SENSITIVITY AND ROBUSTNESS ANALYSIS OF BAYESIAN NETWORKS

Fabio Gagliardi Cozman
fgcozman@usp.br

Escola Politécnica, Universidade de São Paulo,
Av. Prof. Mello Moraes, 2231, Cidade Universitária, São Paulo, SP, Brasil

Abstract – Probabilistic reasoning is one of the fundamental aspects of intelligent systems, and probability theory provides a broad foundation for several methods used in intelligent devices. One of the most flexible tools for probabilistic reasoning is the theory of Bayesian networks. Bayesian networks are graphical structures that represent joint probability measures through collections of conditional probabilities. This paper presents effective procedures for robustness and sensitivity analysis of inferences in Bayesian networks. Robustness analysis evaluates global variations through lower and upper expectations. Sensitivity analysis evaluates local variations through functional derivatives. The main contribution of this paper is the derivation of procedures for efficient and fast sensitivity analysis so that a balance can be found, from a computational point of view, between global and local analysis. Results in this paper can be used during the design phase of intelligent systems associated with probabilistic reasoning.

Keywords: Bayesian networks, convex sets of probability, robustness analysis, sensitivity measures

1 INTRODUCTION

Probabilistic reasoning is one of the fundamental aspects of intelligent systems, as it provides a solid foundation for the manipulation of chance phenomena and subjective uncertainty. Due to this, probability theory currently provides a broad foundation for several methods used in intelligent devices. Amongst these methods, one of the most flexible tools for probabilistic reasoning is the theory of Bayesian networks.

Bayesian networks are quite popular graphical structures that allow efficient manipulation of statistical uncertainty. A Bayesian network represents a single joint probability measure through a directed acyclic graph and a collection of conditional probability functions (Jensen 1996). In practice, such probability values are obtained from experts or estimated from data. It may be difficult to obtain all probability values precisely due to several factors: (1) experts may hesitate, or be unable, to provide exact values; (2) experts in a group may have serious disagreents; (3) statistical probability estimates may only generate intervals; (4) probability estimates may depend on incomplete databases. The objective of this paper is to present tools that evaluate the effect of probability imprecision in Bayesian networks.

A decision-maker may want to obtain information on the global effect of all possible variations of probability values in a network. Such global analysis is called *robustness analysis* (Berger 1990). Section 3 suggests a procedure for robustness analysis based on the theory of convex sets of probability measures.

The decision-maker may instead be interested in a simplified, local analysis of the effects of variations in probability values. This type of local analysis is called *sensitivity analysis* (Gustafson and Wasserman 1995), and is usually associated with differential figures of merit that concentrate on the effect of small variations around some "base model." There are quite a few different approaches to sensitivity analysis in the literature, with distinct objectives and strategies. Section 4 contains a detailed discussion of current approaches and an adaptation of methods available in the literature to the specific characteristics of Bayesian networks.

The main objective ot this paper is to derive algorithms for efficient and fast sensitivity analysis, so that a balance can be found between global and local analysis in Bayesian networks. Algorithms presented here have distinct complexity levels and produce different types of information; together, they create a framework for the study of probability variations and imprecisions in Bayesian networks. This collection of algorithms can be used during the design phase of intelligent systems associated with probabilistic reasoning, where the numbers and relationships in a particular domain are under study.

2 PROBABILISTIC IMPRECISION IN BAYESIAN NETWORKS

A Bayesian network defines a joint probability density through a directed acyclic graph. Each node of this graph is associated with a variable X_i and with a conditional probability density $p(X_i|\mathrm{pa}(X_i))$ (the parents of the node X_i are denoted $\mathrm{pa}(X_i)$). By definition, this graphical structure encodes the following joint density (Pearl 1988):

$$p(\mathbf{X}) = \prod_i p(X_i|\mathrm{pa}(X_i)) \,. \tag{1}$$

Figure 1 depicts a Bayesian network commonly used as an example (Charniak 1991). The network summarizes a story: the variable F indicates whether a family is at home or travelling; the variable B indicates whether the family dog is having health

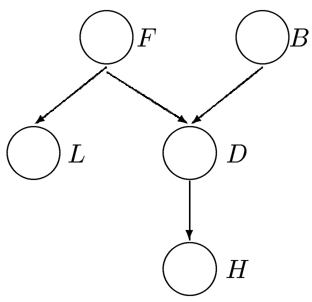

Figure 1: Bayesian network (variables are binary; superscript c indicates negation).

problems; the variable D indicates whether the dog is inside the family's house or in the garden. Presumably D is directly affected by F and B; that is indicated the directed arrows from F and B to D. Variable L indicates whether the lights in the family's house are on or off; this variable depends directly on F. Finally, variable H indicates whether or not one can hear the dog barking, a fact that directly depends on D.

