Structure Information in Decision Trees and Similar Formalisms

Mats Danielson and **Love Ekenberg**
Dept. of Computer and Systems Sciences,
Stockholm University and KTH,
Forum 100, SE-164 40 Kista, SWEDEN

and **David Sundgren**
University of Gävle,
SE-801 76 Gävle, SWEDEN

Abstract

In attempting to address real-life decision problems, where uncertainty about input data prevails, some kind of representation of imprecise information is important and several have been proposed over the years. In particular, first-order representations of imprecision, such as sets of probability measures, upper and lower probabilities, and interval probabilities and utilities of various kinds, have been suggested for enabling a better representation of the input sentences. A common problem is, however, that pure interval analyses in many cases cannot discriminate sufficiently between the various strategies under consideration, which, needless to say, is a substantial problem in real-life decision making in agents as well as decision support tools. This is one reason prohibiting a more wide-spread use. In this article we demonstrate that in many situations, the discrimination can be made much clearer by using information inherent in the decision structure. It is discussed using second-order probabilities which, even when they are implicit, add information when handling aggregations of imprecise representations, as is the case in decision trees and probabilistic networks. The important conclusion is that since structure carries information, the structure of the decision problem influences evaluations of all interval representations and is quantifiable.

Introduction

Decision analysis and evaluation are of interest for various reasons, ranging from decision modules for rational software agents and other autonomous entities to decision support for human decision-makers. When the behavior of an agent in its environment is not forecasted in detail beforehand, i.e. not designed at compile-time, there is a need for dynamic reasoning and decision making. While reasoning leads to the establishment of facts and figures, decision analysis aids in making decisions based on established facts. In this paper, we suggest a decision model allowing various entities to submit input on a given format and enabling discrimination of decision alternatives based on structure and belief. Thus, the generic term 'decision-maker' below can be a software agent or an autonomous entity as well as a human being.

There have been two important paradigms in modern decision analysis. The first can be said to have begun with (von Neumann and Morgenstern 1944), (Savage 1972), and others as they introduced a structured and formal approach to decision making. The second paradigm emerged as different authors, for example (Dempster 1967), (Ellsberg 1961), (Good 1980), and (Smith 1961) among others, expanded the classic theory by introducing other types of uncertainties as well. The motivation behind the second paradigm was that the classical theories were perceived as being too demanding for practical decision making. However, by relaxing the strong requirements, the price paid was that usually the alternatives became more difficult to discriminate between. On the other hand, the relaxation is necessary since the classical theories can be misleading when forcing decision-makers to assert precise values even when they are not available. In this paper we propose a third generation of decision analysis models by showing the importance of the decision structure for an adequate understanding of decision situations and introducing a theory for the quantification of this structure.

First Generation Models

Probabilistic decision models are often given a tree representation. A decision tree consists of a root, representing a decision, a set of intermediary (event) nodes, representing some kind of uncertainty and consequence nodes, representing possible final outcomes.

Usually, probability distributions are assigned in the form of weights in the probability nodes as measures of the uncertainties involved. The informal semantics is simply that given an alternative A_i being chosen, there is a probability p_{ij} that an event e_{ij} occurs. The event is conditional on the alternative being chosen. This event can either be a consequence with a utility v_{ij} assigned to it or it can be another event. Usually, the maximization of the expected utility is used as an evaluation rule.

This is a straight-forward characterization of a multi-level probabilistic decision model in classical decision analysis and a quite widespread opinion is that this captures the concept of rationality. Several variations of this principle have been systematized by, e.g., (Schoemaker 1982). The most basic approach to evaluation is to assign fixed numbers to the probabilities and utilities in the tree which then yields a fixed numeric expected value for each alternative under consideration. The inherent uncertainty and imprecision is then addressed by sensitivity analyses performed as add-on procedures. Despite performing sensitivity analyses, a numerically precise approach often puts too strong a demand on the input capability of the decision-maker, and various alternatives for imprecise reasoning have emerged as a response.

Second Generation Models

During the last 45 years, various methods related to imprecise estimates of probabilities and utilities of any sort (not only numerical values) have been suggested. Early examples include (Hodges 1952), (Hurwicz 1951), and (Wald 1950). Some of the suggestions are based on capacities, evidence theory and belief functions, various kinds of logic, upper and lower probabilities, or sets of probability measures. For overviews, see, e.g., (Weichselberger and Pöhlman 1990), (Walley 1991), and (Ekenberg and Thorbiörnson 2001). The common characteristic of the approaches is that they typically do not include the additivity axiom of probability theory and consequently do not require a decision-maker to model and evaluate a decision situation using precise probability (and, in some cases, utility) estimates. Examples of recent discussions include (Klir 1999), (Cano and Moral 1999), (Danielson et al. 2003), (Danielson 2005), and (Augustin 2001).

