The Same-Decision Probability: A New Tool for Decision Making

Suming Chen and Arthur Choi and Adnan Darwiche
Computer Science Department
University of California, Los Angeles
{suming,aychoi,darwiche}@cs.ucla.edu

Abstract

When using graphical models for decision making, a fundamental question is whether one is ready to make a decision (stopping criteria), and if not, what observations should be made to better prepare for a decision (selection criteria). In this paper, we review the notions of entropy and expected utility, which are commonly used for this purpose, and contrast them with a newly introduced notion, called the *same-decision probability,* which can be used both as a stopping criteria for making decisions and as a selection criteria for choosing additional observations. Furthermore, we show that computing the same-decision probability lies in the same complexity class as a general expectation computation problem that is applicable to a wide variety of queries in graphical models, including the computation of non-myopic value of information.

1 Introduction

Probabilistic graphical models have often been used to model a variety of decision problems, e.g., in medical diagnosis (Pauker and Kassirer, 1980; van der Gaag and Coupé, 1999), fault diagnosis (Lu and Przytula, 2006), troubleshooting (Heckerman et al., 1995), and in intrusion detection (Kruegel et al., 2003). In these and similar applications, there are often unobserved variables, such as the state of a patient's health, or the presence of a security breach, leading to two fundamental questions. The first question is whether, given the current observations, the decision maker is ready to commit to a decision. We will refer to this as the *stopping criteria* for making a decision. Assuming the stopping criteria is not met, the second question is what additional observations should be made before the decision maker is ready to make a decision. This typically requires a *selection criteria* based on some measure for quantifying an observation's *value of information* (VOI).

The literature contains a number of proposals for both stopping and selection criteria. On stopping criteria, one may commit to a decision once the belief about a certain event crosses some threshold, as in (Pauker and Kassirer, 1980; Lu and Przytula, 2006). Alternatively, we may simply perform as many observations as our budget allows, as in (Greiner et al., 1996; Krause and Guestrin, 2009). As for selection criteria, different observations may have different values with respect to the decision we are interested in making, possibly taking into account also the cost of performing an observation (Lindley, 1956; Howard, 1966). We may be interested in making an observation that minimizes our expected uncertainty about an event, or we may be interested in maximizing our expected utility; see, e.g. (Kjærulff and Madsen, 2008; Krause and Guestrin, 2009).

In this paper, we consider the use of a recently introduced notion, called the *Same-Decision Probability (SDP)*, as both a stopping criteria and a selection criteria for the purposes of more *robust* decision making. In short, the SDP is the probability of making the same decision even after knowing the values of a set of, as of yet, unobserved variables. (Darwiche and Choi, 2010). We first extend SDP, which was previously proposed for threshold-based decisions supported by Bayesian networks, to more general decision-making tasks, such as those typically supported by influence diagrams. We next consider the potential value of SDP as a stopping criteria, via concrete examples, illustrating how SDP can quantify the stability of a decision in ways that are not evident when we consider beliefs and utilities alone. Next, we consider the potential value of SDP as a selection criteria, in terms of the gain in confidence (as opposed to information) than an observation may bring. Finally, we analyze the complexity of the generalized notion of SDP that we propose, showing that it is $\mathrm{PP}^{\mathrm{PP}}$-complete. This result generalizes previous complexity results for SDP (Choi et al., 2012), but more importantly, applies to a broader class of VOI and reward functions.

2 Technical Preliminaries

We use standard notation for variables and their instantiations, where variables are denoted by upper case letters X and their instantiations by lower case letters x. Additionally, sets of variables are denoted by bold upper case letters $\mathbf{X}$ and their instantiations by bold lower case letters $\mathbf{x}$. We assume that the state of the world is described over random variables $\mathbf{X}$, where the evidence $\mathbf{E} \subseteq \mathbf{X}$ includes **all** known variables, and where hidden variables $\mathbf{U} \subseteq \mathbf{X}$ include **all** unknown variables. By definition, $\mathbf{E} \cap \mathbf{U} = \emptyset$ and $\mathbf{E} \cup \mathbf{U} = \mathbf{X}$. We often discuss the ramifications of observing a subset of hidden variables $\mathbf{H} \subseteq \mathbf{U}$ on decision making. Furthermore, we use $D \in \mathbf{U}$ to denote the hypothesis variable that forms the basis for making a decision.[1]