Given a Bayesian network, an *inference* is usually understood as the calculation of posterior probabilities or expectations for *queried* variables $\mathbf{X}_Q$, conditional on the event that some variables have been observed (this event is called the *evidence* E) (Jensen 1996). So, in Figure 1 one may contemplate the probability of the family being at home (F = "home") given that the lights are on. This paper considers only categorical variables, as it is common in the literature. A particular value of variable X_i is denoted by X_{ij}; a particular value of $\mathrm{pa}(X_i)$ is denoted by $[\mathrm{pa}(X_i)]_k$. A probability value $p(X_{ij}|\,[\mathrm{pa}(X_i)]_k)$ is denoted by p_{ijk}.

Given a Bayesian network, a decision-maker may be interested on the posterior expected value of a particular function $f(\mathbf{X}_Q)$. Sensitivity and robustness analysis must then focus on the possible variations of the expectation $E_p[f(\mathbf{X}_Q)|E]$ given variations in p_{ijk}. Alternatively, the decision-maker may be interested in computing the whole posterior density $p(\mathbf{X}_Q|E)$. Sensitivity and robustness analysis must then focus on the behavior of the function $p(\mathbf{X}_Q|E)$. These two possibilities can be summarized by stating that a functional $F\,[p(\mathbf{X}_Q|E)]$ must be analyzed with respect to variations in some values p_{ijk}. The difference between robustness and sensitivity analysis, as understood in this paper, is that the first one seeks tight lower and upper bounds for $F\,[p(\mathbf{X}_Q|E)]$, while the second one seeks some differential measure of variation for $F\,[p(\mathbf{X}_Q|E)]$.

Methods that quantify probability imprecision, in global and local senses, are presented in the next sections. But how to *represent* probability imprecision? The basic tool employed in this paper is the theory of convex sets of probability measures, taken as a natural representation for any statement of imprecise probability. This paper follows Giron and Rios' axiomatization for sets of probability measures, called *Quasi-Bayesian* theory (Giron and Rios 1980).

A closed convex set of probability measures is called a *credal set*; similarly, a set of joint measures is called a *joint credal set*. Convexity of a set K means that, given any two measures P_1 and P_2 in K, the convex combination $\alpha P_1 + (1-\alpha)P_2$ (for $\alpha \in [0,1]$) is also in K. To simplify terminology, a convex set $K(X)$ containing probability densities $p(X)$ (defined whenever possible) is also called a credal set. The *conditional credal set* $K(X|A)$ is defined by conditional densities $p(X|A)$. Quasi-Bayesian inference is performed by applying Bayes rule to each measure in a joint credal set. A *posterior credal set* is the union of all posterior measures obtained in this process; a posterior lower expectation is equal to:

$$\underline{E}[f(X)|A] = \min_{p(X)\in K(X)} \frac{E_p[f(X)I_A(X)]}{E_p[I_A(X)]}.$$

If the event A can have probability equal to zero, the posterior credal set $K(X|A)$ is taken *by convention* to be the set of all probability densities $p(X|A)$ (Walley 1991). Otherwise, the vertices of the posterior credal set are obtained by applying Bayes rule to each vertex of the prior credal set (Levi 1980).

Given a Bayesian network, suppose that variations on a density $p(X_i|\,[\mathrm{pa}(X_i)])$ are encoded by a credal set $K(X_i|\mathrm{pa}(X_i))$. The interpretation is that any density $p(X_i|\,[\mathrm{pa}(X_i)]_k)$ in $K(X_i|\,[\mathrm{pa}(X_i)]_k)$ is a valid conditional density in the network. A directed acyclic graph where nodes are associated with variables X_i and credal sets $K(X_i|\mathrm{pa}(X_i))$ is called a *Quasi-Bayesian network* (Cozman 1997); the basic idea in this paper is that robustness and sensitivity analysis operate on a Quasi-Bayesian network.

