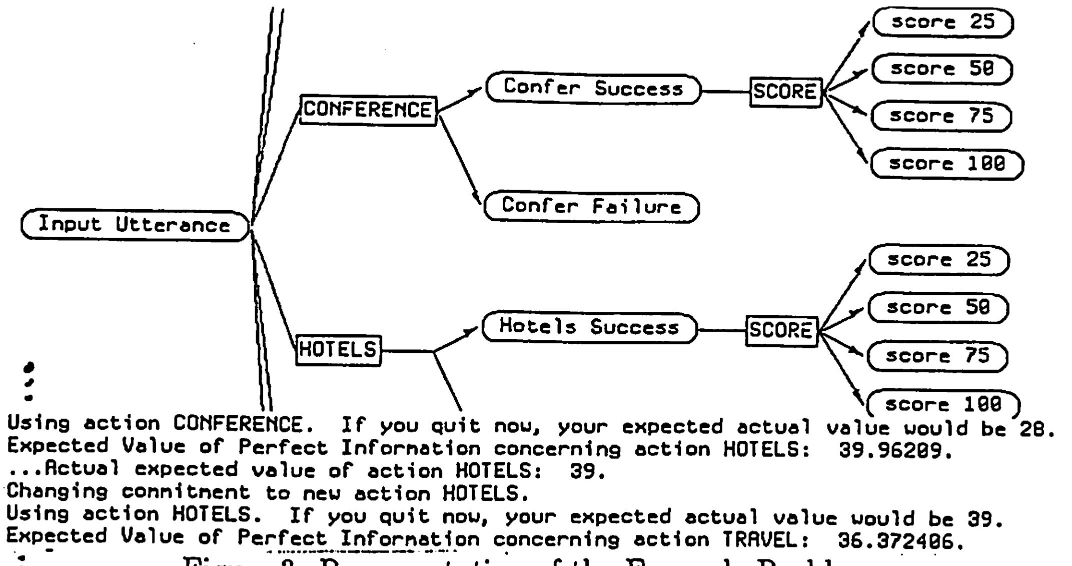


Figure 3: Representation of the Example Problem

# 10 Example

A hypothetical language translation system uses one of several modules to process input utterances, depending upon the microcontext of the conversation. For instance, modules might be specialized for processing parts of conversations about conference registration, hotel registration, travel, sightseeing, etc. A module will uncertainly succeed or fail in processing each utterance. If it succeeds, it will uncertainly produce an output scored at 25, 50, 75, or 100 points. Some utterances, such as "How much does it cost?", may be successfully processed by some or all modules, but only one module will be correct in its understandings and predictions. The system is time-limited, and can only choose one module which will then be used to process repeated utterances.[17]

The problem is modeled, using the B-SURE system, by an uncertain action representing success/failure for each module, and a different uncertain action representing the outcome scores given that the module was successful (See figure 3). Two separate actions are used because the success/failure outcome is a different sort than the score outcome, and to demonstrate conditional actions. Although partition probabilities should be used, the current system only supports event probabilities, and so the event equations will be referenced. It is assumed that the possible current conversational context is completely uncertain, and so equation (18) is used to initialize the second-order event probability distributions for the actions, with $n = 2$ and $n = 4$ respectively. Eq. (8) is used to evaluate the second action and "fold it back" through the first action; each resulting alternative has five event axes, corresponding to "failure" and "($value$)|success".

Eq. (10),(11) are used to rank the alternatives, and the highest one is selected to be the current "best action". Choosing is performed based on maximizing expected value. At this point, the actions are highly uncertain,

[17]This work is designed to attack high-uncertainty problems that have low-confidence probability estimates. It is assumed that, because of the complexities of spontaneous dialog, the system is processing utterances in an original micro-context with uncertain likelihood estimates. High-confidence problems may employ simpler strategies than those presented.

and they all have the same equivalent expectation, so one is chosen at random.