2.1 Same-Decision Probability

The same-decision probability (SDP) was initially defined in the context of threshold-based decisions (Darwiche and Choi, 2010), where a decision is made if the probability $Pr(d \mid \mathbf{e})$ reaches or surpasses a threshold T. Threshold-based decisions are common and can be found in troubleshooting (Heckerman et al., 1995), medical diagnosis (Pauker and Kassirer, 1980; van der Gaag and Coupé, 1999), anomaly detection (Kruegel et al., 2003), and fault diagnosis (Lu and Przytula, 2006).

Although the SDP was defined in the context of threshold-based decisions, we extend the definition to a more general setting. In particular, we assume that $\mathcal{F}$ is a function that outputs some decision given as input a distribution $Pr(D \mid \mathbf{e})$. SDP is thus defined as the probability that the same decision would be made if the hidden states of variables $\mathbf{H}$ were known (Darwiche and Choi, 2010).

Definition 1. Given a decision function $\mathcal{F}$, hypothesis variable D, unobserved variables $\mathbf{H}$, and evidence $\mathbf{e}$, the **same-decision probability** (SDP) is defined as

$$SDP(\mathcal{F}, D, \mathbf{H}, \mathbf{e}) = \sum_{\mathbf{h}} [\mathcal{F}(Pr(D \mid \mathbf{h}, \mathbf{e}))]_{\mathbf{h}} Pr(\mathbf{h} \mid \mathbf{e}) \quad (1)$$

where $[\mathcal{F}(Pr(D \mid \mathbf{h}, \mathbf{e}))]_{\mathbf{h}}$ is an indicator function that is equal to 1 when $\mathcal{F}(Pr(D \mid \mathbf{h}, \mathbf{e})) = \mathcal{F}(Pr(D \mid \mathbf{e}))$, and equal to 0 otherwise.

Note that the original SDP definition assumed that we had a binary decision and would perform one decision if $Pr(d \mid \mathbf{e}) \geq T$ and the alternative decision otherwise (Darwiche and Choi, 2010).

2.2 Value of Information

We follow (Krause and Guestrin, 2009) by using a general definition of VOI based on reward functions. In particular, given an arbitrary reward function R,[2] hypothesis D, hidden variables $\mathbf{H}$, and evidence $\mathbf{e}$, the VOI is defined as

$$\mathcal{V}(R, D, \mathbf{H}, \mathbf{e}) = \mathcal{ER}(R, D, \mathbf{H}, \mathbf{e}) - R(Pr(D \mid \mathbf{e})) \quad (2)$$

where

$$\mathcal{ER}(R, D, \mathbf{H}, \mathbf{e}) = \sum_{\mathbf{h}} R(Pr(D \mid \mathbf{h}, \mathbf{e})) Pr(\mathbf{h} \mid \mathbf{e}) \quad (3)$$

is the expected reward of observing variables $\mathbf{H}$ and $R(Pr(D \mid \mathbf{e}))$ is the reward if we do

[1] One can extend the analysis to multiple hypothesis variables, but we focus here on the case of one hypothesis variable for simplicity.

[2] We thoroughly discuss R in Section 5.

D	$Pr(D)$
$+$	0.5
$-$	0.5

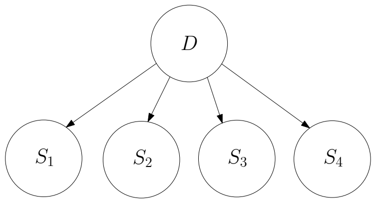

Figure 1: A Bayesian network for intrusion detection, with its CPTs given in Table 1.

Table 1: CPTs for the network in Figure 1.