Intervals

The primary evaluation rules of an interval decision tree model are based on the expected utility. Since neither probabilities nor utilities are fixed numbers, the evaluation of the expected utility yields multi-linear expressions.

Definition 1. *Given a decision tree with r alternatives A_i for $i = 1, \ldots, r$, the expression*

$$E(A_i) = \sum_{i_1=1}^{n_{i_0}} p_{ii_1} \sum_{i_2=1}^{n_{i_1}} p_{ii_1i_2} \cdots \sum_{i_{m-1}=1}^{n_{i_{m-2}}} p_{ii_1i_2\ldots i_{m-2}i_{m-1}} \sum_{i_m=1}^{n_{i_{m-1}}} p_{ii_1i_2\ldots i_{m-2}i_{m-1}i_m} v_{ii_1i_2\ldots i_{m-2}i_{m-1}i_m},$$

where m is the depth of the tree corresponding to A_i, n_{i_k} is the number of possible outcomes following the event with probability p_{i_k}, $p_{\ldots i_j \ldots}$, $j \in [1, \ldots, m]$, denote probability variables and $v_{\ldots i_j \ldots}$ denote utility variables as above, is the expected utility *of alternative A_i.*

Maximization of such non-linear objective functions subject to linear constraint sets (statements on probability and utility variables) are computationally demanding problems to solve for a real-time software agent or an interactive decision tool in the general case, using techniques from the area of non-linear programming. In, e.g., (Danielson and Ekenberg 1998), (Ding et al. 2004), and (Danielson 2004), there are discussions about computational procedures reducing the evaluation of non-linear decision problems to systems with linear objective functions, solvable with ordinary linear programming methods. The approach taken is to model probability and utility intervals as constraint sets, containing statements on the upper and lower bounds. Furthermore, normalisation constraints for the probabilities are added (representing that the consequences from a parent node are exhaustive and pairwise disjoint). Such constraints are always on the form $\sum_{i=1}^{n} x_i = 1$. Consequently, the solution sets of probability and utility constraint sets are polytopes. The evaluation procedures then yield first-order interval estimates of the evaluations, i.e. upper and lower bounds for the expected utilities of the alternatives.

An advantage of approaches using upper and lower probabilities is that they do not require taking particular probability distributions into consideration. On the other hand, it has then often been difficult to find reasonable decision rules that select an alternative out of a set of alternatives and at the same time fully reflect the intensions of a decision-maker. Since the probabilities and utilities are represented by intervals, the expected utility range is also an interval. Consequently, it has then often not been possible to discriminate between the alternatives. In effect, such a procedure keeps all alternatives with overlapping expected utility intervals, even if the overlap is indeed small.

Some approaches for extending the representation using distributions over classes of probability and utility measures have been suggested in, e.g., (Gärdenfors and Sahlin 1982) and (Gärdenfors and Sahlin 1983). These have been developed into various hierarchical models, but in general, no detailed procedures or suggestions are provided for how to represent or how to evaluate aggregations of belief distributions.

Third Generation Models

As discussed above, it is known that a second-order representation adds information to a decision model (Ekenberg and Thorbiörnson 2001), (Ekenberg, Danielson, and Thorbiörnson 2006). In this paper, we show that the *structure* of the decision tree itself also adds information. Below, we discuss this more general view of a decision tree in which the structure of the tree itself leads to second-order effects even when no second-order information is explicitly provided. Hence, we propose a third generation of decision models consisting of models taking decision structure effects into account in addition to representation and evaluation issues already considered in the previous generations. Such models are able to handle structure effects in pure interval decision trees as well as in trees containing second-order statements.