In this example, "perfect information" consists of knowing with complete confidence the real (first-order) likelihood of each module, and thus the corresponding expected value [Rai68, p.28,168]. Assume that there is an estimator function for each module that, for a known cost in time and resources (e.g. 0.5 sec.), can analyze the current dialog context and predict the likelihoods that this module will be successful and will obtain particular answer scores. Assume that these functions are ideal, i.e. they return perfectly correct first-order predictions (complete confidence). Assume that the system is willing to spend about two seconds on finding the best module, i.e. there is a $[time, V]$ curve representing the value of stopping immediately that connects the points [0,0], [1,0], [3,100], $[\infty, 100]$. Then the following strategy can be followed: (a) Maintain the current "best action" and its equivalent expectation; (b) Select the best action out of the others that have not been estimated yet; (c) Evaluate the EVPI of executing this action's estimator, using eq. (23); (d) Compare this against the current value of stopping immediately: (e) If greater, then update the "best action"; (f) Else stop and use the current "best action" to process the next series of utterances. Note that the amount of investigation is a function of the amount of free time available.

In this experiment, the average best expectation is 67, and the average number of modules analyzed is 4.2 (2.1 sec.). This compares against an expectation of 31 if only the *a priori* best module is chosen blindly.

# 11 Discussion

Some people have argued against the use of probability in some situations when an agent is ignorant and is unable to formulate a specific (doxastic) probability. The argument goes that assignment of even a second-order probability implicitly assigns an equivalent first-order probability to the events, and these authors feel uncomfortable with this. In other words, they *do not feel confident* of the resulting probability.

However, our theory answers this concern. Even if the agent has *no* doxastic probability at all (which he or she is free to do), the situation's *real probability* does exist in all cases–and this demands that a *normative* probability *should* exist. It does not matter whether the agent "believes in" the normative probability or not; the point is that it offers the best *model* of the *real* situation, given all the available information. Second-order probabilities offer an explicit representation of confidence, and a method of determining *when* to decide. Some people would have the agent not decide at all when faced with completely uncertain probabilities–however, this may not be an option in a resource-limited world. "Not to decide" is in fact a meta-decision, and it may be the *wrong* meta-decision if constantly applied in an indiscriminate manner. Our theory allows the agent to meta-decide whether to decide or to prefer gathering more information if possible, based on the confidence in its models. Sometimes it is not possible to wait, and in this case a decision must be made with information in which the agent has low confidence–but at least this is represented explicitly.

Fagin and Halpern [FH89] also argue that some events have probabilities that are nonmeasurable. This is a difficult philosophical point, and one that requires further research.

Another concern that people have pointed out is that a second-order probability distribution based on maximum entropy does not capture all of the prior information that a human expert has, based on experience and "common-sense" logic. There are two responses to this point. The first is to note that the method of maximum entropy is *not* restricted to simply equation 17 which is used for completely unknown situations, but can be used to include any constraints known to the agent that can be expressed mathematically. The mathematics gets rather complex, however, and is beyond the scope of this paper (see e.g. [Pap84, p.535-544]). The second response is to observe that this problem may indeed occur when attempting to accurately install the existing knowledge of a human into a computer expert system. However, when creating an artificial intelligence that must plan, function in the world, and learn from its experiences, a completely-unknown probability distribution may *accurately* describe the state of its knowledge. Human children and even intelligent adults make mistakes in judgement based on lack of experience, especially when encountering new situations; to be fair, an artificial intelligence must be given a similar range of experiences if its judgements are expected to be comparable.

Distributions are easily implemented in neural nets [Alb81]. It would be interesting to find out whether people actually use second-order probability distributions to represent likelihood and confidence.

The current algorithms assume that path evaluation time is negligible when compared with the time required to perform an action. If this is not the case, the algorithms should take into account the amount of time required to do inferencing and evaluation, as well as the domain action time, when computing the cutoff for limited-resource reasoning.

When many instances of the same action type are incorporated in a plan, dynamically updating the action-type's outcome probability distributions from observed execution outcomes will change likelihoods, which may change decision preferences. The system supports such dynamic updating. However, in theory it is possible to predict the changes in believed likelihood based on possible outcomes, and to use these a priori in different downstream timelines. The system does not yet support such calculations.

This work is based on actions having stationary real probabilities. The questions of how to deal with actions having nonstationary probabilities or probabilities that vary depending on the current situation, how to estimate the probabilities of such actions when proposed in a new situation type, and how to recognize and distinguish such actions and situations, basically comprises the learning problem and must be left for future research.

This work assumes that the starting situation is known with complete confidence. If the starting situation is known uncertainly, it is necessary to use game theory against a fair opponent (nature) to decide which actions to perform. Discussion of such theory is beyond the scope of this paper.