D	S_1	$Pr(S_1 \mid D)$
$+$	$+$	0.55
$+$	$-$	0.45
$-$	$+$	0.45
$-$	$-$	0.55

D	S_3	$Pr(S_3 \mid D)$
$+$	$+$	0.60
$+$	$-$	0.40
$-$	$+$	0.40
$-$	$-$	0.60

D	S_2	$Pr(S_2 \mid D)$
$+$	$+$	0.55
$+$	$-$	0.45
$-$	$+$	0.45
$-$	$-$	0.55

D	S_4	$Pr(S_4 \mid D)$
$+$	$+$	0.65
$+$	$-$	0.35
$-$	$+$	0.35
$-$	$-$	0.65

not observe $\mathbf{H}$. Note that as in (Krause and Guestrin, 2009), VOI can refer to the expected reward as defined in Equation 3 since $R(Pr(D \mid \mathbf{e}))$ does not depend on variables $\mathbf{H}$.

For example, depending on the reward function used, VOI can be used to identify observations that best increase expected utility or reduce entropy (Kjærulff and Madsen, 2008). The mean-squared error and margins of confidence have also been used as reward functions (Krause and Guestrin, 2009).

3 SDP as a Stopping Criteria

We illustrate in this section the use of SDP as a stopping criteria in the context of threshold-based decisions and expected-utility decisions (i.e., influence diagrams).

3.1 Threshold-Based Decisions

Consider the sensor network in Figure 1, which may correspond to an intrusion detection application as discussed in (Kruegel et al., 2003). Here, the hypothesis variable is $D = \{+, -\}$ with $D = +$ implying an intrusion. Suppose we commit to a decision, and stop performing observations, when our belief in the event $D = +$ surpasses some threshold T, say $T = 0.55$. There are four sensors in this model, S_1, S_2, S_3 and S_4, whose readings may affect this decision.

Consider the two following scenarios: (1) $S_1 = +$ and $S_2 = +$, and (2) $S_3 = +$ and $S_4 = +$. Since $Pr(D = + \mid S_1 = +, S_2 = +) = 0.60 > 0.55$ and $Pr(D = + \mid S_3 = +, S_4 = +) = 0.74 > 0.55$, it is clear that in both cases that the threshold has been crossed. We deem that no further observations are necessary based on our beliefs surpassing our threshold, as in (Kruegel et al., 2003; Lu and Przytula, 2006). Hence, when using thresholds as a stopping criteria, the two scenarios are identical.

From the viewpoint of SDP, however, these two scenarios are very different. In particular, the first scenario leads to an SDP of 52.97%. This means that there is a 47.03% chance that a different decision would be made if we were to further observe the two unobserved sensors S_3 and S_4. The second scenario, however, leads to an SDP of 100%. That is, we would with certainty know that we would make the same decision, if we were to also observe the two unobserved sensors S_1 and S_2: no matter what the readings of S_1 and S_2 could be, our beliefs in the event $D = +$ would always surpass our threshold 0.55. Indeed, as we can see in Table 1, the sensors S_1 and S_2 are not as strong as sensors S_3 and S_4, and in this example, they are not strong enough to reverse our decision.

This example provides a clear illustration of the usefulness of the SDP as a stopping criteria. First, the SDP can pinpoint situations where further observations are unnecessary as they would never reverse the decision under consideration. Second, the SDP can also identify situations where the decision to be made is not robust, and is likely to change upon making further observations.

3.2 Expected-Utility Decisions

We now consider the use of SDP as a stopping criteria in the context of influence diagrams (Howard and Matheson, 1984).

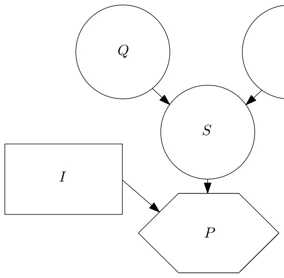

Figure 2: An influence diagram for an investment problem.