Algorithms in this paper assume that credal sets $K(X_i|\mathrm{pa}(X_i))$ are defined as the convex hull of a finite number of probability measures; such *finitely generated credal sets* are polytopes in the space of probability measures. As an extreme example, consider a situation of complete ignorance, where a credal set is taken to contain every possible density $p(X_i|\,[\mathrm{pa}(X_i)]_k)$. This "vacuous" credal set is the convex hull of the densities

$$p_j(X_i|\,[\mathrm{pa}(X_i)]_k) = \begin{cases} 1 & \text{if } X_i = X_{ij}, \\ 0 & \text{otherwise,} \end{cases}$$

for the finitely many values of X_i.

3 ROBUSTNESS ANALYSIS

Suppose that a given bounded function $f(\mathbf{X}_Q)$ is specified and a decision-maker wishes to know whether the posterior expectation $E_p[f(\mathbf{X}_Q)|E]$ is robust to variations in a Quasi-Bayesian network. The concepts of *lower* and *upper expectatations* are appropriate for this type of analysis.

Lower and upper expectations for any bounded function $f(X)$ are respectively defined as:

$$\underline{E}[f(X)] = \min_{p(X)\in K(X)} E_p[f(X)]\,,$$

$$\overline{E}[f(X)] = \max_{p(X)\in K(X)} E_p[f(X)]\,.$$

Due to the linearity of expectations, lower and upper expectations are always attained at the vertices of a credal set. Bounds on the probability of any event A can be obtained by taking the lower expectation of the indicator function $I_A(X)$, which is one if $X \in A$ and zero otherwise. For a credal set $K(X)$, the function $\underline{P}(A) = \min_{p(X)\in K(X)} P(A) = \min_{p(X)\in K(X)} E_p[I_A(X)]$ is called a *lower envelope* and the function $\overline{P}(A) = \max_{p(X)\in K(X)} P(A) = \max_{p(X)\in K(X)} E_p[I_A(X)]$ is called a *upper envelope*.

Given a Quasi-Bayesian network, evidence E and a function $f(\mathbf{X}_Q)$, the "robustness of $E_p[f(\mathbf{X}_Q)|E]$" is measured by the difference between upper and lower bounds of $E_p[f(\mathbf{X}_Q)|E]$. To obtain lower and upper posterior expectations for $f(\mathbf{X}_Q)$, it is necessary to calculate $E_p[f(\mathbf{X}_Q)|E]$ for every vertex of $K(\mathbf{X})$. What exactly are the vertices of a given Quasi-Bayesian network? The answer to this question is somewhat open, as there may be different joint credal sets $K(\mathbf{X})$ that are represented by a Quasi-Bayesian network (i.e., which have the correct conditional credal sets and satisfy relations of irrelevance and independence) (Cozman 1998).

The simplest approach, from the standpoint of computational complexity, is to take the joint credal set $K(\mathbf{X})$ to be the *type-1 extension* of the network (Cozman 1997):

Definition 1 *The type-1 extension of a Quasi-Bayesian network is the convex hull of all joint densities generated by Expression (1), where each conditional probability $p(X_i|\mathrm{pa}(X_i))$ is one of the vertices of the conditional credal set $K(X_i|\mathrm{pa}(X_i))$.*

This convention is followed by a variety of authors (Andersen and Hooker 1994, Cano *et al.* 1993, Tessem 1992). Generating the vertices of a type-1 extension is straightforward: every combination of vertices of the various conditional credal sets $K(X_i|\mathrm{pa}(X_i))$ is valid. Exact inferences with a type-1 extension can be performed by applying any standard Bayesian network algorithm to each one of the vertices of the type-1 extension and minimizing/maximizing over the results (Cano *et al.* 1993, Cano *et al.* 1994, Cano and Moral 1996, Cozman 1997). Type-1 extensions have the useful property that all variables that affect an inference can be detected by polynomial-time algorithms (Cozman 1998).

The fundamental difficulty with type-1 extensions, and with robustness analysis in general, is the potentially enormous number of vertices that a joint credal set may have. Execution of a comprehensive round of robustness analysis, where all possible variations in a model are taken into account simultaneously, may become too expensive computationally. A viable solution is to employ sensitivity analysis to guide the selection of those conditional probability values that affect inferences the most, and perform robustness analysis only with respect to those values.