## 12 Conclusion

This paper has presented a brief introduction into the theory and practice of planning, decision-making, and meta-decision-making using second-order probabilities to explicitly represent uncertainty. Decisions can be made when necessary, with little or no frequency information, but the confidence may be low; information-gathering actions are indicated when required. Second-order probabilities offer a method to determine when it is useful to gather more information, and when it is time to make a decision.

# A Proof of Some Equations

This section presents a proof of equations (21) and (22).

Equation (20) is repeated here for convenience.

$$q_{newA_i}(p_i) = \frac{p_i^{(k_i)}(1-p_i)^{(m-k_i)} \quad q_{oldA_i}(p_i)}{\left(\int_0^1 p_i^{(k_i)}(1-p_i)^{(m-k_i)} \, q_{oldA_i}(p_i)\, dp_i\right)} \tag{25}$$

However, if (18) is used for initialization, then

$$q_{old\,A_i\,|\,m=0}(p_i) = (n-1)\,(1-p_i)^{n-2} \tag{26}$$

so

$$\begin{aligned} q_{newA_i|m,k_i}(p_i) &= \frac{p_i^{(k_i)}(1-p_i)^{(m-k_i)} \quad (n-1)\,(1-p_i)^{n-2}}{\left(\int_0^1 p_i^{(k_i)}(1-p_i)^{(m-k_i)}\,(n-1)\,(1-p_i)^{n-2}\,dp_i\right)} \\ &= \frac{p_i^{(k_i)}(1-p_i)^{(m+n-2-k_i)}}{\left(\int_0^1 p_i^{(k_i)}(1-p_i)^{(m+n-2-k_i)} \quad dp_i\right)} \end{aligned}$$

Using the identity

$$\int_0^1 p^A(1-p)^{B-A}dp = \frac{A!\,(B-A)!}{(B+1)!} \tag{27}$$

with $A = k_i$ and $B = m+n-2$, the denominator is evaluated and inverted. This resolves into equation (21):

$$q_{newA_i\,|\,m,k_i}(p_i) = \frac{(m+n-1)!}{k_i!\;((m+n-2)-k_i)!}\,p_i^{(k_i)}(1-p_i)^{((m+n-2)-k_i)} \quad \boxed{\text{q.e.d}} \tag{28}$$

Repeating equation (3) for convenience, the equivalent first-order probability of a second-order event distribution is:

$$p(A_i) = \int_0^1 q_{A_i}(p_i)\;p_i\;\;dp_i \tag{29}$$

Using equation (28) for $q(p)$, this is:

$$\begin{aligned} p(A_i) &= \int_0^1 \frac{(m+n-1)!}{k_i!\;((m+n-2)-k_i)!}\,p_i^{(k_i)}(1-p_i)^{((m+n-2)-k_i)}\,p_i\;\;dp_i \\ &= \frac{(m+n-1)!}{k_i!\;((m+n-2)-k_i)!} \quad \int_0^1 p_i^{(k_i+1)}(1-p_i)^{((m+n-1)-(k_i+1))}\;dp_i \end{aligned}$$

Again using identity (27) with $A = (k_i+1)$ and $B = (m+n-1)$ to convert the integral, we get:

$$p(A_i) = \frac{(m+n-1)!}{k_i!\;((m+n-2)-k_i)!} \quad \frac{(k_i+1)!\;(m+n-1-(k+1))!}{(m+n)}$$

$$= \frac{(m+n-1)!}{k_i!} \cdot \frac{(k_i+1)!}{(m+n)}$$

$$= \frac{k_i+1}{(m+n)} \quad \boxed{\text{q.e.d}}$$

# B Experimental Verification of the Estimator

## B.1 Discussion of the Experiment

The experiment is based on the theory of real and normative believed probabilities outlined in sections 1.4 and 5. The experiment consists of an observer attempting to estimate the real probability of an outcome of a stationary random process. For each experiment, first the number of possible outcomes $n$ is chosen. It is assumed that the observer knows $n$ correctly. Next, out of all the possible assignments of probabilities to the outcomes a probability assignment set $p(A_1), ...p(A_n)$ is chosen at random, such that $\Sigma_1^n A_i = 1$. This represents the unknown real probability of the problem. The observer arbitrarily picks the $i$'th outcome of the process, $A_i$, for observation. Without loss of generality, let $i = 1$. The goal is to estimate the probability $p(A_i)$ as closely as possible, by performing repeated trials of the process and observing the results, where $m$ is the number of trials that have been performed so far, and $k_i$ is the number of the successes of outcome $A_i$ observed so far. It is assumed that the identity of the process outcome is crisp, and is consistently and correctly observed with certainty.