Consider the influence diagram in Figure 2, which consists of a Bayesian network with three variables (C, Q and S), a decision node I, and a utility node P that is a direct function of the utility function u. This influence diagram models an investment problem in which a venture capital firm is deciding whether to invest an amount of \$5 million in a tech startup ($I = T$) or to allow the money to collect interest in the bank ($I = F$). In this example, the profit of the investment (P) depends on the decision (I) and the success of the company (S), which in turn depends on two factors: (1) whether the existing competition is successful (C) and (2) whether the co-founders of the startup have a high quality, original idea (Q). Both C and Q are unobserved initially and independent of each other. Variable S is the hypothesis variable in this case and cannot be observed. Variables C and Q, however, can be observed for a price.

The goal here is to choose the decision $I = i$ with the maximum expected utility:

$$\mathcal{EU}(i \mid \mathbf{e}) = \sum_{s} Pr(s \mid \mathbf{e}) u(i, s)$$

where $u(i, s)$ is the utility of decision i given s, whether the company is successful or not.

Figures 3 and 4 contain two different parameterizations of the influence diagram in Figure 2. We will refer to these as different scenarios of the investment problem.

In both scenarios, given no evidence on variables C and Q, the best decision is $I = F$, with an expected utility of \$500K. A decision maker may commit to this decision or decide to observe variables C and Q, with the hope of finding a better decision in light of the additional information. The classical stopping criteria here is to compute the *maximum* expected utility given that we observe variables C and Q (Heckerman et al., 1993; Kjærulff and Madsen, 2008):

$$\sum_{c,q} \max_{i} \mathcal{EU}(i \mid c, q) Pr(c, q)$$

In both scenarios, the maximum expected utility comes out to \$1,180K,[3] showing that further observations may lead to a better decision of $I = T$, i.e. investing in the company.

Up to this point, the above two scenarios are indistinguishable from the viewpoint of classi-

Q	$Pr(Q)$
T	0.4
F	0.6

Q	C	$Pr(S = T \mid .)$
T	T	0.60
T	F	0.90
F	T	0.20
F	F	0.30

C	$Pr(C)$
T	0.6
F	0.4

I	S	$u(I, S)$
T	T	$\$5 \times 10^6$
T	F	$-\$5 \times 10^6$
F	T	$\$5 \times 10^5$
F	F	$\$5 \times 10^5$

Figure 3: A parameterization of the influence diagram in Figure 2.

Q	$Pr(Q)$
T	0.1
F	0.9

Q	C	$Pr(S = T \mid .)$
T	T	0.05
T	F	0.98
F	T	0.01
F	F	0.05

C	$Pr(C)$
T	0.9
F	0.1

I	S	$u(I, S)$
T	T	$\$7 \times 10^7$
T	F	$-\$5 \times 10^6$
F	T	$\$5 \times 10^5$
F	F	$\$5 \times 10^5$

Figure 4: A parameterization of the influence diagram in Figure 2.

[3]According to the formulation of (Krause and Guestrin, 2009), we have computed the VOI for variables C and Q using the reward function.

cal decision making tools. The SDP, however, finds these two scenarios very different. In particular, with respect to variables C and Q, the SDP is 60% in the first scenario and is 99% in the second scenario. That is, even though we stand to make a better decision of $I = T$ in both scenarios upon observing certain instantiations of C and Q, (at least with respect to utility), and even though the expected benefit from such observations is the same in both scenarios, it is very unlikely that we would change the current decision of $I = F$ in the second scenario in comparison to the first. Hence, given the additional information provided by the SDP, a decision maker may act quite differently in these two scenarios. Indeed, when we take a closer look at the second scenario, there is a state of the world that has very high utility (when $I = T$ and $S = T$). However, the chance of this state manifesting itself is extremely small (analogous to a lottery).

This illustrates the usefulness of SDP as a stopping criteria in the context of expected-utility decisions and influence diagrams. Namely, using SDP, we can distinguish between two very different scenarios, that are otherwise indistinguishable when we consider utilities alone.

4 SDP as a Selection Criteria

We now turn our attention to the use of SDP as a criteria for deciding which variables to observe next, assuming that some stopping criteria indicates that further observations are necessary.

