Globalisation of Belief Distributions

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

Johan Thorbiörnson
Department of Physics and Mathematics
Mid Sweden University
SE 851 70 Sundsvall, SWEDEN
email: johant@fmi.mh.se

Abstract

We present a theory for evaluating decisions under risk when the available information is indeterminate. The probability and utility estimates involved in such a situation is expressed as sets of distributions, representing beliefs in various vectors in the decision space. We also demonstrates some consistency requirements between beliefs in local values, i.e. vectors representing values for a singleton probability or utility variable, and beliefs in global values, i.e. beliefs in vectors in the decision space. The evaluation of the different strategies is performed with respect to a generalisation of the principle of maximising the expected utility. We show that, despite the possible complexity of the various inputs, the computational efforts for evaluating strategies are tractable.

Background

A number of models with representations allowing imprecise statements have been suggested. Some of them use standard probability theory while others contain some specialised formalism, c.f. [Choquet, 1953/54], [Huber, 1973, Huber and Strassen, 1973], [Good, 1962, Smith, 1961], [Dempster, 1967]. One particular modelling approach is fuzzy set theory for handling the vagueness in subjective estimates of values and probabilities. As argued in [Ekenberg and Thorbiörnsson, 1997], theories based on this will have several consistency problems as soon as linear constraints are involved. This is the case already when representing probabilities, so the fuzzy set theory is less suitable for theories based on classical decision analysis. Fuzzy approaches are also restricted in the sense that they do not handle qualitative aspects such as, e.g., comparisons between different components involved in decision situations. Other approaches are investigated in [Gördenfors and Sahlin, 1982, 1983, Levi, 1974, 1980]. The authors take a more global viewpoint of beliefs, but nevertheless restrict themselves to the probability case and, like the fuzzy models, interval representations. Another limitation is that they neither investigate the relation between global and local distributions nor introduce methods for determining the consistency of user-asserted sentences. The work by [Danielson and Ekenberg, 1997, Ekenberg, et. al., 1996, 1997, Malmnös, 1994] is restricted in the sense that no distributions over the intervals are taken into account.

The next sections describe a representation model for impreciseness, and discuss some general properties of global and local belief distributions. In particular, it is described how global belief distributions can be defined and in what sense such distributions can define solution sets to a set of constraints, and how classes of admissible local belief distributions can be derived from projections of such global distributions. Thereafter we show how sets of belief distributions can be evaluated with respect to a generalisation of the expected utility.

Representation

The motivation behind the present work is to extend the expressibility when representing and evaluating vague and numerically imprecise information in decisions situations. To achieve a basic intuition of what will be presented below, consider a decision situation 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 of m_i consequences. We will refer to the latter as a consequence set. In such a decision situation, numerically imprecise sentences like "the probability of consequence c_{11} is greater than 40%" or comparative sentences like "consequence c_{11} is preferred to consequence c_{12}" occur. These sentences can be represented in a numerical format [Danielson and Ekenberg, 1997]. Examples of vague sentences in that model are: "The consequence c_{ij} is probable" or "The event c_{ij} or c_{ik} is possible". Such sentences are represented by suitable intervals. Another kind of sentences are interval sentences of the form: "The probability of c_{ij} lies between the numbers a_k and b_k", which are translated to $p_{ij} \in [a_k, b_k]$. Finally, comparative sentences are of the form: "The probability of c_{ij} is greater than the probability of c_{kl}". Such a sentence is translated into an inequality $p_{ij} \geq p_{kl}$. Each statement is thus represented by one or more constraints. The conjunction of constraints of the types above, together with $\sum_{j=1}^{m_i} p_{ij} = 1$ for each consequence set $\{c_{ij}\}_{j=1,\ldots,m_i}$, involved, is a probability base $\mathcal{P}$. A value base $\mathcal{V}$ consists of similar translations of vague and numerically imprecise value estimates. In a sense, a probability base can be interpreted as constraints defining the set of all possible probability measures.