The verification uses the two estimators $\widehat{p1_i} = \frac{k_i}{m}$ and $\widehat{p2_i} = \frac{(k_i+1)}{(m+n)}$. The trial sequence for each experiment starts at $m = 0$ and continues sequentially up to $m = 100$ trials using the same real probability assignments. Since $\widehat{p1_i}$ is undefined at $m = 0$, it is arbitrarily specified as 0 in this case.

The interval between 0 and 1 is divided up by sequentially assigning subintervals to outcomes $A_i$, where the length of each subinterval corresponds to the size of its probability $p(A_i)$. For each trial, a (flat-distribution) random number $R$ between 0 and 1 is picked to determine the outcome of the process. Whichever subinterval the number falls in designates the outcome of this trial. Since the observer is only interested in the first outcome, this can be reduced to signaling a success when the random number is $0 \leq R < p(A_i)$, and a failure otherwise.

For each trial, the distances $|\widehat{p1_i} - p_i|$ and $|\widehat{p2_i} - p_i|$ are collected as a function of $m$.

The experiment is performed 2000 times, and the results for all the experiments are collected as a function of $m$ and then divided by 2000 to get the average value. Finally, the average curve for $\widehat{p2_i}$ is subtracted from that of $\widehat{p1_i}$ to determine how much closer $\widehat{p2_i}$ is to the real probability, as a function of $m$.

The results are shown in figure 4 for $n$ equal to two, three, ten, and one hundred.

Comments There are two methods for selecting probability assignments randomly for $n$ outcomes. The first involves picking $(n-1)$ random numbers from a flat distribution between 0 and 1, to represent the $n-1$ degrees of freedom of the problem. These numbers are discarded if they sum to more than one. Otherwise, the last probability is determined by (1-sum). This method has the disadvantage that huge numbers of probability assignments must be discarded as invalid before a valid combination can be found, especially for higher $n$. The second method depends on using equation (18) to determine the expected distribution of an outcome's probability from a randomly-chosen probability assignment, where all possible assignments are equally likely. In this case, an appropriate outcome probability can be picked directly by transforming a flat-distribution random number $R$ between 0 and 1:

$$p_1 = 1 - e^{\frac{\ln(1-R)}{n-1}} \tag{30}$$

This is the method used for the experiment. The two methods are equivalent.