Formally, our proposal is based on using VOI as the selection criteria (see Equation 2), while choosing the SDP as the reward function (see Equation 1). We next define SDP gain.

Definition 2. Given Definition 1 of SDP, the **SDP gain** of observing variables $\mathbf{H}$ out of variables $\mathbf{U}$ is defined as

$$\mathcal{G}(\mathbf{H}) = \sum_{\mathbf{h}} SDP(\mathcal{F}, D, \mathbf{U} \setminus \mathbf{H}, \mathbf{he}) Pr(\mathbf{h}|\mathbf{e}) - SDP(\mathcal{F}, D, \mathbf{U}, \mathbf{e}). \quad (4)$$

where $\mathbf{he}$ denotes the joint instantiation of $\mathbf{h}$ and $\mathbf{e}$.

D	$Pr(D)$
$+$	0.5
$-$	0.5

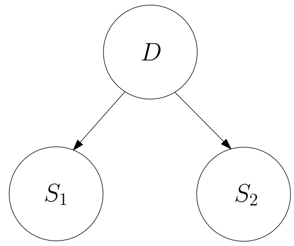

Figure 5: A Bayesian network with its CPTs given in Table 2.

Table 2: CPTs for the Bayesian network in Figure 5.

D	S_1	$Pr(S_1 \mid D)$
$+$	$+$	0.8
$+$	$-$	0.2
$-$	$+$	0.2
$-$	$-$	0.8

D	S_2	$Pr(S_2 \mid D)$
$+$	$+$	0.75
$+$	o	0.2
$+$	$-$	0.05
$-$	$+$	0.05
$-$	o	0.2
$-$	$-$	0.75

The goal here is to observe those variables which, on average, will lead to the highest SDP. That is, we want to maximize the probability of making a decision that is unlikely to change even if we observe the remaining variables.

We will next provide an example of using SDP as a selection criteria, contrasting it with two other selection criteria: One based on reducing entropy of the hypothesis variable D, and another based on maximizing the margin between the first and second most likely states. While both criteria can be motivated as reducing uncertainty, we will show that both can indeed lead to less stable decisions in contrast to SDP.

The example is given by the Bayesian network in Figure 5, where D is the hypothesis variable and S_1/S_2 are sensors. A decision is triggered when $Pr(D = + \mid \mathbf{e}) \geq .80$, where evidence $\mathbf{e}$ is over sensors S_1 and S_2. With no observations (empty evidence $\mathbf{e}$), the SDP is 0.595, suggesting that further observations may be needed. Assuming a limited number of observations (Heckerman et al., 1995), and observing one variable at a time (Dittmer and Jensen, 1997), we need now to select the next variable to observe.

Note that maximizing VOI with negative entropy as the reward function amounts to maximizing mutual information (Cover and Thomas, 1991). The mutual information between variable D and sensor S_2 is 0.53 whereas the mutual information between D and sensor S_1 is 0.278. Hence, observing S_2 will reduce the entropy of D the most. In terms of margin of confidence, another reward function (Krause and Guestrin, 2009), observing S_2 will on average lead to a 0.7 margin between the $D = +$ and $D = -$, whereas observing S_1 will only lead to a 0.6 margin between the two states.

However, if we compute the corresponding SDP gains, $\mathcal{G}(S_1)$ and $\mathcal{G}(S_2)$, we find that observing S_1 will, on average, lead to improving the decision stability the most. In particular, observing S_1 would give us an SDP of either 1 or 0.81, resulting in an expected SDP of 0.905. Observing S_2 would give us an SDP of either 0.7625, 0.5, or 1, resulting in an expected SDP of 0.805. Therefore, $\mathcal{G}(S_1) = 0.31$ and $\mathcal{G}(S_2) = 0.21$. Hence, observing S_1 will on average allow us to make a decision that is less likely to change due to additional information.

Some intuition to why this is the case is that in threshold-based decisions, we make a decision solely based on whether $Pr(D \mid \mathbf{e})$ is above or below the threshold. Selection criteria such as entropy and margins of confidence will not consider the threshold. This example demonstrates the usefulness of SDP as a selection criteria for threshold-based decisions, as the SDP can be used to select observations that lead to more robust decisions.