However, a decision maker does not necessarily believe with the same intensity in all the epistemologically possible probability distributions E. To enable a refinement of the model to allow for a differentiation of distributions in this respect, a global distribution expressing various beliefs can be defined over the set E. In the following subsections, we define and investigate some features of global and local distributions and how these are related to sets of linear constraints.

Global Belief Distributions

The basic entities in the kinds of decision situations we will consider are the sets of consequences involved. Over these sets, different functions can be defined, expressing, for instance, classes of probability or value measures. To enable for a differentiation of functions and to take constraints into account, a global distribution expressing various beliefs can be defined over a multi-dimensional unity cube. Each dimension corresponds to a consequence.

Definition 1 *Let a unity cube $B = (b_1, \ldots, b_k)$ be given. By a global belief distribution over B, we mean a positive distribution*[1] *g defined on the unity cube B such that*

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

where V_B is some k–dimensional Lebesgue measure on B. The set of all global belief distributions over B is denoted by GBD (B).

Global belief distributions can be used to represent subsets of a unity cube by considering the support of the distributions. However, if we want to represent a subset which is of lower dimension than the unity cube itself we cannot use distributions that are upper bounded since a mass under such a distribution will be 0 while integrating with respect to some Lebesgue measure defined on the unity cube. This problem is solved in detail in [Ekenberg and Thorbiörnson, 1997], and will not be treated here. We will be content to demonstrate the general idea by the example below.

Example 1: Let a unity cube $B = (b_1, b_2, b_3)$ be given. Let A denote the subset of B, where $x_1 + x_2 + x_3 = 1$, and let $f(x_1, x_2) = x_1 \cdot x_2$ be defined on A. Then $f \in$ GBD (A) with respect to the 2–dimensional Lebesgue measure on A, but $\bar{f}_A(x) \notin$ GBD (B), because $\int_B \bar{f}_A(x)\, dV_B(x) = 0$. However, $\int_B \bar{f}_A(x) g_A(x)\, dV_B(x) = 1$, so $\bar{f}_A(x) g_A(x) \in$ GBD (B), and $\bar{f}_A(x_1, x_2, x_3) g_A(x_1, x_2, x_3) = 0$, except when $x_1 + x_2 + x_3 = 1$. To put this informally; $\bar{f}_A \cdot g_A$ represents the same proportional belief over B as f does over A.

[1] A distribution on a set Ω is a linear functional defined on $C_0^\infty(\Omega)$ which is continuous with respect to a certain topology.

Linear Constraints

It should be noted that one property of a global belief distribution is that it in some sense defines the solution set to a set of constraints.

Definition 2 *Let a unity cube $B = (b_1, \ldots, b_k)$ be given. We will use the term constraints for the union of the following:*

- *I–constraints are constraints on the form $a \geq x_i$ or $a \leq x_i$, where a is a real number in $[0, 1]$, and x_i is a variable.*
- *L–constraints are constraints on the form $\sum x_i = a$, where a is a real number in $[0, 1]$.*
- *C–constraints are constraints on the form $x_i \leq x_j + a$, where a is a real number in $[0, 1]$.*

Definition 3 *Let a unity cube $B = (b_1, \ldots, b_k)$ and a distribution g over B be given. The support of g (supp g) is the closure of the set $\{(x_1, \ldots, x_k) : g(x_1, \ldots, x_k) > 0\}$.*

Derived Belief Distributions

This section introduces local distributions and briefly discusses consistency of user-asserted sentences.

Definition 4 *Let a unity cube $B = (b_1, \ldots, b_k)$ be given. By a local belief distribution over B, we mean a positive distribution f defined on the unity cube b_i such that*

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

where V_{b_i} is some Lebesgue measure on b_i. The set of all local belief distributions over b_i is denoted by LBD (b_i).

Local belief distributions over the axes of a unity cube B can be derived from a global belief distribution over B.

Definition 5 *Let a unity cube $B = (b_1, \ldots, b_k)$ and $F \in$ GBD (B) be given. Let*

$$f_i(x_i) = \int_{B_i^-} F(x)\, dV_{B_i^-}(x)$$

where $B_i^- = (b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_k)$. We say that $f_i(x_i)$ is derived from F.[2]

[2] In the following, we use x to denote the vector $(x_1, \ldots, x_k)$.