Structural Information

Earlier, second generation attempts to study second-order effects do not take tree depth into account. The characteristic of a decision tree with some depth is that the local probabilities of the event nodes are multiplied in order to obtain the global probability of the combined events, i.e. of the path from the root to each leaf. It is important to note that second-order distributions do not have to be explicitly introduced for the local or global probabilities. Second-order effects always exist in a decision tree, even in pure interval representations, independent of whether distributions of belief are explicitly stated or not. In the evaluation of decision trees, the operations involved are *multiplications* and *additions*. When considering distributions over intervals, it is well-known that the results of additions have a tendency to concentrate around the expected values of the distributions. Thus, we will here concentrate on the effects of multiplications in such trees, i.e. the effect of tree structure on the decision evaluation. For calculating expected utilities in decision trees, the two effects are present at the same time, i.e. additive effects for global probabilities aggregated together with the utilities at the leaf nodes and multiplicative effects for intermediate local probabilities. The third generation of models is capable of handling these effects and turn them into advantages in evaluations.

To begin with, consider an upper triangular distribution of belief, see Figure 1.

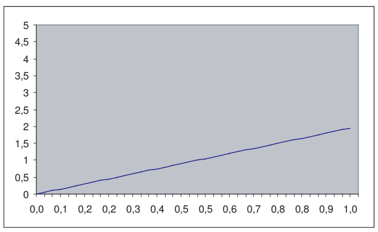

Figure 1: Upper triangular distribution of belief

Figure 2 shows the multiplication of two and three intervals for the upper triangular distributed variables, i.e. two and three levels of tree depth. Even though we started with distributions laying heavy on the right (upper) values (Figure 1), they have become warped after only a few multiplications.

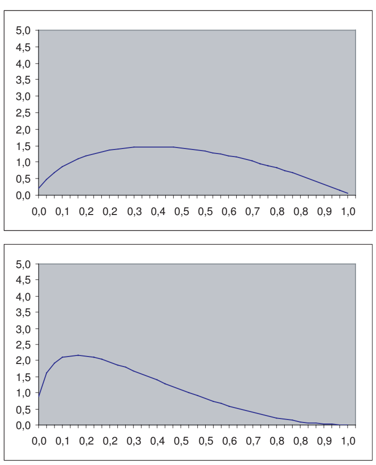

Figure 2: Multiplication of two and three upper triangular variables, respectively

Already from these low-dimensional examples, it is clear that the global distributions resulting from multiplications have shapes very different from their marginal components. Whether there are explicit statements on the distributions of belief or not, the warp of belief in multiplications is a strong effect and demonstrates that an analysis using only upper and lower bounds is not taking all structural information into account. As will be discussed below, the results of multiplication on dependent variables such as probabilities (with the normalization constraint of having to sum to one) are even more warped.

The intuition shown above is that multiplied (global) distributions become considerably warped compared to the corresponding component (marginal) distributions. Such multiplications occur in obtaining the expected utility in interval decision trees and probabilistic networks, and the warp effect can be used as an advantage, enabling discrimination while still allowing overlap. Properties of additions of components follow from ordinary convolution, i.e., there is a strong tendency to concentrate towards the middle. We will exemplify the combined effect below.

In conclusion, the observed warp in resulting global distributions is due to structural effects, i.e. the tree structure is carrying information. When there is information on belief in marginal probabilities, this will further enhance the warping. Below, we will look more formally at the structure effect.

Belief Distributions

Probability and utility estimates in a decision tree can be expressed by sets of probability distributions and utility functions. Second-order estimates, such as distributions expressing various beliefs, can be defined over a multi-dimensional space, where each dimension corresponds to, for instance, possible probabilities of events or utilities of consequences. In this way, the distributions can be used to express varying strength of beliefs in different vectors in the polytopes. But this is not the only use for distributions. They can also be employed in an analysis of the effects the structures themselves have on aggregated belief.

Traditional interval estimates (lower and upper bounds) can be considered as special cases of representations based on belief distributions. For instance, a belief distribution can be defined to have a positive support only for $x \leq y$. More formally, the solution set to a probability or utility constraint set is a subset of a unit cube since both variable sets have $[0, 1]$ as their ranges. This subset can be represented by the support of a distribution over a cube.

Definition 2. *Let a* unit cube *be represented by* $B = (b_1, \ldots, b_k)$*. The* b_i *are explicitly written out to make the labels of the dimensions clearer. (More rigorously, the unit cube should be represented by all the tuples* $(x_1, \ldots, x_k)$ *in* $[0, 1]^k$*.)*

Definition 3. *By a* belief distribution *over* B*, we mean a positive distribution* F *defined on the unit cube* B *such that*

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

where V_B *is some* k*-dimensional Lebesque measure on* B*. The set of all belief distributions over* B *is denoted by* $BD(B)$.

