A Deterministic Algorithm for Solving Imprecise Decision Problems

Håkan L. Younes
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213, U.S.A.
lorens@cs.cmu.edu

Love Ekenberg
Department of Information Technology
Mid Sweden University
SE-851 70 Sundsvall, SWEDEN
love@ite.mh.se

Abstract

Today there are numerous tools for decision analysis, suitable both for human and artificial decision makers. Most of these tools require the decision maker to provide precise numerical estimates of probabilities and utilities. Furthermore, they lack the capability to handle inconsistency in the decision models, and will fail to deliver an answer unless the formulation of the decision problem is consistent. In this paper we present an algorithm for evaluating imprecise decision problems expressed using belief distributions, that also can handle inconsistency in the model. The same algorithm can be applied to decision models where probabilities and utilities are given as intervals or point values, which gives us a general method for evaluating inconsistent decision models with varying degree of expressiveness.

Introduction

Today there are numerous tools for decision analysis, suitable for both human and artificial decision makers[1] (Younes & Boman 1999). Most of these tools are based on classical decision analysis, thus requiring precise numerical estimates of probabilities and utilities. This has often been considered unrealistic in real-life situations, and a number of models with representations allowing imprecise statements have been suggested (see, e.g., (Good 1962; Smith 1961; Dempster 1967; Shafer 1976; Chen & Hwang 1992; Lai & Hwang 1994)). In this paper we will focus on the theory of belief distributions (Ekenberg & Thorbiörnson 1997; 1998)—in particular the numerical realization of the computational procedures of the theory.

A further problem with current tools is that they require the decision maker to be consistent. In classical models, cf. (Savage 1954), the only constraint is usually that possible outcomes of a probabilistic event must be exhaustive and mutually exclusive—i.e., the sum of the probabilities must equal 1. With the increased expressibility of models allowing imprecise information, it becomes very hard for the decision maker to be consistent. The algorithm we present in this paper does not completely enforce consistency. Instead a measure of the inconsistency in a decision model is used to automatically generate a consistent model that can be properly evaluated. The same inconsistency measure can be used in simpler models as well, where probabilities and utilities are given as intervals or point values, which gives us a general method for evaluating inconsistent decision models with varying degree of expressiveness.

The first section provides the reader with a brief theoretical background, introducing the fundamental concepts of the theory of belief distributions, and explaining how the evaluation of the arising decision models is carried out. In the following section, we describe an algorithm for numerically realizing the computational procedures given by the theory, and in the section after that we describe how the algorithm applies to a range of decision models of varying complexity. We conclude the paper by discussing possible variations of the inconsistency measure, as well as directions of future research.

Theory

In this section we introduce the basic concepts of the theory of belief distributions, necessary for understanding the algorithm presented in this paper. For details, the reader is referred to (Ekenberg & Thorbiörnson 1997).

When faced with a decision problem, the decision maker must first model the problem before it can be analyzed. Consider a decision situation (D) consisting of a set of n alternatives

$$\{\{c_{ij}\}_{j=1,\ldots,m_i}\}_{i=1,\ldots,n}$$

where each alternative is represented by a set C_i of m_i consequences. In classical decision theory, probabilities and utilities of consequences are assigned precise numerical values. The decision model would then contain statements like "the probability of c_{11} is 0.17". This often gives a pretense of accuracy, as decision makers in most realistic situations lack foundation for preferring a certain value (e.g., 0.17) to other values close by.

[1] With an *artificial* decision maker we mean a software agent, or robot, that applies decision theory to the decision problems it is faced with. In the rest of the paper, unless specifically noted, the term decision maker will denote both human and artificial decision makers.

This problem is addressed in *supersoft decision theory*, which allows assessments of probabilities and utilities to be represented by vague and imprecise statements (Malmnäs 1995). These could be *qualitative* statements like "consequence c_{11} is very probable", or *comparative* statements like "consequence c_{11} is at least as desirable as consequence c_{12}". Such linguistic statements can be translated to a numerical format (Malmnäs 1994), where qualitative statements are represented by intervals (e.g., $p_{11} \in [a, b]$), and comparative statements are represented by inequalities (e.g., $u_{11} \geq u_{12}$). Thus, each statement is translated into one or more constraints. Danielson and Ekenberg (1998) investigates methods for evaluating decision models of this type.