In the section *Evaluation* we show how the use of centroids logarithmically reduces the computational complexity in the evaluations of a generalised expected utility. Intuitively, the centroid of a distribution is a point in space where some of the geometrical properties of the distribution can be regarded as concentrated.

Definition 6 *Let a unity cube $B = (b_1, \ldots, b_k)$ and $g_B \in \text{GBD}(B)$ be given. The centroid of g_B is the point $x_{g_B} = (\beta_1, \ldots, \beta_k)$ in B whose i:th component is*

$$\beta_i = \int_B x_i \cdot g_B(x)\, dV_B(x).$$

Definition 7 *Let a unity cube $B = (b_1, \ldots, b_k)$ and $f_{b_i} \in \text{LBD}(b_i)$ be given. The centroid of f_{b_i} is the point in b_i defined by*

$$x_{f_{b_i}} = \int_{b_i} x_i \cdot f_{b_i}(x_i)\, dV_{b_i}(x_i).$$

Centroids are invariant under projections on the local unity cubes in the sense that the projection of the centroid of the global belief distribution on the local unity cube has the same coordinates as the centroids of the corresponding derived local belief distributions.

Lemma 1 *Let a unity cube $B = (b_1, \ldots, b_k)$ and $F \in \text{GBD}(B)$ be given. Let $f_i(x_i)$ be derived from F. Furthermore, let*

$$G(x_1, \ldots, x_k) = f_1(x_1) \cdot \ldots \cdot f_k(x_k).$$

Then

(i) $f_i(x_i) \in \text{LBD}(b_i),\ i = 1, \ldots, k$.

(ii) $G \in \text{GBD}(B)$[3]

(iii) $x_G = x_F$

(iv) If $x_G = (\alpha_1, \ldots, \alpha_k)$ *then* $\alpha_i = x_{f_{b_i}}$.

(v) $f_i(x_i)$ *is derived from G for all* $i = 1, \ldots, k$.

Relations Between Constraints and Belief Distributions

Of particular interest is to what extent local belief distributions can combine to a global belief distribution, so that the global distribution in some sense represents the local belief distributions as well as a set of constraints imposed on the decision situation.

Definition 8 *Let a unity cube $B = (b_1, \ldots, b_k)$ and a consistent set C of constraints in B be given. The global belief distribution F is called C-admissible iff*

x is a solution vector to C iff $x \in \text{supp}\, F$.

Usually a decision maker has access only to local information and a set of relations between different parameters and, consequently, has no explicit idea about the global distribution. In many cases, it may be that the only accessible relations between the local distributions are in terms of constraints.

[3]In general, measure properties defined locally are not necessarily preserved globally, cf. [Thorbiörnson, 1996].

If the decision maker is able to define a set of local belief distributions and a set of constraints describing the decision problem, these must be congruent in a certain respect. Given a set of constraints, a decision maker is restricted concerning which combinations of local belief distributions that are possible to impose, if she wants to be consistent in a reasonable sense. This is expressed by the following definition.

Definition 9 *Let a unity cube $B = (b_1, \ldots, b_k)$ and a consistent set C of constraints in B be given. A set $L = \{f_i(x_i) \in \text{LBD}(b_i)\}_{i=1,\ldots,k}$ is called C-admissible iff the vector $(x_{f_1}, \ldots, x_{f_k})$ is a solution vector to C, where x_{f_i} denotes the centroid of f_i.*

From Lemma 1, we can derive the following theorem.

Theorem 1 *Let a unity cube $B = (b_1, \ldots, b_k)$ and a consistent set C of constraints, such that $s(C) \subseteq B$ be given. Let G be a C-admissible global distribution and let $g_i(b_i), i = 1, \ldots k$, be derived from G. Then $\{g_i(x_i)\}_{i=1,\ldots,k}$ is a C-admissible set of local belief distributions.*

Theorem 1 implies that if a decision maker defines a set of local belief distributions describing a problem, and if these are admissible w.r.t. the constraints involved, a global belief distribution can be determined. This distribution has the property of having the same centroid (and the same support relative to the local belief distributions) as any global belief distribution from which the user-asserted local belief distributions can be derived.