5 Computational Complexity

The SDP was shown to be a $\mathrm{PP}^{\mathrm{PP}}$-complete problem in (Choi et al., 2012). The $\mathrm{PP}^{\mathrm{PP}}$ class can be thought of as a counting variant of the $\mathrm{NP}^{\mathrm{PP}}$ class, for which the MAP problem is complete (Park and Darwiche, 2004).

We show in this section that a general problem of computing expectations is also $\mathrm{PP}^{\mathrm{PP}}$-complete, with non-myopic VOI being an instance of such an expectation. We also show that the SDP is another instance of this computation. Thus, the development of algorithms for SDP will be beneficial to problems in the complexity class $\mathrm{PP}^{\mathrm{PP}}$, which in turn benefits computing an assortment of expectations, including non-myopic VOI.

The proposed expectation computation is based on using a reward function R with some properties that we review next. In particular, the function R is assumed to map a probability distribution $Pr(D \mid \mathbf{e})$ to a numeric value. We also assume that the minimum l and maximum u of this range are polytime computable. These assumptions are not too limiting—for example, both entropy and utility can be expressed using reward functions that fall in this category (Krause and Guestrin, 2009).

We now consider the following computation of expectations.

D-EPT: Given reward function R, hypothesis variable D, unobserved variables $\mathbf{H}$, evidence $\mathbf{e}$, a real number N, and a distribution Pr induced by a Bayesian network over variables $\mathbf{X}$,[4] the **expectation decision problem** asks: Is

$$\mathbb{E} = \sum_{\mathbf{h}} R(Pr(D \mid \mathbf{h}, \mathbf{e}))Pr(\mathbf{h} \mid \mathbf{e})$$

greater than N?

It should be clear that the SDP falls as a special case when the reward function R is the SDP indicator function (see Definition 1). For example, in (Choi et al., 2012), the decision function outputs one of two decisions depending on whether $Pr(d|\mathbf{e}) > T$ for some value d of D and some threshold T.

We now have the following theorems, with proofs in the Appendix.

Theorem 1. ***D-EPT*** *is* $\mathrm{PP}^{\mathrm{PP}}$*-hard.*

Theorem 2. ***D-EPT*** *is in* $\mathrm{PP}^{\mathrm{PP}}$*.*

This shows that **D-EPT** is $\mathrm{PP}^{\mathrm{PP}}$-complete. This also implies that computing the SDP is $\mathrm{PP}^{\mathrm{PP}}$-complete, as are other computational problems such as non-myopic VOI using a variety of reward functions.

[4]This proof also holds for influence diagrams constrained to have only one decision node.

6 Conclusion

In this paper, we extended the recently introduced notion of same decision probability (SDP), as a way to quantify the robustness of a decision, to a broader class of decision-making problems. Through concrete examples, we illustrated the usefulness of SDP as a stopping criterion, where SDP is capable of distinguishing scenarios that are otherwise indistinguishable based on thresholds or utilities alone. We further illustrated the usefulness of SDP as a selection criterion, in the terms of a confidence gain (in contrast to information gain). Finally, we provided a general analysis of the complexity of SDP, which extends itself to a broad class of functions formulated as expectations, including a general class of VOI computations.

7 Appendix

See (Kwisthout, 2009; Choi et al., 2012) for more on the complexity class PP^{PP} in the context of reasoning in Bayesian networks.

Proof of Theorem 1. We show **D-EPT** is PP^{PP}–hard by reduction from the following decision problem **D-SDP**, which corresponds to the originally proposed notion of same-decision probability for threshold-based decisions (Darwiche and Choi, 2010).

D-SDP: Given a decision based on probability $Pr(d \mid \mathbf{e})$ surpassing a threshold T, a set of unobserved variables $\mathbf{H}$, and a probability p, is the same-decision probability:

$$\sum_{\mathbf{h}} [Pr(d \mid \mathbf{h}, \mathbf{e}) \geq T] Pr(\mathbf{h} \mid \mathbf{e}) \qquad (5)$$

greater than p?