All constraints involving probabilities, together with $\sum_{j=1}^{m_i} p_{ij} = 1$ for each set of consequences, form a probability base ($\mathcal{P}$). Similarly, the constraints involving utilities form a value base ($\mathcal{V}$). The solution set of $\mathcal{P}$ is the set of all possible probability distributions over the consequences ($E_{\mathcal{P}}$). However, a decision maker does not necessarily believe equally much in all of the distributions in $E_{\mathcal{P}}$. To allow for a differentiated belief, a distribution can be defined for $E_{\mathcal{P}}$, assigning a belief intensity to each of the probability distributions. The support of the belief distribution—a subset of $E_{\mathcal{P}}$—is then taken to be the set of all possible probability measures for the decision maker.

A similar belief distribution can be defined for the set of all epistemologically possible utility distributions ($E_{\mathcal{V}}$), where $E_{\mathcal{V}}$ is the solution set of $\mathcal{V}$.

Global Belief Distributions

As was mentioned above, a decision maker might not believe with the same intensity in all possible probability or utility measures. For each set C_i of consequences, we define the *cell* for C_i to be the unity cube $[0, 1]^{m_i}$. We further define the cell for a consequence c_{ij} to be the interval $[0, 1]$. Given a cell $B = (b_1, \ldots, b_k)$, a global belief distribution over B is a positive distribution g such that

$$\int_B g(x) dV_B(x) = 1,$$

where V_B is a k-dimensional Lebesgue measure on B.

Note that if the decision maker can specify a global belief distribution g over a cell B, representing a class of probability distributions, then there is no need for specifying an explicit set of constraints $\mathcal{P}$. This is because the support of a global belief distribution g is the solution set to the probability base, i.e., $E_{\mathcal{P}}$ is the support of g. Similarly, $E_{\mathcal{V}}$ is the support of a global belief distribution g' defined over a cell B' representing a class of utility measures.

Local Belief Distributions and Constraints

However, it is seldom possible for a decision maker to specify global belief distributions. Often, the decision maker has access to local information and various relations between different variables only. When this is the case, local belief distributions have to be defined over the cells for each single consequence c_{ij}.

Given a cell $B = (b_1, \ldots, b_k)$, a local belief distribution over B is a positive distribution f_i such that

$$\int_{b_i} f_i(x_i) dV_{b_i}(x_i) = 1,$$

where V_{b_i} is a Lebesgue measure on b_i. In addition, relations between local distributions can be expressed with a set of constraints. In this paper, we will only consider linear constraints of the form

$$\sum_i a_i x_i \ \Re\ b,$$

where $\Re$ is any of the relations $=$, $\leq$, or $\geq$.

Evaluation

The evaluation principle presented here is based on the principle of maximizing the expected utility. Given a decision situation D, a probability base $\mathcal{P}$, and a value base $\mathcal{V}$, the expected utility, denoted $E(C_i)$, is $\sum_{j=1}^{m_i} p_{ij} u_{ij}$, where p_{ij} and u_{ij} are variables in $\mathcal{P}$ and $\mathcal{V}$ respectively.

Ekenberg and Thorbiörnson (1997) suggest how belief distributions can be accounted for in the evaluations of $E(C_i)$. A generalized expected mean value $G(C_i)$ is defined. $G(C_i)$ can be evaluated by the use of centroids.[2] If $x_{\mathbf{p}_i}$ is the centroid of the belief distribution for $(p_{i1}, \ldots, p_{im_i})$, and $x_{\mathbf{u}_i}$ is the centroid of the belief distribution over the corresponding utility variables $(u_{i1}, \ldots, u_{im_i})$, then $G(C_i)$ is $\langle x_{\mathbf{p}_i}, x_{\mathbf{u}_i} \rangle$. Furthermore, given certain restrictions, the centroids of the local distributions can be used in a similar way to evaluate the generalized expected mean (Ekenberg & Thorbiörnson 1997). Thereby, the complexity of the evaluations can be logarithmically reduced.