Evaluation

The evaluation principle treated in this section is based on the principle of maximising the expected value. Given a decision situation D, let $\mathcal{P}$ and $\mathcal{V}$ be a probability base and a value base for D, respectively. The expected value $E(C_i)$ denotes the expression $\sum_{j=1}^{m_i} p_{ij} v_{ij}$, where p_{ij} and v_{ij} are variables in $\mathcal{P}$ and $\mathcal{V}$. To evaluate the expected value [Danielson and Ekenberg, 1997, Ekenberg, et al., 1996, Malmnös, 1994] investigates the set $\{E(C_i)\}_{i=1,\ldots,n} \cup \mathcal{P} \cup \mathcal{V}$ in a variety of respects. However, in the present framework, distributions are included, and we will suggest how these can be taken into account in the evaluations.

Generalised Mean Values

We will now describe how belief functions can be imposed on evaluations of a generalised expected mean value.

First, we a define a decision scenario as containing a number of global belief distributions. Informally, these express various beliefs in vectors in subsets of the solution sets to probability-, and utility bases. For a given consequence set, there is one global belief distribution for the probabilities, and one global belief distribution for the utilities, with respect to this set.

Definition 10 *A decision scenario is a structure* $(D, P, V, \{\mathbf{p}_i\}_{i=1,\ldots,n}, \{\mathbf{v}_i\}_{i=1,\ldots,n})$, *where*

- *D is a decision situation* $\{\{c_{ij}\}_{j=1,\ldots,m_i}\}_{i=1,\ldots,n}$.
- $P = (p_{11}, p_{12}, \ldots, p_{mn})$ *is a unity cube.*
- $V = (v_{11}, v_{12}, \ldots, v_{mn})$ *is a unity cube.*
- $\mathbf{p}_i$ *is a global belief distribution over the unity cube* $P_i = (p_{i1}, \ldots, p_{im_i})$ *such that* $\mathbf{p}_i(x) = 0$, *when* $\sum_{j=1}^{m_i} p_{ij} \neq 1$.
- $\mathbf{v}_i$ *is a global belief distribution over the unity cube* $V_i = (v_{i1}, \ldots, v_{im_i})$.

The next definition suggests a generalised expected value. This is summed over all possible expected values weighted by the global belief distributions over the solution sets to the probability-, and utility bases.

Definition 11 *Let a decision scenario* $(\{C_i\}_{i=1,\ldots,n}, P, V, \{\mathbf{p}_i\}_{i=1,\ldots,n}, \{\mathbf{v}_i\}_{i=1,\ldots,n})$ *be given. The expression*

$$\int_{P_i \times V_i} \left(\sum_{j=1}^{m_i} x_{ij} y_{ij} \right) \mathbf{p}_i(x_{i1}, \ldots, x_{im_i}) \cdot \mathbf{v}_i(y_{i1}, \ldots, y_{im_i}) \, dV(x_{i1}, \ldots, x_{im_i}) dV(y_{i1}, \ldots, y_{im_i})$$

is called the generalised expected value for C_i, *and is denoted by* $G(C_i)$.

The next theorem shows how the generalised expected value of a consequence set can be calculated by using only the centroids of the global belief distributions.

Theorem 2 *Let a decision scenario* $(D, P, V, \{\mathbf{p}_i\}_{i=1,\ldots,n}, \{\mathbf{v}_i\}_{i=1,\ldots,n})$ *be given. Then*

$$G(C_i) = \langle x\mathbf{p}_i, x\mathbf{v}_i \rangle,$$

where $\langle x, y \rangle$ *is the standard inner product of* x *and* y.