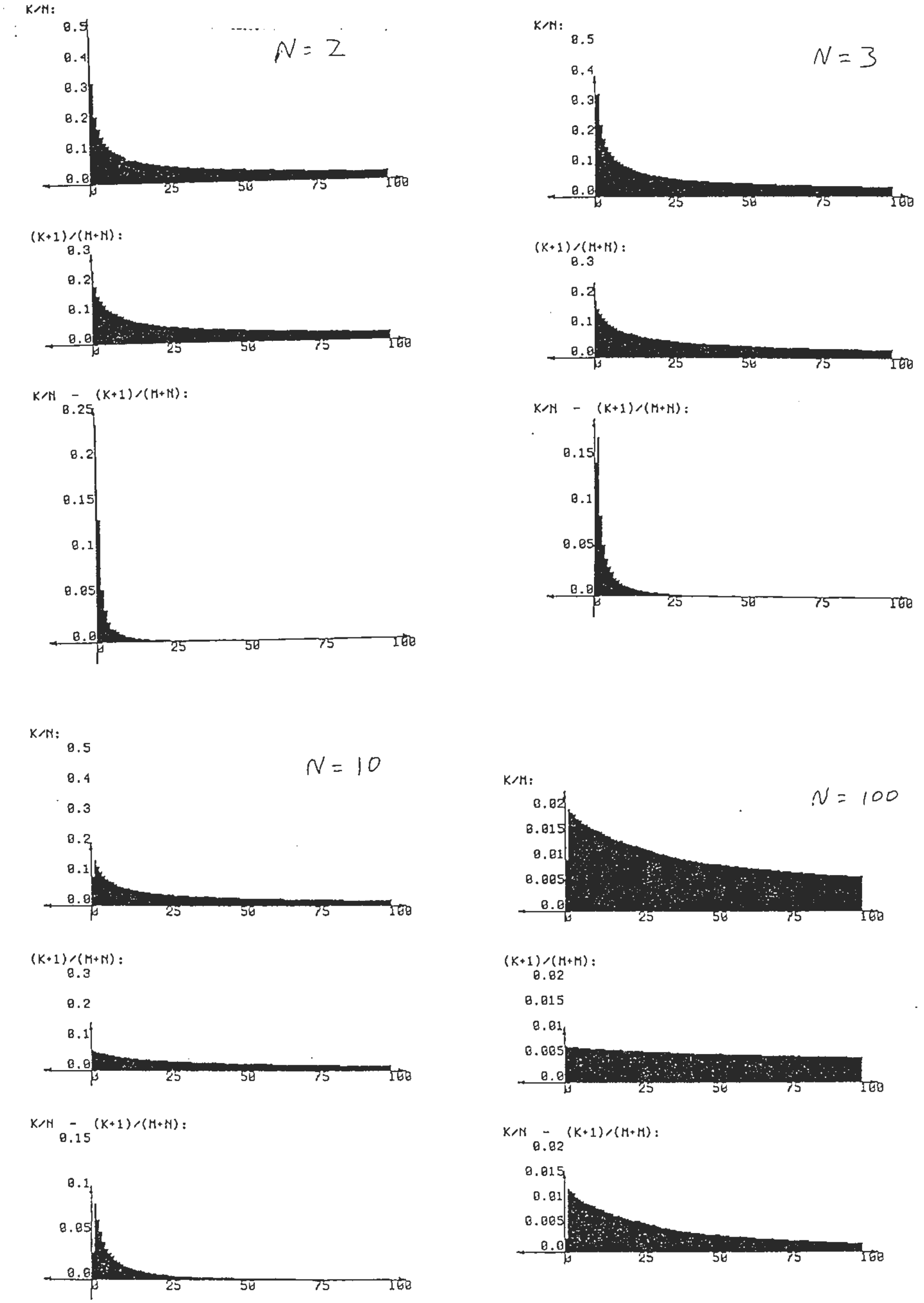


Figure 4: Mean Distance from the Real Probability and Improvement, for N=2, 3, 5, and 100

# References

[AC87] Philip E. Agre and David Chapman. Pengi: An implementation of a theory of activity. In *AAAI'87: The Sixth National Conference on Artificial Intelligence*, pages 268–272, Seattle, WA, 1987.

[AC90] Philip E. Agre and David Chapman. What are plans for? In Pattie Maes, editor, *Designing Autonomous Agents*, pages 17–34. The MIT Press, Cambridge, MA, 1990.

[AK83] James F. Allen and Johannes A. Koomen. Planning using a temporal world model. In *IJCAI'83: The Eighth International Joint Conference on Artificial Intelligence*, pages 741–747, Karlsruhe, West Germany, 1983.

[Alb81] James S. Albus. *Brains, Behavior, & Robotics*. BYTE Books (McGraw-Hill), Peterborough, NH., 1981.

[App80] Douglas E. Appelt. A planner for reasoning about knowledge and action. In *First Annual National Conference on Artificial Intelligence*, 1980.

[Bac90a] Fahiem Bacchus. Lp–a logic for statistical information. In *Uncertainty in Artificial Intelligence 5*, pages 3–39. North-Holland, Amsterdam, The Netherlands, 1990.

[Bac90b] Fahiem Bacchus. *Representing and Reasoning with Probabilistic Knowledge*. The MIT Press, Cambridge, MA, 1990.

[Bar88] Jonathan Baron. *Thinking and Deciding*. Cambridge University Press, Cambridge, Mass., 1988.

[BD89] Mark Boddy and Thomas Dean. Solving time-dependent planning problems. In *IJCAI'89: Eleventh International Joint Conference on Artificial Intelligence*, pages 979–984, Detroit, MI, 1989.

[BDKL91] Kenneth Basye, Thomas Dean, Jak Kirman, and Moises Lejter. Decision-theoretic architectures for planing, perception and control. Technical report, Brown University, Providence RI, 1991.

[BGD87] Piero P. Bonissone, Steven S. Gans, and Keith S. Decker. Rum: A layered architecture for reasoning with uncertainty. In *IJCAI'87: The Tenth International Joint Conference on Artificial Intelligence*, pages 891–898, Milan, 1987.

[BP83] Jon Barwise and John Perry. *Situations and Attitudes*. The MIT Press, Cambridge, Mass., 1983.