Inconsistency Measure

The use of local belief distributions and constraints might however introduce a problem. Given local belief distributions $f_{\mathbf{p}_i(j)}$ for the probability variables, and a set of linear constraints $\mathcal{P}$, it may be the case that the centroids $x_{\mathbf{p}_i(j)}$ of the local belief distributions are not consistent with the constraints. Let $x_{\mathbf{p}}$ denote the vector

$$(x_{\mathbf{p}_1(1)}, \ldots, x_{\mathbf{p}_1(m_1)}, x_{\mathbf{p}_2(1)}, \ldots, x_{\mathbf{p}_n(m_n)}).$$

We have a conflict if $x_{\mathbf{p}}$ is not on the polytope defined by $\mathcal{P}$, i.e. there is a positive belief in a vector not consistent with the constraints. This means that the decision maker, to be consistent, has to modify either the belief distributions or the constraints (or both). When an inconsistency occurs, we could still calculate the generalized expected mean value by choosing a vector $x'_{\mathbf{p}}$ consistent with $\mathcal{P}$, but the problem is how this vector

[2] Intuitively, the centroid of a distribution is where the belief mass is concentrated.

should reasonably be selected. A prima facie solution is to let $x'_{\mathbf{p}}$ be the vector, consistent with $\mathcal{P}$ and with the least Euclidean distance to $x_{\mathbf{p}}$. Since the vector $x_{\mathbf{p}}$ represents the center of belief mass, by minimizing the Euclidean distance we express a bias towards vectors consistent with $\mathcal{P}$ that are close to the vector where the belief mass is centered. The rationale for this choice is that, unless the belief distributions are very irregular, these are likely to be the vectors consistent with $\mathcal{P}$ in which the decision maker has the highest belief.

One inconsistency measure implementing this strategy is

$$\frac{1}{2}\|x'_{\mathbf{p}} - x_{\mathbf{p}}\|^2. \tag{1}$$

The vector $x'_{\mathbf{p}} - x_{\mathbf{p}}$ can be used as guidance for a decision maker on how to adjust the model. To lower the inconsistency of the probability base, the decision maker could, for example, modify the belief distributions so that $x_{\mathbf{p}}$ is moved in the direction of this vector.

Equation 1 works only when the set of linear constraints in the probability base is consistent. Furthermore, only the constraints $\sum_{j=1}^{m_i} p_{ij} = 1$ are forced upon the decision maker by the axioms of probability theory. All other constraints are, just as the belief distributions, expressions of the decision maker's beliefs. This suggests that relaxing some of the user specified constraints could be just as appropriate as moving the centroid. We can incorporate this strategy into our inconsistency measure by adding a new positive slack variable ξ_i to each of the soft constraints (cf. (Cortes & Vapnik 1995)).[3] A soft constraint $\sum_i a_i x_i \leq b$ would be substituted by $\sum_i a_i x_i \leq b + \xi_i$, and $\sum_i a_i x_i \geq b$ by $\sum_i a_i x_i \geq b - \xi_i$. A soft equality constraint $\sum_i a_i x_i = b$ would be substituted by the two inequality constraints $\sum_i a_i x_i \leq b + \xi_i$ and $\sum_i a_i x_i \geq b - \xi_{i+1}$. We can now choose to minimize

$$\frac{1}{2}\|x'_{\mathbf{p}} - x_{\mathbf{p}}\|^2 + C\sum_{i=1}^{\ell_p} \xi_{\mathbf{p}(i)} \tag{2}$$

which expresses a bias towards small modifications of the soft constraints, in addition to the bias for small movements of the centroid already expressed by Equation 1. The parameter C allows the decision maker to control the penalty for modifying constraints, a larger C corresponding to assigning a higher penalty to relaxing constraints.