Proof: Let $x\mathbf{p}_i = (\alpha_{i1}, \ldots, \alpha_{im_i})$ in P_i and let $x\mathbf{v}_i = (\beta_{i1}, \ldots, \beta_{im_i})$ in V_i. Since $\mathbf{p}_i \in \mathrm{GBD}(P_i)$, $\mathbf{v}_i \in \mathrm{GBD}(V_i)$, then by definition 6

$$\int_{P_i} x_{ij} \cdot \mathbf{p}_i \, dV(x_{i1}) \cdots dV(x_{im_i}) = \alpha_{ij}.$$

Analogously we get

$$\int_{V_i} y_{ij} \cdot \mathbf{v}_i \, dV(y_{i1}) \cdots dV(y_{im_i}) = \beta_{ij}.$$

Thus, by the independence of P_i and V_i we get

$$\int_{P_i \times V_i} x_{ij} \cdot y_{ij} \mathbf{p}_i(x_{i1}, \ldots, x_{im_i}) \cdot \mathbf{v}_i(y_{i1}, \ldots, y_{im_i}) \, dV(x_{i1}, \ldots, x_{im_i}) dV(y_{i1}, \ldots, y_{im_i}) = \alpha_{ij} \cdot \beta_{ij}.$$

Thus

$$G(C_i) = \alpha_{i1} \cdot \beta_{i1} + \ldots + \alpha_{im_i} \cdot \beta_{im_i} = \langle x\mathbf{p}_i, x\mathbf{v}_i \rangle.$$

However, as mentioned above, usually a decision maker does only have access to local information and a set of relations between different parameters, and has no explicit idea about the global distributions. If the decision maker is able to define a set of local belief distributions describing the decision problem, and if these are congruent with the constraints involved, the general expected value can be determined. Theorem 3 shows that this is equal to the generalised expected value of C_i involving any global belief distribution, which have a positive support on the solutions sets to $\mathcal{P}$ and $\mathcal{V}$ only, and from which the local belief distributions is derived.

Definition 12 *A potential decision scenario is a structure*

$$(D, P, V, \{\{f_{p_{ij}}\}_{j=1,\ldots,m_i}\}_{i=1,\ldots,n}, \{\{f_{v_{ij}}\}_{j=1,\ldots,m_i}\}_{i=1,\ldots,n}, C),$$

where

- *D is a decision situation* $\{\{c_{ij}\}_{j=1,\ldots,m_i}\}_{i=1,\ldots,n}$.
- $P = (p_{11}, p_{12}, \ldots, p_{mn})$ *is a unity cube.*
- $V = (v_{11}, v_{12}, \ldots, v_{mn})$ *is a unity cube.*
- $f_{p_{ij}} \in \mathrm{LBD}(p_{ij})$.
- $f_{v_{ij}} \in \mathrm{LBD}(v_{ij})$.
- *C is the sets* $\{C_{P_i}\}_{i=1,\ldots,n}$ *and* $\{C_{V_i}\}_{i=1,\ldots,n}$, *where* C_{P_i} *is a set of constraints in the* p_{ij} *variables, and* C_{V_i} *is a set of constraints in the* v_{ij} *variables.*

Theorem 3 *Let a potential decision scenario*

$$(D, P, V, \{\{f_{p_{ij}}\}_{j=1,\ldots,m_i}\}_{i=1,\ldots,n}, \{\{f_{v_{ij}}\}_{j=1,\ldots,m_i}\}_{i=1,\ldots,n}, C)$$

be given. If $p_i \in \mathrm{GBD}(P_i)$ *is* C_{P_i}*-admissible,* $v_i \in \mathrm{GBD}(V_i)$ *is* C_{V_i}*-admissible, and* $f_{p_{ij}}$ *and* $f_{v_{ij}}$ *are derived from* p_i *and* v_i *respectively. Then* $\{f_{p_{ij}}\}_{j=1,\ldots,m_i}$ *is* C_{P_i} *admissible,* $\{f_{v_{ij}}\}_{j=1,\ldots,m_i}$ *is* C_{V_i}*-admissible, and*