For our purposes here, second-order probabilities are an important sub-class of these distributions and will be used below as a measure of belief, i.e. a second-order joint probability distribution. Marginal distributions are obtained from the joint ones in the usual way.

Definition 4. *Let a unit cube* $B = (b_1, \ldots, b_k)$ *and* $F \in BD(B)$ *be given. Furthermore, let* $B_i^- = (b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_k)$*. Then*

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

is a marginal distribution *over the axis* b_i*.*

Distributions over Normalized Planes

Regardless of the actual shape of a global distribution, constraints like $\sum_{i=1}^{n} x_i = 1$ must be satisfied, since it is not reasonable to believe in an inconsistent point such as, e.g., $(0.1, 0.3, 0.5, 0.4)$ if the vector is supposed to represent a probability distribution. Consequently, the Dirichlet distribution is a convenient and general way of expressing distributions, even if this is not the only feasible candidate. This distribution has the general form:

$$\frac{\Gamma(\sum_{i=1}^{k} \alpha_i)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} p_1^{\alpha_1 - 1} p_2^{\alpha_2 - 1} \cdots p_k^{\alpha_k - 1}$$

on a set $\{p_1, p_2, \ldots, p_k \geq 0 : \sum p_i = 1\}$, where $\alpha_1, \alpha_2, \ldots, \alpha_k$ are positive parameters and $\Gamma(x)$ is a gamma-distribution over x.

A marginal distribution of a Dirichlet distribution in a cube $B = (b_1, \ldots, b_k)$ is a beta-distribution. For instance, if the distribution is uniform, the resulting marginal distribution (over an axis) is a polynomial of degree $n-2$, where n is the dimension of B.

Example 1. *An example is the marginal distribution* $f(x_i)$ *of a uniform distribution over the surface* $\sum_{i=1}^{n} x_i = 1$ *in a 4-dimensional cube, which is*

$$f(x_i) = \int_0^{1-x_i} \int_0^{1-y-x_i} 6\, dz\, dy = 3(1-2x_i+x_i^2) = 3(1-x_i)^2\,.$$

This tendency is the result of a general phenomenon that becomes more emphasized as the dimension increases. This observation of marginal probabilities is important for a structural analysis of decision trees and networks.

For simplicity, a uniform a priori belief representation will be used in the paper for studying the structure effects, i.e., $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 1$ in the Dirichlet distribution. Other approaches, such as preferences for and valuation of a gamble, lead to a decision-maker's support of it, in which case the second-order distribution could be interpreted as varying (or non-varying, i.e. uniform) support. Thus, while 'belief' is used throughout the paper, it should be thought of as the decision-maker's support.

Evaluations

Now, we will consider how to put structural and belief information into use in order to further discriminate between alternatives that evaluate into overlapping expected utility intervals when using first-order interval evaluations. The main idea is not to require a total lack of overlap but rather allowing overlap by interval parts carrying little belief mass, i.e. representing a very small part of the decision-maker's belief. Then, the non-overlapping parts can be thought of as being the core of the decision-maker's appreciation of the decision situation, thus allowing discrimination. Again, in this paper we mainly discuss the structural effects. In addition, effects from varying belief (i.e. differing forms of belief distribution) should be taken into account.

Evaluation of expected utilities in interval decision trees lead to multiplication of probabilities using a type of "multiplicative convolution" of two densities from (Ekenberg, Danielson, and Thorbiörnson 2006).

First some terminology. Let G be a belief distribution over the two cubes A and B. Assume that G has a positive support on the feasible probability distributions at level i in a decision tree, i.e., is representing these (the support of G in cube A), as well as on the feasible probability distributions of the children of a node x_{ij}, i.e., $x_{ij1}, x_{ij2}, \ldots, x_{ijm}$ (the support of G in cube B). Let $f(x)$ and $g(y)$ be the projections of $G(z)$ on A and B, respectively.

Definition 5. *The* cumulative distribution *of the two belief distributions* $f(x)$ *and* $g(y)$ *is*

$$H(z) = \iint_{\Gamma_z} f(x)g(y)\, dx\, dy = \int_0^1 \int_0^{z/x} f(x)g(y)\, dy\, dx =$$
$$\int_0^1 f(x)G(z/x)\, dx = \int_z^1 f(x)G(z/x)\, dx\,,$$

where G *is a primitive function to* g*,* $0 \leq z \leq 1$ *and* $\Gamma_z = \{(x,y) : x \cdot y \leq z\}$.