Similarly,

$$\frac{1}{2}\|x'_{\mathbf{u}} - x_{\mathbf{u}}\|^2 + C\sum_{i=1}^{\ell_u} \xi_{\mathbf{u}(i)} \tag{3}$$

can be used as inconsistency measure for the value base, with the vector $x'_{\mathbf{u}} - x_{\mathbf{u}}$, as well as the values of the slack variables $\xi_{\mathbf{u}(i)}$, functioning as guidance for the decision maker in trying to lower the inconsistency of the value base.

[3] The soft constraints are here taken to be all constraints except $\sum_{j=1}^{m_i} p_{ij} = 1$, but the decision maker could be allowed to specify whether any further constraints should be hard.

Algorithm

Next, an algorithm for numerically realizing the computational procedures given in the previous section, is provided. In this section, we present the general algorithm for computing generalized expected mean values given a decision problem involving belief distributions and linear constraints. The next section demonstrates how the algorithm also applies to decision problems involving interval values and constraints only, as well as to classical models using point values. As will be noted, the algorithm can be much simplified in these latter cases. The general algorithm consists of the following steps:

1. Calculate the centroids of the belief distributions.
2. Find vectors, consistent with the constraints, that minimize the inconsistency measures given in Equations 2 and 3.
3. Calculate the generalized expected mean value for each consequence.

The last step is just a straightforward calculation of $\langle x'_{\mathbf{p}_i}, x'_{\mathbf{u}_i}\rangle$ and will not be discussed below. The first two steps, however, call for further elaboration.

Calculating Centroids

Calculating the centroids of the belief distributions involves numerical integration. Although there are few restrictions on the belief distributions, they will typically be smooth and continuous functions. Thus, if belief distributions are defined locally for only one probability or utility variable, then any standard algorithm for numerical integration of single variable functions can be used (see, e.g., (Davis & Rabinowitz 1984; Zwillinger 1992)). One of the simplest being the trapezoidal rule, which gives high accuracy after only a few iterations when the integrand is smooth and continuous. For more irregular and only piecewise continuous belief distributions, more sophisticated methods, such as integration by parts and adaptive integration are required (Davis & Rabinowitz 1984).

With belief distributions defined over more than one variable, methods for integration of multivariate functions are needed (see, e.g., (Davis & Rabinowitz 1984; Sloan & Joe 1994)), increasing the complexity of the algorithm substantially. On the other hand, with belief distributions defined over several variables, there will usually be fewer explicit constraints.

Minimizing Inconsistency

As was noted above, centroids may be inconsistent with the stated constraints when the specification of the decision problem involves local distributions. It may also be the case that the set of linear constraints lacks feasible solutions. In such cases, we find the vectors minimizing the inconsistency measure given in Equation 2 for the probability base and Equation 3 for the value base respectively, and use these instead of the centroids when calculating the generalized mean values. These vectors are found by solving the convex quadratic programming problems with the inconsistency measures as objective functions. Because the problems are convex, this is relatively easy, and we are guaranteed to find global minimizers (Fletcher 1987; Luenberger 1989).

Special Cases

Above, we have presented an algorithm for evaluating decision problems with belief distributions defined for the probability and value bases. It should also be noted that the algorithm can be used to evaluate decision models with interval values but no belief distributions, and classical models with probabilities and utilities given as point values.

Interval Values

Danielson and Ekenberg (1998) investigates different methods for evaluating decision models with interval probabilities and utilities. Here, we suggest an alternative approach. The algorithm presented in the previous section requires belief distributions in order to compute the vectors used for evaluating the generalized expected mean values. Now, we have no belief distributions, so how can we use the same algorithm? Simply by assuming a symmetric belief distribution for the intervals. The coordinates of the centroids then become the midpoints of the intervals. The complexity of the algorithm is reduced since the integral computations become trivial.