Here, [.] denotes an indicator function which evaluates to 1 if the enclosed expression is satisfied, and 0 otherwise. **D-SDP** was shown to be PP^{PP}–complete in (Choi et al., 2012).

This same-decision probability corresponds to an expectation with respect to the distribution $Pr(\mathbf{H} \mid \mathbf{e})$, using the reward function:

$$R(Pr(D \mid \mathbf{h}, \mathbf{e})) = \begin{cases} 1 & \text{if } Pr(d \mid \mathbf{h}, \mathbf{e}) \geq T \\ 0 & \text{otherwise.} \end{cases}$$

Thus the same-decision probability is greater than T iff this expectation is greater than T. □

Proof of Theorem 2. To show that **D-EPT** is in PP^{PP}, we provide a probabilistic polynomial-time algorithm, with access to a PP oracle, that answers the decision problem **D-EPT** correctly with probability greater than $\frac{1}{2}$. This proof generalizes and simplifies the proof given in (Choi et al., 2012) for **D-SDP**.

Consider the following probabilistic algorithm that determines if $\mathbb{E} > N$:

1. Sample a complete instantiation $\mathbf{x}$ from the Bayesian network, with probability $Pr(\mathbf{x})$. We can do this in linear time, using forward sampling (Henrion, 1986).

2. If $\mathbf{x}$ is compatible with $\mathbf{e}$, we can use a PP–oracle to compute $t = R(Pr(D \mid \mathbf{h}, \mathbf{e}))$. First, the reward function R can be computed in polynomial time, by definition. Second, $Pr(D \mid \mathbf{h}, \mathbf{e})$ can be computed using a PP–oracle, since the decision problem for marginals is PP–complete (Roth, 1996), and since $\text{P}^{\text{PP}} = \text{P}^{\#\text{P}}$.

3. Define a function $a(t) = \frac{1}{2} + \frac{1}{2}\frac{t-N}{u-l}$, which defines a probability used by our probabilistic algorithm to guess whether $\mathbb{E} > N$ (see Lemma 1).

4. Declare that $\mathbb{E} > N$ with probability:
   - $a(t)$ if $\mathbf{x}$ is compatible with $\mathbf{e}$;
   - $\frac{1}{2}$ if $\mathbf{x}$ is not compatible with $\mathbf{e}$.

The probability of declaring $\mathbb{E} > N$ is:

$$r = \sum_{\mathbf{h}} a(t) Pr(\mathbf{h}, \mathbf{e}) + \frac{1}{2}(1 - Pr(\mathbf{e})) \qquad (6)$$

which is greater than $\frac{1}{2}$ iff the following set of equivalent statements hold:

$$\sum_{\mathbf{h}} a(t) Pr(\mathbf{h}, \mathbf{e}) > \frac{Pr(\mathbf{e})}{2}$$

$$\sum_{\mathbf{h}} a(t) Pr(\mathbf{h} \mid \mathbf{e}) > \frac{1}{2}$$

$$\sum_{\mathbf{h}} \left(\frac{1}{2} + \frac{1}{2} \frac{t-N}{u-l} \right) Pr(\mathbf{h} \mid \mathbf{e}) > \frac{1}{2}$$

$$\sum_{\mathbf{h}} \left(\frac{1}{2} \frac{t-N}{u-l} \right) Pr(\mathbf{h} \mid \mathbf{e}) > 0$$

$$\sum_{\mathbf{h}} (t-N) Pr(\mathbf{h} \mid \mathbf{e}) > 0$$

$$\sum_{\mathbf{h}} R(Pr(D \mid \mathbf{h}, \mathbf{e})) Pr(\mathbf{h} \mid \mathbf{e}) > N.$$

Thus $r > \frac{1}{2}$ iff $\mathbb{E} > N$. □

Lemma 1. *The function* $a(t) = \frac{1}{2} + \frac{1}{2}\frac{t-N}{u-l}$ *maps a reward* t *to a probability in* $[0, 1]$.