Let $h(z)$ *be the corresponding* density function*. Then*

$$h(z) = \frac{d}{dz} \int_z^1 f(x)G(z/x)\, dx = \int_z^1 \frac{f(x)g(z/x)}{x}\, dx\,.$$

Informally, it means that the beliefs of x and y are multiplied and then added for all values $z = x \cdot y$. The addition of such products is analogous to the product rule for first-order (standard) probabilities. Similarly, addition is the ordinary convolution of two densities.

Definition 6. *The distribution* h_2 *on a sum* $2_2 = x + y$ *of two independent variables associated with belief distributions* $f(x)$ *and* $g(y)$ *is given by a convolution*

$$h_2(z_2) = \int_{-\infty}^{\infty} f(x)g(y-x)\, dx.$$

Using these combination rules, there are two main cases. The linearly independent case (utility variables) and the linearly dependent case (probability variables).

Assume that the assertions (statements) involved are made through intervals and that the constraint sets are linearly independent. If the marginal belief distributions are uniform, the multiplications of uniform distributions over intervals $[0, 1]$ result in the following Theorem.

Theorem 1. *Let $f_1(x_1) = 1, \ldots, f_m(x_m) = 1$, be belief distributions over the intervals $[0, 1]$. The product $h_m(z_m)$ over these m factors is the distribution.*

$$h_m(z_m) = \frac{(-1)^{m-1}(\ln(z_m))^{m-1}}{(m-1)!} .$$

Here, the strong structure effect introduced by multiplications can be seen. The mass of the resulting belief distributions becomes more concentrated to the lower values, the more factors are involved. Indeed, already after one multiplication, this effect is clearly seen $(-\ln(z))$. From initially uniform distributions, the resulting distribution has quite differing properties. It still has a non-zero support on the entire interval $[0, 1]$, but the relative beliefs in the various feasible points are shifted towards the lower bound. For instance, the belief distribution $-(\ln(x))^3/6$ has about 97% of its belief mass over the sub-interval $[0, 0.3]$, a result that deviates from the initial equal belief in the entire interval $[0, 1]$.

Next, we consider dependencies within the constraint set. A *local* constraint set is a constraint set where all the labels are children of the same node and when only linear constraints are given, a belief distribution $F(x_1, x_2, \ldots, x_n)$ is assumed to be constant over n nodes. As demonstrated above, the projections of a uniform distribution over the solution set to a probability constraint set, without other constraints than the default normalizations, are polynomials of degree $n - 2$.

Example 2. *For instance, when having a 4-ary tree of depth 3, with all initial belief being uniform over [0,1], the resulting distribution becomes as in Figure 3:*

$$\frac{27}{2}\left(24(z-1)^2 - 9(z^2-1)\ln(z) - (z^2+8z-1)(\ln(z))^2\right) .$$

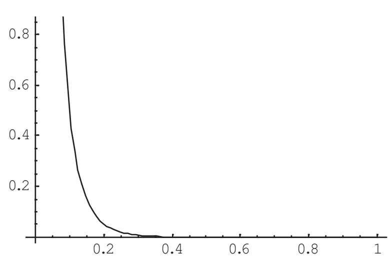

Figure 3: The projection of the distribution over a 4-tree of depth 3

In general, the structure effects are noticeable when evaluating imprecise decision problems. The most important sub-intervals to consider are the supports of the distributions where the most mass is concentrated. This can be seen in Figure 3, where about 95% of the mass is concentrated to the interval $[0, 0.07]$. Thus, even if we have no information on the probabilities, the resulting multiplied (joint) distribution does contain information and depends on the depth of the tree. This means that different trees contain different distributions of belief and thus of expected utility even before probabilities or utilities are assigned in the tree. This difference is only attributable to the structure of the tree. This can be compared to the ordinary multiplication of extreme points (bounds) which generates an interval $[0, 1]$ regardless of tree structure without any further discrimination within the interval which could be seen as rather misleading given the mass concentration at the lower end of the resulting distribution. This shows the importance of including the effects of structure in the evaluation of interval decision trees, regardless of there being explicit distributions of belief or not.