4 SENSITIVITY ANALYSIS

Suppose a decision-maker wishes to study the rate of variation of the expectation $E_p[f(\mathbf{X}_Q)|E]$ with respect to a probability value p_{ijk}. The following partial derivative is a natural measure:

$$S^E_{ijk} = \frac{\partial E_p[f(\mathbf{X}_Q)|E]}{\partial p_{ijk}}.$$

This sensitivity measure must be slightly modified to guarantee that the joint density $p(\mathbf{X})$ adds up to one even in the presence of perturbations. This is indeed the path taken by Laskey (Laskey 1995), who suggests a possible correction to the joint density $p(\mathbf{X})$ to compensate for changes in p_{ijk}.

The sensitivity measure S^E_{ijk} is somewhat cumbersome due to its narrowness. A complete sensitivity analysis may involve a potentially large number of values, one for each value of X_i and each value of $\mathrm{pa}(X_i)$. If several conditional credal sets are present in a Quasi-Bayesian network, a numbing variety of indexes may have to be analyzed. This may be made even worse, as the decision-maker may be interested in the lower and upper envelopes of $p(\mathbf{X}_Q|E)$.

Instead, a more concise sensitivity measure would be a measure S^E_i defined as a summary of the differential influence in $E_p[f(\mathbf{X}_Q)|E]$ exerted by the conditional probability associated with variable X_i. Such a sensitivity measure would be well-suited to the analysis of Bayesian networks, as S^E_i would concentrate exactly on the basic blocks of Bayesian networks, the conditional densities $p(X_i|\mathrm{pa}(X_i))$.

The problem then is how to define a meaningful measure S^E_i, or how to establish a meaningful "rate of variation" of an expectation with respect to a function. To do so, the next section adapts ideas from functional analysis that are commonly employed in Statistics (Gustafson and Wasserman 1995).

4.1 Quantifying sensitivity by functional derivatives

Consider first a single variable X and a probability density $p(X)$. In the space of probability densities, one can pick a vector going from $p(X)$ to another density $q(X)$. A density in this vector can be written as:

$$q_\epsilon(X) = (1-\epsilon)p(X) + \epsilon q(X).$$

Now suppose that an event E is observed. We can measure the local change of the expectation $E_p[f(X)|E]$ by the limit

$$S^E_q = \lim_{\epsilon \to 0} \frac{|\ E_p[f(X)|E] - E_{q_\epsilon}[f(X)|E]\ |}{\epsilon}.$$

To calculate S^E_q, note that

$$q_\epsilon(X|E) = \frac{(1-\epsilon)p(X|E)\,p(E) + \epsilon q(X|E)q(E)}{(1-\epsilon)p(E) + \epsilon q(E)}.$$

Due to the linearity of expectations:

$$\begin{aligned} S^E_q &= \lim_{\epsilon \to 0} \left| \frac{\epsilon}{(1-\epsilon)p(E) + \epsilon q(E)} \right. \\ &\qquad \left. (E_p[f(X)|E] - E_q[f(X)|E]) \right| / \epsilon \\ &= \frac{q(E)}{p(E)}\ |\ E_p[f(X)|E] - E_q[f(X)|E]\ |\ . \end{aligned}$$

It may be reasonable to normalize the change of $p(X|E)$ with respect to the variation of the prior $p(X)$ and use the following sensitivity measure:

$$s^E_q = \frac{S^E_q}{||p(X), q(X)||}, \tag{2}$$

where $||p,q||$ is some distance measure between $p(X)$ and $q(X)$ (not necessarily a norm).

The sensitivity measures S^E_q and s^E_q capture the local change of $p(X)$ in the particular direction of $q(X)$. To obtain a sensitivity measure that depends only on the density $p(X)$, it is common to maximize with respect to some set of densities $K(X)$ (Gustafson and Wasserman 1995):

$$S^E_K = \max_{q(X)\in K(X)} S^E_q;$$

similarly,

$$s^E_K = \max_{q(X)\in K(X)} s^E_q.$$

4.2 Functional derivatives in Bayesian networks

The basic expressions used to define S_K^E and s_K^E in the previous section are now elaborated upon and applied to Bayesian networks. Denote by $q_l(\mathbf{X})$ the joint density induced by a Bayesian network when the conditional density $p(X_l|\mathrm{pa}(X_l))$ is replaced by a different density $q(X_l|\mathrm{pa}(X_l))$:

$$q_l(\mathbf{X}) = q(X_l|\mathrm{pa}(X_l)) \times \prod_{i \neq l} p(X_i|\mathrm{pa}(X_i)) .$$