Proof. Values u and l are given, and denote upper and lower bounds on the reward t, but also the threshold N. Thus $\frac{t-N}{u-l}$ is in $[-1, 1]$. □

Note that $a(t)$ denotes a probability used by our algorithm to declare whether $\mathbb{E} > N$, which is higher or lower depending on the value of the reward $t = R(Pr(D \mid \mathbf{h}, \mathbf{e}))$.

References

Arthur Choi, Yexiang Xue, and Adnan Darwiche. 2012. Same-decision probability: A confidence measure for threshold-based decisions. *International Journal of Approximate Reasoning (IJAR)*.

Thomas M. Cover and Joy A. Thomas. 1991. *Elements of information theory*. Wiley-Interscience, New York, NY, USA.

Adnan Darwiche and Arthur Choi. 2010. Same-decision probability: A confidence measure for threshold-based decisions under noisy sensors. In *Proceedings of the Fifth European Workshop on Probabilistic Graphical Models*, pages 113–120. PGM.

Soren Dittmer and Finn Jensen. 1997. Myopic value of information in influence diagrams. In *Proceedings of the Thirteenth Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-97)*, pages 142–149.

Russell Greiner, Adam J. Grove, and Dan Roth. 1996. Learning active classifiers. In *Proceedings of the Thirteenth International Conference on Machine Learning (ICML96*.

David Heckerman, Eric Horvitz, and Blackford Middleton. 1993. An approximate nonmyopic computation for value of information. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(3):292–298.

David Heckerman, John S. Breese, and Koos Rommelse. 1995. Decision-theoretic troubleshooting. *Commun. ACM*, 38(3):49–57, March.

Max Henrion. 1986. Propagating uncertainty in bayesian networks by probabilistic logic sampling. In *Uncertainty in Artificial Intelligence 2 Annual Conference on Uncertainty in Artificial Intelligence (UAI-86)*, pages 149–163, Amsterdam, NL. Elsevier Science.

Ronald A. Howard and James E. Matheson, editors. 1984. *Readings on the Principles and Applications of Decision Analysis*. Strategic Decision Group, Menlo Park, CA.

Ronald A. Howard. 1966. Information value theory. *IEEE Trans. Systems Science and Cybernetics*, 2(1):22–26.

Uffe B. Kjærulff and Anders L. Madsen. 2008. *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis*. Information Science and Statistics. Springer.

Andreas Krause and Carlos Guestrin. 2009. Optimal value of information in graphical models. *Journal of Artificial Intelligence Research (JAIR)*, 35:557–591.

Christopher Kruegel, Darren Mutz, William Robertson, and Fredrik Valeur. 2003. Bayesian Event Classification for Intrusion Detection. In *Proceedings of the Annual Computer Security Applications Conference (ACSAC)*, Las Vegas, NV USA, 12.

Johan Kwisthout. 2009. *The Computational Complexity of Probabilistic Networks*. Ph.D. thesis, University of Utrecht.

Dennis V. Lindley. 1956. On a measure of the information provided by an experiment. *Annals of Mathematical Statistics*, 27(4):986–1005.

Tsai-Ching Lu and K. Wojtek Przytula. 2006. Focusing strategies for multiple fault diagnosis. In *FLAIRS Conference*, pages 842–847.

James D. Park and Adnan Darwiche. 2004. Complexity Results and Approximation Strategies for MAP Explanations. *J. Artif. Intell. Res. (JAIR)*, 21:101–133.

Steven G. Pauker and Jerome P. Kassirer. 1980. The threshold approach to clinical decision making. *N Engl J Med*, 302(20):1109–17.

Dan Roth. 1996. On the hardness of approximate reasoning. *Artificial Intelligence*, 82:273 – 302.

Linda C. van der Gaag and Veerle M. H. Coupé. 1999. Sensitive analysis for threshold decision making with bayesian belief networks. In *AI*IA*, pages 37–48.