[CG89] Eugene Charniak and Robert Goldman. A semantics for probabilistic quantifier-free first-order languages, with particular application to story understanding. In *IJCAI'89: Eleventh International Joint Conference on Artificial Intelligence*, pages 1074–1079, Detroit, MI, 1989.

[Che85] Peter Cheeseman. In defense of probability. In *IJCAI'85: The Ninth International Joint Conference on Artificial Intelligence*, pages 1002–1009, Los Angeles, CA, 1985.

[DB88] Thomas Dean and Mark Boddy. An analysis of time dependent planning. In *AAAI'88: The Seventh National Conference on Artificial Intelligence*, St. Paul, MN, 1988.

[DB90] Mark Drummond and John Bresina. Anytime synthetic projection: Maximizing the probability of goal satisfaction. In *AAAI'90: The Eighth National Conference on Artificial Intelligence*, pages 138–144, Boston, MA, 1990. The MIT Press.

[Dea91] Thomas Dean. Decision-theoretic control of inference for time-critical applications. *International Journal of Intelligent Systems*, 6:417–441, 1991.

[DF89] Bruce D'Ambrosio and Michael Fehling. Resource-bounded agents in an uncertain world. In *AI and Limited Rationality 1989 Spring Symposium*, 1989.

[dK86] Johan de Kleer. An assumption-based tms. *Artificial Intelligence*, 28(2):127–162, March 1986.

[DK88] Thomas Dean and Keiji Kanazawa. Probabilistic temporal reasoning. In *AAAI'88: The Seventh National Conference on Artificial Intelligence*, pages 524–528, St. Paul, MN, 1988.

[DK91a] Thomas Dean and Keiji Kanazawa. A model for reasoning about persistence and causation. Technical report, Brown University, Providence RI, 1991.

[DK91b] Thomas Dean and Jak Kirman. Representation issues in bayesian decision theory for planning and active perception. Technical report, Brown University, Providence RI, 1991.

[FC86] Robert M. Fung and Chee Yee Chong. Metaprobability and dempster-shafer in evidential reasoning. In *Uncertainty in Artificial Intelligence (1)*, pages 295–302. North-Holland, Amsterdam, The Netherlands, 1986.

[FH89] Ronald Fagin and Joseph Y. Halpern. Uncertainty, belief, and probability. In *IJCAI'89: Eleventh International Joint Conference on Artificial Intelligence*, pages 1161–1167, Detroit, MI, 1989.

[FK86] Barry R. Fox and Karl G. Kempf. Planning, scheduling, and uncertanty in the sequence of future events. In *Uncertainty in Artificial Intelligence Workshop*, pages 77–81, Philadelphia, PA, 1986.

[FN71] Richard E. Fikes and Nils J. Nilsson. Strips: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, 2:189–208, 1971.

[Fre71] John E. Freund. *Mathematical Statistics*. Prentice-Hall Inc., Englewood Cliffs, NJ, second edition, 1971.

[FS77] Jerome A. Feldman and Robert F. Sproull. Decision theory and artificial intelligence ii: The hungry monkey. *Cognitive Science*, 1:158–192, 1977. Reprinted in "Readings in Planning", pp. 207-224.

[Gai86] Haim Gaifman. A theory of higher order probabilities. In Joseph Y. Halpern, editor, *Reasoning About Knowledge*, pages 275–292. Morgan Kaufmann Publishers, Inc., Los Altos, CA, 1986.

[GI89] Michael P. Georgeff and Francois F. Ingrand. Decision-making in an embedded reasoning system. In *IJCAI'89: Eleventh International Joint Conference on Artificial Intelligence*, pages 972–978, Detroit, MI, 1989.

[Hal89] Joseph Y. Halpern. An analysis of first-order logics of probability. In *IJCAI'89: Eleventh International Joint Conference on Artificial Intelligence*, pages 1375–1381, Detroit, MI, 1989.

[Hal90] Joseph Y. Halpern. An analysis of first-order logics of probability. *Artificial Intelligence*, 46:311–350, 1990.

[Ham89] Kristian J. Hammond. *Case-Based Planning*. Academic Press, Orlando, FL, 1989.

[Han87] Steve Hanks. Temporal reasoning about uncertain worlds. In *Third Workshop on Uncertainty in Artificial Intelligence*, pages 114–122, Seattle, WA, 1987.