A perturbation $q_{l,\epsilon}(\mathbf{X})$ to the network is defined as:

$$\begin{aligned} q_{l,\epsilon}(\mathbf{X}) &= (1-\epsilon)p(\mathbf{X}) + \epsilon q_l(\mathbf{X}) \\ &= ((1-\epsilon)p(X_l|\mathrm{pa}(X_l)) + \epsilon q(X_l|\mathrm{pa}(X_l))) \times \\ &\quad \prod_{i \neq l} p(X_i|\mathrm{pa}(X_i)) . \end{aligned}$$

A meaningful sensitivity measure can then be written as:

$$\begin{aligned} S_{q_l}^E &= \lim_{\epsilon \to 0} \frac{|\ E_p[f(\mathbf{X}_Q)|E] - E_{q_{l,\epsilon}}[f(\mathbf{X}_Q)|E]\ |}{\epsilon} \\ &= \frac{q_l(E)}{p(E)} \ |\ E_p[f(\mathbf{X}_Q)|E] - E_{q_l}[f(\mathbf{X}_Q)|E]\ | . \end{aligned}$$

Suppose that conditional credal sets $K(X_i|\mathrm{pa}(X_i))$ are specified for a number of variables. Each one of these credal sets specifies the possible variations for the corresponding conditional densities. The objective of sensitivity analysis is to determine the local influence of such densities as they move inside their credal sets.

For any given credal set $K(X_i|\mathrm{pa}(X_i))$, assume that a particular density $p(X_i|\mathrm{pa}(X_i))$ in the set is chosen as the "base" density. Differential changes from the base density must be captured by $S_{q_l}^E$. In the expression for $S_{q_l}^E$, the densities $p(X_i|\mathrm{pa}(X_i))$ are base densities, while the "perturbation" $q(X_l|\mathrm{pa}(X_l))$ is an arbitrary density in $K(X_l|\mathrm{pa}(X_l))$.

To remove the explicit dependency on a particular $q(X_l|\mathrm{pa}(X_i))$ from the sensitivity measure $S_{q_l}^E$, consider the set $K'(X_l|\mathrm{pa}(X_l))$ of all possible functions $q(X_l|\mathrm{pa}(X_l))$ selected from the conditional credal sets $K(X_l|\,[\mathrm{pa}(X_l)]_k)$. Then define $S_{K_l}^E$, the sensitivity measure with respect to $K'(X_l|\mathrm{pa}(X_i))$:

$$S_{K_l}^E = \max_{q(X_l|\mathrm{pa}(X_l)) \in K'(X_l|\mathrm{pa}(X_l))} S_{K_l}^E .$$

The linearity of $S_{q_l}^E$ guarantees:

Theorem 1 *The maximum value of $S_{q_l}^E$ is attained at one of the vertices of $K'(X_l|\mathrm{pa}(X_l))$.*

This theorem indicates that calculation of $S_{K_l}^E$ is straightforward. For each vertex $q(X_l|\mathrm{pa}(X_l))$ of $K'(X_l|\mathrm{pa}(X_l))$, it is necessary to obtain $q(E)$ and $q(\mathbf{X}_Q|E)$ using standard Bayesian network algorithms, and then calculate $S_{K_l}^E$.

Following the rationale employed in Expression (2) to define s_q^E, it is appropriate to normalize $S_{K_l}^E$ for each $q(X_l|\mathrm{pa}(X_l))$. This operation finally leads to the main objective of this section, a sensible definition for the sensitivity measure S_l^E:

$$\begin{aligned} S_l^E = &\max_{q(X_l|\mathrm{pa}(X_l)) \in \mathrm{ext} K'(X_l|\mathrm{pa}(X_l))} \\ &\frac{q_l(E)}{p(E)} \frac{|E_p[f(\mathbf{X}_Q)|E] - E_{q_l}[f(\mathbf{X}_Q)|E]|}{\|p(X_l|\mathrm{pa}(X_l)), q(X_l|\mathrm{pa}(X_l))\|} . \end{aligned}$$