Point Values

Considering point values, we have the constraints $\sum_{j=1}^{m_i} p_{ij} = 1$, so we need to consider inconsistency in the probability assessments only. If we interpret the point values as being the only points with positive belief, we can use the same approach as in our original algorithm. The computational complexity, however, is substantially reduced. The vector $x'_{\mathbf{p}_i}$ is now given by the equation

$$x'_{\mathbf{p}_i} = x_{\mathbf{p}_i} - \frac{\langle \mathbf{n}, (x_{\mathbf{p}_i} - \mathbf{v}) \rangle}{\|\mathbf{n}\|^2}\mathbf{n},$$

where $\mathbf{v}$ is an arbitrary vector on the hyperplane defined by $\sum_{j=1}^{m_i} p_{ij} = 1$, and $\mathbf{n}$ is a normal of that hyperplane. Since $\mathbf{n} = k\mathbf{1}$ for some $k \neq 0$ where $\mathbf{1}$ is a vector of size m_i with all elements equal to 1, we get:

$$\begin{aligned}
-\frac{\langle \mathbf{n}, (x_{\mathbf{p}_i} - \mathbf{v}) \rangle}{\|\mathbf{n}\|^2}\mathbf{n} &= \frac{\langle k\mathbf{1}, (\mathbf{v} - x_{\mathbf{p}_i}) \rangle}{\|k\mathbf{1}\|^2}k\mathbf{1} \\
&= \frac{\langle \mathbf{1}, (\mathbf{v} - x_{\mathbf{p}_i}) \rangle}{m_i}\mathbf{1} \\
&= \frac{\sum_{j=1}^{m_i}(v_j - x_{\mathbf{p}_i(j)})}{m_i}\mathbf{1} \\
&= \frac{\sum_{j=1}^{m_i} v_j - \sum_{j=1}^{m_i} x_{\mathbf{p}_i(j)}}{m_i}\mathbf{1} \\
&= \frac{1 - \sum_{j=1}^{m_i} x_{\mathbf{p}_i(j)}}{m_i}\mathbf{1}.
\end{aligned}$$

From this follows that the jth element of $x'_{\mathbf{p}_i}$ simply is

$$x'_{\mathbf{p}_i(j)} = x_{\mathbf{p}_i(j)} + \frac{1 - \sum_{j=1}^{m_i} x_{\mathbf{p}_i(j)}}{m_i}.$$

Discussion

We have presented an algorithm that implements the computational procedures for evaluating imprecise decision problems suggested by Ekenberg and Thorbiörnson (1997; 1998). There are, however, many important issues that are not addressed in this paper. One is how decision makers can benefit from the increased expressibility. Developing an intuitive semantics for belief distributions, similar to that already existing for interval probabilities and utilities, is essential. Without it, the decision maker would have little guidance for what constitutes a good belief distribution, which in particular would make it hard for human decision makers to fully make use of the added expressiveness. For artificial decision makers, procedures for learning and updating belief distributions from observations must be developed. For example, ignorance can very naturally be represented by a uniformly distributed belief for the whole domain of a variable, but once evidence for that variable becomes available, it is not as obvious how this should be incorporated into the model.

Furthermore, when evaluating the decision problem, probabilities and utilities are reduced to single points by means of integration. This makes the computations less complex, but much of the information in the problem specification is lost. A topic for future research is to develop algorithms for sensitivity analysis. Such an algorithm could try to identify assignments to the probability and utility variables in which the decision maker has a high belief, but that would result in a different alternative being the one with highest expected utility.

Finally, our bias towards vectors close to the centroids of the belief distributions does not necessarily give us the vectors in which the decision maker has the highest belief. Alternatively, we could choose an inconsistency measure expressing a bias towards vectors with a high belief. One such measure for the probability base could be

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i}(f_{\mathbf{p}_i(j)}(x_{\mathbf{p}_i(j)}) - f_{\mathbf{p}_i(j)}(x'_{\mathbf{p}_i(j)})).$$

This will, however, make the efficiency of the algorithm dependent on the decision maker's choices of belief distributions. Unless these are all linear functions (which in most cases will not be the case), finding $x'_{\mathbf{p}_i(j)}$ and $x'_{\mathbf{u}_i(j)}$ will amount to solving a nonlinear programming problem, which in the general case is hard (see, e.g., (Luenberger 1989)). Finding a good bias by evaluating the current inconsistency measure, and alternative measures, is an important task for future research.