[Han90a] Steve Hanks. Controlling inference in planning systems: Who, what, when, why, and how. Technical Report 90-04-01, Dept. of CS and Eng., University of Washington, Seattle, WA, 1990.

[Han90b] Steve Hanks. Practical temporal projection. In *AAAI'90: The Eighth National Conference on Artificial Intelligence*, pages 158–163, Boston, MA, 1990.

[Han90c] Steven John Hanks. *Projecting Plans for Uncertain Worlds.* PhD thesis, Yale University Department of Computer Science, January 1990.

[Har91] Leo B. Hartman. Uncertainty and the cost of planning. Technical Report 372, University of Rochester, Rochester NY, 1991.

[HCH89] Eric J. Horvitz, Gregory F. Cooper, and David E. Heckerman. Reflection and action under scarce resources: Theoretical principles and empirical study. In *IJCAI'89: Eleventh International Joint Conference on Artificial Intelligence*, pages 1121–1127, Detroit, MI, 1989.

[HF90a] Peter Haddawy and Alan M. Frisch. Modal logics of higher-order probability. In *Uncertainty in Artificial Intelligence 4*, pages 79–92. North-Holland, Amsterdam, The Netherlands, 1990.

[HF90b] Joseph Y. Halpern and Ronald Fagin. Two views of belief: Belief as generalized probability and belief as evidence. In *AAAI'90: The Eighth National Conference on Artificial Intelligence*, pages 112–119, Boston, MA, 1990.

[HF90c] Steve Hanks and R. James Firby. Issues and architectures for planning and execution. Technical report, Dept. of CS and Eng., University of Washington, Seattle, WA, 1990.

[HH90] Peter Haddawy and Steve Hanks. Issues in decision-theoretic planning: Symbolic goals and numeric utilities. In *DARPA Planning Workshop*, 1990.

[HJ87] David Heckerman and Holly Jimison. A perspective on confidence and its use in focusing attention during knowledge acquisition. In *Third Workshop on Uncertainty in Artificial Intelligence*, pages 123–131, Seattle, WA, 1987.

[HM83] Ronald A. Howard and James E. Matheson, editors. *The Principles and Applications of Decision Analysis.* Strategic Decisions Group, Menlo Park, CA, 1983.

[HM91] Steve Hanks and Drew McDermott. Modeling a dynamic and uncertain world 1: Symbolic and probabilistic reasoning about change. Technical report, Dept. of CS and Eng., University of Washington, Seattle, WA, 1991. Submitted to Artificial Intelligence journal.

[How66] Ronald A. Howard. Information value theory. *IEEE Trans. on Systems Science and Cybernetics*, SSC-2(1), 1966. reprinted in "The Principles and Applications of Decision Analysis", pp.779-783.

[jH90] Jane Yung jen Hsu. Partial planning with incomplete information. In James Hendler et al., editor, *Working Notes of Planning in Uncertain, Unpredictable or Changing Environments, AAAI Spring Symposium Series*, pages 62–66, Stanford University, March 1990. Also PRICAI '90, pp. 474-479.

[Jus84] James Justice, editor. *Maximum entropy and Bayesian Methods in Applied Statistics.* Cambridge University Press, Cambridge, Mass., 1984.

[KD89] Keiji Kanazawa and Thomas Dean. A model for projection and action. In *IJCAI'89: Eleventh International Joint Conference on Artificial Intelligence*, pages 985–990, Detroit, MI, 1989.

[Kyb89] H. E. Kyburg. Higher order probabilities. In *Uncertainty in Artificial Intelligence 3*, pages 15–22. North-Holland, Amsterdam, The Netherlands, 1989.

[Kyb91a] Henry E. Kyburg. Evidential probability. In *IJCAI'91: 12th International Joint Conference on Artificial Intelligence*, pages 1196–1202, Sydney, Australia, 1991.

[Kyb91b] Henry E. Kyburg. Knowledge representation and uncertainty in artificial intelligence. Technical Report 395, University of Rochester, Rochester NY, 1991.

[MA90] Nathaniel G. Martin and James F. Allen. Abstraction in planning: A probabilistic approach. Technical report, University of Rochester, Rochester NY, 1990.

[McD82] Drew McDermott. A temporal logic for reasoning about processes and plans. *Cognitive Science*, 6(2):101–155, April-June 1982.

[MFD85] David Miller, R. James Firby, and Thomas Dean. Deadlines, travel time, and robot problem solving. In *IJCAI'85: The Ninth International Joint Conference on Artificial Intelligence*, pages 475–479, Los Angeles, CA, 1985.