A reasonable distance measure, among several possibilities, is (Gustafson and Wasserman 1995):

$$\begin{aligned} \|p(X_l|\mathrm{pa}(X_l)), q(X_l|\mathrm{pa}(X_l))\| = &\max_{X_l, \mathrm{pa}(X_l)} \quad (3) \\ &|p(X_l|\mathrm{pa}(X_l)) - q(X_l|\mathrm{pa}(X_l))| (4) \end{aligned}$$

A different situation occurs if the decision-maker wishes to quantify the sensitivity of the whole density $p(\mathbf{X}_Q|E)$, instead of focusing on a particular expectation. A solution is to use a distance function and define:

$$\begin{aligned} S_{q_l}^D &= \lim_{\epsilon \to 0} \frac{\|p(\mathbf{X}_Q|E), q_{l,\epsilon}(\mathbf{X}_Q|E)\|}{\epsilon} \\ &= \frac{q(E)}{p(E)} \|p(\mathbf{X}_Q|E), q_l(\mathbf{X}_Q|E)\|. \end{aligned}$$

Now this sensitivity measure can be extended to the credal sets $K(X_l|\mathrm{pa}(X_l))$ by maximization. At first it may seem that maximization of $S_{q_l}^D$ is quite complex, as the distance $\|p(\mathbf{X}_Q|E), q_l(\mathbf{X}_Q|E)\|$ may be nonlinear. For the max-distance (Expression (3)), the triangle inequality is valid and $S_{q_l}^D$ is a concave function in q_l; consequently:

Theorem 2 *The maximum value of $S_{q_l}^D$ is attained at one of the vertices of $K'(X_l|\mathrm{pa}(X_l))$.*

This theorem indicates that the same maximization and normalization strategies can be used with $S_{q_l}^D$, yielding a sensitivity measure

$$\begin{aligned} S_l^D = &\max_{q(X_l|\mathrm{pa}(X_l)) \in \mathrm{ext} K'(X_l|\mathrm{pa}(X_l))} \\ &\frac{q(E)}{p(E)} \frac{\|p(\mathbf{X}_Q|E), q_l(\mathbf{X}_Q|E)\|}{\|p(X_l|\mathrm{pa}(X_l)), q(X_l|\mathrm{pa}(X_l))\|} . \end{aligned}$$

5 Performing robustness and sensitivity analysis

Techniques for robustness and sensitivity analysis presented in the previous sections can be conducted with the *JavaBayes* system, a publicly available package for operations with Bayesian and Quasi-Bayesian networks. The *JavaBayes* system is composed of a user interface and an inference engine. The inference engine has the ability to produce posterior marginals and expectations; if requested, the inference engine combines conditional credal sets into type-1 extensions and reports upper and lower posterior envelopes or expectations. Sensitivity measures can also be obtained from the posterior values computed by the engine. More information about the *JavaBayes* distribution, available through the GNU-license, can be found in the World-Wide-Web at http://www.cs.cmu.edu/~javabayes.

Consider again, as an example, the network in Figure 1 (Charniak 1991). Suppose the following probability assessments are made:

$$
\begin{array}{ll}
K(F) = [0.45, 0.55], & K(B) = [0.5, 0.7], \\
K(L|F) = [0.55, 0.65], & K(D|F, B) = [0.7, 0.9], \\
K(L|F^c) = [0.02, 0.08], & K(D|F, B^c) = [0.05, 0.15], \\
K(H|D) = [0.5, 0.7], & K(D|F^c, B) = [0.05, 0.15], \\
K(H|D^c) = [0.1, 0.5], & K(D|F^c, B^c) = [0.65, 0.75].
\end{array}
$$

The type-1 extension combining the credal sets in the network has 1024 vertices. Inferences usually involve a smaller number of vertices due to d-separation. For example, the value of $p(L|F)$ does not affect $p(H|F)$, because L and H are d-separated given F. But computation of the lower envelope $\underline{P}(H|F)$ and the upper envelope $\overline{P}(H|F)$ still involves 512 probability measures. *JavaBayes* produces the following posterior envelopes for H:

```
table lower-envelope 0.25 0.365 ;
table upper-envelope 0.635 0.75 ;
```