Acknowledgments. Most of the work described in this document was done while the first author was working at the Department of Computer and Systems Sciences (DSV) at KTH in Sweden. The authors are grateful to Magnus Boman at DSV for making this work possible, and to the NUTEK Promodis programme for additional support during that time. Work performed at CMU was sponsored in part by grants from the Swedish Fulbright Commission and the Foundation BLANCEFLOR Boncompagni-Ludovisi, née Bildt. Ekenberg gratefully acknowledge the support from the program *Fiber Science and Communication Network*. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies or endorsements, either expressed or implied, of the sponsors.

References

Chen, S.-J., and Hwang, C.-L. 1992. *Fuzzy Multiple Attribute Decision Making*, volume 375 of *Lecture Notes in Economics and Mathematical Systems*. Berlin: Springer-Verlag.

Cortes, C., and Vapnik, V. 1995. Support-vector networks. *Machine Learning* 20:273–297.

Danielson, M., and Ekenberg, L. 1998. A framework for analysing decisions under risk. *European Journal of Operational Research* 104:474–484.

Davis, P. J., and Rabinowitz, P. 1984. *Methods of Numerical Integration*. Orlando, Florida: Academic Press, second edition.

Dempster, A. P. 1967. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics* 38(2):325–339.

Ekenberg, L., and Thorbiörnson, J. 1997. Modeling indeterminacy in decision situations. Interim Report IR-97-044, IIASA, Laxenburg, Austria.

Ekenberg, L., and Thorbiörnson, J. 1998. Globalisation of belief distributions. In Cook, D. J., ed., *Proceedings of the Eleventh International Florida Artificial Intelligence Research Symposium Conference*. Menlo Park, CA: The AAAI Press.

Fletcher, R. 1987. *Practical Methods of Optimization*. New York: John Wiley & Sons, second edition.

Good, I. J. 1962. Subjective probability as the measure of a non-measureable set. In Suppes; Nagel; and Tarski., eds., *Logic, Methodology and Philosophy of Science (Proc. 1960 Internat. Congr.)*. Stanford, CA: Stanford University Press. 319–329.

Lai, Y.-J., and Hwang, C. L. 1994. *Fuzzy Multiple Objective Decision Making*, volume 404 of *Lecture Notes in Economics and Mathematical Systems*. Berlin: Springer-Verlag.

Luenberger, D. G. 1989. *Linear and Nonlinear Programming*. Reading, Mass.: Addison-Wesley, second edition.

Malmnäs, P.-E. 1994. Towards a mechanization of real life decisions. In Prawitz, D., and Westerståhl, D., eds., *Logic and Philosophy of Science in Uppsala*. Kluwer Academic Publishers.

Malmnäs, P.-E. 1995. Methods of evaluation in supersoft decision theory. Unpublished manuscript, available on the WWW using URL `http://www.dsv.su.se/DECIDE`.

Savage, L. J. 1954. *Foundations of Statistics*. New York: John Wiley & Sons.

Shafer, G. 1976. *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton University Press.

Sloan, I., and Joe, S. 1994. *Lattice Methods for Multiple Integration*. Oxford: Clarendon Press.

Smith, C. A. B. 1961. Consistency in statistical inference and decision. *Journal of the Royal Statistics Society* B 23:1–37.

Younes, H. L., and Boman, M. 1999. Tools for artificial decision making. MAAMAW'99 Poster Presentations, published on CD-ROM, Universidad Politècnica de Valencia.

Zwillinger, D. 1992. *Handbook of Integration*. Boston: Jones and Bartlett Publishers.