[MN86] Paul H. Morris and Robert A. Nado. Representing actions with an assumption-based truth maintenance system. In *AAAI'86: The Fifth National Conference on Artificial Intelligence*, Philadelphia, PA, 1986.

[Mye91] John K. Myers. Plan inference with probabilistic-outcome actions. In *Conf. Proc. Information Processing Society of Japan*, volume 3, pages 168–169, Tokyo, Japan, March 1991.

[Mye92] John K. Myers. The b-sure manual. Technical report, ATR Interpreting Telephony Research Laboratories, Kyoto, Japan, 1992.

[Pap84] Athanasios Papoulis. *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill Book Company, New York, NY, 1984. First edition 1965.

[Pea88] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufmann Publishers, Inc., Los Altos, CA, 1988.

[Pea89] Judea Pearl. Do we need higher-order probabilities and, if so, what do they mean? In *Uncertainty in Artificial Intelligence 3*. North-Holland, Amsterdam, The Netherlands, 1989.

[Rac89] Howard Rachlin. *Judgment, Decision, and Choice.* W. H. Freeman and Company, New York, NY, 1989.

[Rai68] Howard Raiffa. *Decision Analysis.* Addison-Wesley Publishing Co., Menlo Park, CA, 1968.

[RG91a] Anad S. Rao and Michael P. Georgeff. Asymmetry thesis and side-effect problems in linear-time and branching-time intention logics. In *Proceedings of IJCAI-91*, Sydney, Australia, 1991.

[RG91b] Anad S. Rao and Michael P. Georgeff. Deliberation and intentions. Technical Report 10, The Australian Artificial Intelligence Institute, 1991.

[RW89] Stuart Russell and Eric Wefald. On optimal game-tree search using rational meta-reasoning. In *IJCAI'89: Eleventh International Joint Conference on Artificial Intelligence*, pages 334–340, Detroit, MI, 1989.

[RW91] Stuart Russell and Eric Wefald. Principles of metareasoning. *Artificial Intelligence*, 49:361–395, 1991.

[SC87] Matthew Self and Peter Cheeseman. Bayesian prediction for artificial intelligence. In *Third Workshop on Uncertainty in Artificial Intelligence*, pages 61–69, Seattle, WA, 1987.

[Sch87] Marcel Schoppers. Universal plans for reactive robots in unpredictable domains. In *IJCAI'87: The Tenth International Joint Conference on Artificial Intelligence*, Milan, 1987.

[Seg88] Alberto Maria Segre, editor. *Machine Learning of Robot Assembly Plans.* Kluwer Academic, Boston, MA, 1988.

[Sha79] Glen Shafer. *Mathematical theory of Evidence.* Princeton University Press, Princeton, NJ, 1979.

[Sho89] Yoav Shoham. Time for action. In *IJCAI'89: Eleventh International Joint Conference on Artificial Intelligence*, pages 954–959,1173, Detroit, MI, 1989.

[Spe86] David J. Speigelhalter. Probabilistic reasoning in predictive expert systems. In *Uncertainty in Artificial Intelligence (1)*, pages 47–67. North-Holland, Amsterdam, The Netherlands, 1986.

[SS90] Prakash P. Shenoy and Glenn Shafer. Axioms for probability and belief-function propagation. In *Uncertainty in Artificial Intelligence 4*, pages 169–198. North-Holland, Amsterdam, The Netherlands, 1990.

[THD90] Austin Tate, James Hendler, and Mark Drummond. A review of ai planning techniques. In James Allen, James Hendler, and Austin Tate, editors, *Readings in Planning*, pages 26–49. Morgan Kaufmann Publishers, Inc., Los Altos, CA, 1990.

[Ver83] Steve A. Vere. Planning in time: Windows and durations for activities and goals. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5(3):246–251, May 1983.

[vM57] Richard von Mises. *Probability, Statistics and Truth.* Dover Publications, New York, NY, 1957.

[Wel90a] Michael P. Wellman. Fundamental concepts of qualitative probabilistic networks. *Artificial Intelligence*, 44:257–303, 1990.

[Wel90b] Michael P. Wellman. The strips assumption for planning under uncertainty. In *AAAI'90: The Eighth National Conference on Artificial Intelligence*, pages 198–203, Boston, MA, 1990.