Sensitivity measures can be computed with much less effort. Suppose that a "base" Bayesian network is created by taking the middle point of all credal sets previously specified. For example, take $p(F)$ to be 0.5, $p(H|D)$ to be 0.6. Suppose also the variable of interest is H (i.e., $\mathbf{X}_Q = \{H\}$) and the evidence is that the variable F is true. Using *JavaBayes*, the sensitivity measures can be easily computed. The sensitivity measures S_L^D and S_F^D are equal to zero, reflecting the fact that the evidence "separates" the probability values from the variable of interest. Variables B and D do affect $p(H|F)$ and display sensitivity measures $S_B^D = 0.21$ and $S_D^D = 0.24$. The computational effort spent in obtaining these measures is negligible for a network of this size; answers are obtained instantly. In general, the sensitivity measures are produced in time that is similar to the time spent in standard inferences, so they can be obtained with reasonable effort in general.

6 Conclusion

Together, the tools for sensitivity and robustness analysis developed in this paper allow the designers of intelligent systems to properly handle imprecision in probability values. The analysis proposed on this paper is based on closed convex sets of probability measures, which have a relatively simple semantics. Also, sets of probability measures provide a unifying language that can smoothly handle global (robustness) and local (sensitivity) analysis in graphical statistical models.

Sensible strategies for robustness and sensitivity analysis should require moderate computational effort. Robustness analysis, due to its global character, is certain to be more expensive than sensitivity analysis, and may be prohibitively expensive in some applications. The main contribution of this paper is the development of the sensitivity measures S_i^E and S_i^E by the adaptation of other procedures in the literature. These quantities provide the necessary tools for an quick assessment of sensitivity, before any comprehensive analysis is conducted in a Bayesian network. As soon as a Bayesian network is properly tuned and analyzed, the statistical model embodied by the network can be easily employed in automation tasks (Jensen 1996).

ACKNOWLEDGEMENTS

The author was partially supported by CNPq through grant 300183/98.

REFERENCES

Andersen, K. A. and J. N. Hooker (1994). Bayesian logic. *Decision Support Systems* **11**, 191–210.

Berger, J. O. (1990). Robust Bayesian analysis: Sensitivity to the prior. *Journal of Statistical Planning and Inference* **25**, 303–328.

Cano, A. and S. Moral (1996). A genetic algorithm to approximate convex sets of probabilities. *Proceedings of the Fifth Internatial Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems* **2**, 859–864.

Cano, A., J. E. Cano and S. Moral (1994). Convex sets of probabilities propagation by simulated annealing. *Proceedings of the Fifth Internatial Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems* pp. 4–8.

Cano, J., M. Delgado and S. Moral (1993). An axiomatic framework for propagating uncertainty in directed acyclic networks. *International Journal of Approximate Reasoning* **8**, 253–280.

Charniak, E. (1991). Bayesian networks without tears. *AI Magazine* pp. 50–63.

Cozman, F. (1997). Robustness analysis of Bayesian networks with local convex sets of distributions. *XIII Uncertainty in Artificial Intelligence Conference* pp. 108–115.

Cozman, F. (1998). Irrelevance and independence in Quasi-Bayesian networks. *XIV Uncertainy in Artificial Intelligence Conference* pp. 89–96.

Giron, F. J. and S. Rios (1980). Quasi-Bayesian behaviour: A more realistic approach to decision making?. In: *Bayesian Statistics* (J. M. Bernardo, J. H. DeGroot, D. V. Lindley and A. F. M. Smith, Eds.). pp. 17–38. University Press. Valencia, Spain.

Gustafson, P. and L. Wasserman (1995). Local sensitivity diagnostics for Bayesian inference. *Annals of Statistics* **23**(6), 2153–2167.

Jensen, F. V. (1996). *An Introduction to Bayesian Networks*. Springer Verlag. New York.

Laskey, K. B. (1995). Sensitivity analysis for probability assessments in Bayesian networks. *IEEE Transactions on Systems, Man, and Cybernetics* **25**(6), 901–909.

Levi, I. (1980). *The Enterprise of Knowledge*. The MIT Press. Cambridge, Massachusetts.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kauffman. San Mateo, CA.

Tessem, B. (1992). Interval probability propagation. *International Journal of Approximate Reasoning* **7**, 95–120.

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall. New York.