$$\langle x_{p_i}, x_{v_i} \rangle = \int_{P_i \times V_i} \left(\sum_{i=1}^{m_i} x_{ij} y_{ij} \right) \cdot f_{p_{i1}}(x_{i1}) \cdots f_{p_{im_i}}(x_{im_i}) f_{v_{i1}}(y_{i1}) \cdots f_{v_{im_i}}(y_{im_i}) \, dV(x_{i1}, \ldots, x_{im_i}) dV(y_{i1}, \ldots, y_{im_i}) = \langle (x_{f_{p_{i1}}}, \ldots, x_{f_{p_{im_i}}}), (y_{f_{v_{i1}}}, \ldots, y_{f_{v_{im_i}}}) \rangle,$$

where $\langle x, y \rangle$ *is the standard inner product of* x *and* y.

Proof: The first part is a direct consequence of theorem 1. The second part follows from Lemma 1.

References

Choquet, G. 1953/54. "Theory of Capacities," Ann. Inst. Fourier, vol. 5, pp. 131-295.

Danielson, M., and Ekenberg, L. 1997. "A Framework for Analysing Decisions under Risk," European Journal of Operational Research, vol. 104/3.

Dempster, A. P. 1967. "Upper and Lower Probabilities Induced by a Multivalued Mapping," Annals of Mathematical Statistics, vol. xxxviii, pp. 325- 339.

Ekenberg, L., Danielson, M., and Boman, M. 1996. "From Local Assessments to Global Rationality," International Journal of Intelligent and Cooperative Information Systems, vol. 5, nos. 2 & 3, pp. 315-331.

Ekenberg, L., Danielson, M., and Boman, M. 1997. "Imposing Security Constraints on Agent-Based Decision Support," Decision Support Systems 20, pp. 3-15.

Ekenberg, L., Oberoi, S., and Orci, I. 1995. "A Cost Model for Managing Information Security Hazards," Computers & Security, vol. 14, pp. 707- 717.

Ekenberg, L., and Thorbiörnson, J. 1997. "Modelling Indeterminacy in Decision Situations," IIASA report series IR-97-044.

Good, I. J. 1962. "Subjective Probability as the Measure of a Non- measurable Set," in Logic, Methodology, and the Philosophy of Science, Suppes, Nagel, and Tarski, Eds.: Stanford University Press, pp. 319-329.

Gördenfors, P., and Sahlin, N.-E. 1982. "Unreliable Probabilities, Risk Taking, and Decision Making," Synthese, vol. 53, pp. 361-386.

Gördenfors, P., and Sahlin, N.-E. 1983. "Decision Making with Unreliable Probabilities," British Journal of Mathematical and Statistical Psychology, vol. 36, pp. 240-251.

Huber, P. J. 1973. "The Case of Choquet Capacities in Statistics," Bulletin of the International Statistical Institute, vol. 45, pp. 181-188.

Huber, P. J., and Strassen, V. 1973. "Minimax Tests and the Neyman-Pearsons Lemma for Capacities," Annals of Statistics, vol. 1, pp. 251-263.

Levi, I. 1974. "On Indeterminate Probabilities," The Journal of Philosophy, vol. 71, pp. 391-418.

Levi, I. 1980. The Enterprise of Knowledge: MIT Press.

Malmnös, P. E. 1994. "Towards a Mechanization of Real Life Decisions," in Logic and Philosophy of Science in Uppsala, Prawitz and Westerstähl, Eds.: Kluwer Academic Publishers.

Smith, C. A. B. 1961. "Consistency in Statistical Inference and Decision," Journal of the Royal Statistic Society, Series B, vol. xxiii, pp. 1-25.

Thorbiörnson, J., and Kolodziej, S. 1996. "Separately Harmonic and Subharmonic Functions," Potential Analysis, vol. 5, No 5.

Thorbiörnson, J. 1998. "Constructions of Extremal Plurisubharmonic Functions for Reinhardt Domains," accepted for publication in Michigan Mathematical Journal as "Counterexamples to Strong Subadditivity in Pluripotential Theory".