Summary and Concluding Remarks

In this paper, we show the importance of structure. We have demonstrated that evaluations incorporating structure and belief can supply important insights for the decision-maker (software agent or human being) when handling aggregations of interval representations, such as in decision trees or probabilistic networks, and that interval estimates (upper and lower bounds) in themselves are not complete, especially not when it comes to handling information inherent in the structure of decision trees. This applies to all kinds of probabilistic decision trees and criteria weight trees (and to probabilistic networks as well) since they all use multiplications.

We have therefore suggested a new generation of methods of structural and belief analyses as a complement to non-structural methods such as interval analyses. The latter is sometimes necessary in, e.g., low probability/high consequence analyses, but in most situations, it covers too wide intervals including cases with too little belief support as well and omits structural information. The rationale behind this fact is that we have demonstrated that multiplied global distributions warp compared to their marginal component distributions. The multiplied distributions concentrate their mass to the lower values compared to their component distributions.

References

Augustin, Th. 2001. On Decision Making under Ambiguous Prior and Sampling Information, *Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications.*

Cano, A., and Moral, S. 1999. A Review of Propagation Algorithms for Imprecise Probabilities. *Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications.*

Danielson, M. 2004. Handling imperfect user statements in real-life decision analysis. *International Journal of Information Technology and Decision Making, vol. 3, no. 3*, pages 513–534.

Danielson, M. 2005. Generalized Evaluation in Decision Analysis, *European Journal of Operational Research* 162(2), pages 442–449.

Danielson, M. and Ekenberg, L. 1998. A framework for analysing decisions under risk. *European Journal of Operational Research*, Vol. 104, Issue 3, pages 474–484.

Danielson, M., Ekenberg, L., Johansson, J., and Larsson, A. 2003. The *Decide*IT Decision Tool. *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications*, pages 204–217. Carleton Scientific.

Dempster, A.P. 1967. Upper and Lower Probabilities Induced by a Multivalued Mapping, *Annals of Mathematical Statistics 38*, pages 325–339.

Ding, X. S., Danielson, M., and Ekenberg, L. 2004. Nonlinear programming solvers for decision analysis support systems. *Operations Research Proceedings 2003 - Selected Papers of the International Conference on Operations Research (OR 2003)*, pages 475–482. Springer-Verlag.

Ekenberg, L. and Thorbiörnson, J. 2001. Second-order decision analysis. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 9, No 1, pages 13–38.

Ekenberg, L., Danielson M. and Thorbiörnson J. 2006. Multiplicative Properties in Evaluation of Decision Trees. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol. 14, No. 3.

Ellsberg, D. 1961. Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics*, 75, pages 643–669.

Good, I. 1980. Some history of the hierarchical Bayesian methodology. *Bayesian Statistics*, Vol 1, pages 489–519. Valencia University Press.

Gärdenfors, P., and Sahlin, N-E. 1982. Unreliable Probabilities, Risk Taking, and Decision Making. *Synthese* 53, pages 361–386.

Gärdenfors, P., and Sahlin, N-E. 1983. Decision Making with Unreliable Probabilities. *British Journal of Mathematical and Statistical Psychology* 36, pages 240–251.

Hodges, J.L., and Lehmann, E.L. 1952. The Use of Previous Experience in Reaching Statistical Decisions, *The Annals of Mathematical Statistics* 23, pages 396–407.

Hurwicz, L. 1951. Optimality Criteria for Decision Making under Ignorance. *Cowles Commission Discussion Paper* 370.

Jaffray, J.-Y. 1999. Rational decision making with imprecise probabilities. *Proceedings of ISIPTA'99*.

Klir, G.J. 1999 Uncertainty and Information Measures for Imprecise Probabilities: An Overview. *Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications*.

von Neumann, J. and Morgenstern, O. 1944. *Theory of Games and Economic Behaviour*. Princeton University Press.

Savage, L.J. 1972. *The Foundation of Statistics*, 2nd ed, John Wiley and Sons.

Schoemaker, P. 1982. The expected utility model: Its variants, purposes, evidence and limitations. *Journal of Economic Literature*, vol. XX, pages 529–563.

Smith, C. A. B. 1961. Consistency in statistical inference and decision. *Journal of the Royal Statistical Society*, Series B, xxiii, pages 1–25.

Wald, A. 1950. *Statistical Decision Functions*, John Wiley and Sons.

Walley, P. 1991. *Statistical reasoning with Imprecise Probabilities*. Chapman and Hall.

Weichselberger, K. and Pöhlman, S. 1990. *A Methodology for Uncertainty in Knowledge-Based Systems*. Springer-Verlag.