MATHEMATICAL MODELS FOR HANDLING PARTIAL KNOWLEDGE IN ARTIFICIAL INTELLIGENCE

Edited by
Giulianella Coletti
Didier Dubois and
Romano Scozzafava

Mathematical Models for Handling Partial Knowledge in Artificial Intelligence

Edited by

Giulianella Coletti

University of Perugia
Perugia, Italy

Didier Dubois

University of Toulouse
Toulouse, France

and

Romano Scozzafava

University "La Sapienza"
Rome, Italy

Plenum Press • New York and London

Library of Congress Cataloging-in-Publication Data

On file

Proceedings of the International School of Mathematics "G Stampacchia", Workshop on Mathematical Models for Handling Partial Knowledge in Artificial Intelligence, held June 19–25, 1994, in Erice, Italy

ISBN 0-306-45076-3

A Division of Plenum Publishing Corporation
233 Spring Street, New York, N. Y. 10013

10 9 8 7 6 5 4 3 2 1

Printed in the United States of America

Mathematical Models for Handling Partial Knowledge in Artificial Intelligence

PREFACE

Knowledge acquisition is one of the most important aspects influencing the quality of methods used in artificial intelligence and the reliability of expert systems. The various issues dealt with in this volume concern many different approaches to the handling of partial knowledge and to the ensuing methods for reasoning and decision making under uncertainty, as applied to problems in artificial intelligence.

The volume is composed of the invited and contributed papers presented at the Workshop on Mathematical Models for Handling Partial Knowledge in Artificial Intelligence, held at the Ettore Majorana Center for Scientific Culture of Erice (Sicily, Italy) on June 19-25, 1994, in the framework of the International School of Mathematics "G.Stampacchia". It includes also a transcription of the roundtable held during the workshop to promote discussions on fundamental issues, since in the choice of invited speakers we have tried to maintain a balance between the various schools of knowledge and uncertainty modeling.

Choquet expected utility models are discussed in the paper by Alain Chateauneuf: they allow the separation of perception of uncertainty or risk from the valuation of outcomes, and can be of help in decision making.

Petr Hájek shows that reasoning in fuzzy logic may be put on a strict logical (formal) basis, so contributing to our understanding of what fuzzy logic is and what one is doing when applying fuzzy reasoning.

The mathematical foundations of evidence theory are expounded by Jürg Kohlas, leading to belief and plausibility functions of the kind introduced and studied by G.Shafer: however it is not confined to finite frames but is entirely general.

Those uncertain inferences based on statistical knowledge, that are valid if their conclusions are true in a large proportion of the models in which the relevant premises are true, are dealt with by Henry Kyburg.

Frank Lad, in his first paper (co-authored with Ian Coope), goes through prospects and problems in applying the fundamental theorem of prevision as an expert system, by means of a detailed discussion of an example of learning about parole decisions. His second paper presents coherent prevision as a linear functional without an underlying measure space, based on the purely arithmetic structure of logical relations among conditional quantities.

Revision rules for convex sets of probabilities are discussed in the paper by Serafin Moral (co-authored with Nic Wilson), emphasizing the differences between revision and focusing. These two procedures are expressed by using the logical language of gambles.

The appropriate mathematical tools for decision making (including subjective probability, lower probabilities, the Choquet integral, random sets, measure-free representation of conditionals, rule-based procedures) are considered in the paper by Hung Nguyen, depending upon the form of the available knowledge (data).

Judea Pearl demonstrates in his paper the use of graphs as a mathematical tool for expressing independencies, and as a formal language for communicating and processing causal information for decision analysis and for organizing claims about external interventions and their interactions.

The ten contributed papers deal with the generalized concept of atoms for conditional events (A.Capotorti), the checking of coherence of conditional probabilities in expert systems (G.Di Biase and A.Maturo), the study of an hyperstructure of conditional events for artificial intelligence (S.Doria and A.Maturo), an overview on the application of possibility theory to automated reasoning (D.Dubois and H.Prade), a formulation of probability logic as fuzzy logic (G.Gerla, abstract), algorithms for precise and imprecise conditional probability assessments (A.Gilio), a valuation-based architecture for assumption-based reasoning (R.Haenni, abstract), the computation of symbolic support functions by classical theorem-proving techniques (U.Hänni, abstract), inconsistent knowledge integration in a probabilistic model (R.Jiroušek and J.Vomlel), and the use of conditional and comparative probabilities in artificial intelligence (P.Vicig).

Giulianella Coletti, Didier Dubois, Romano Scozzafava

CONTENTS

INVITED PAPERS

CONTRIBUTED PAPERS

ELLSBERG PARADOX INTUITION AND CHOQUET EXPECTED UTILITY

Alain Chateauneuf

Cermsem
Université Paris I
75634 Paris Cedex 13
France

INTRODUCTION

The aim of this paper is to introduce through two interpretations of Ellsberg paradox, Choquet expected utility (C.E.U.) models, a popular class of models introduced separately by Quiggin (1982), Yaari (1987) and Schmeidler (1982, 1989). Such models based on Choquet integral (Choquet (1954)) offer flexible but simple formulas, explain paradoxes of Allais (1953) under risk and of Ellsberg (1961) under uncertainty ; moreover they allow to separate perception of uncertainty or risk from the valuation of outcomes. Here we emphasize the intuitive and appealing meaning of "comonotonicity", and show the ability of C.E.U. models in modeling attitudes towards uncertainty and risk. A simple characterization of C.E.U. model, as in Chateauneuf (1994 a), is proposed, under the restrictive assumption of constant marginal utility for wealth. The general model, removing this restriction, is evocated in conclusion. Some economic applications are also quoted.

In section 1 (Ellsberg paradox intuition and comonotonicity) we analyze Ellsberg paradox through the concept of comonotonicity, and show that typical preferences in Ellsberg's experiment, can be explained by

Mathematical Models for Handling Partial Knowledge in Artificial Intelligence
Edited by G. Coletti *et al.*, Plenum Press, New York, 1995

hedging effects i.e. "violation of comonotonicity". Interpreting Ellsberg paradox in Auscombe-Aumann (1963) framework (section 1.1.2.), as initiated by Schmeidler (1982, 1989), leads to the fundamental Choquet expected utility model under uncertainty introduced by Schmeidler (1982, 1989). Interpreting Ellsberg paradox in the simpler framework of choices between acts X which are real-valued functions defined on a set S of states of nature (section 1.1.1.), leads to a simplified version of Schmeidler's model as exposed in Chateauneuf (1994 a). These two interpretations are related in section 1.2. to behaviors modelizable through axioms of uncertainty aversion (Schmeidler (1989), Chateauneuf (1994 a)) or pessimism (Wakker (1990 a)).

In section 2 (Choquet expected utility and comonotonicity) as in Chateauneuf (1994 a) we present in a unified framework the Choquet expected utility model under uncertainty of Schmeidler (1982, 1989) and under risk of Yaari (1987). This will be done by assuming as Yaari that our decision maker displays a constant marginal utility of wealth. Using stimulating results of Landsberger and Meilijson (1993), a recent interpretation of the central comonotonic independence axiom as in Chateauneuf, Kast and Lapied (1994) is proposed, based on the intimate connection between comonotonicity and the Bickel-Lehmann (1979) dispersion order. As for economic applications, under uncertainty we focus on the ability of Schmeidler's model to explain the gap between selling and buying prices of financial assets (Dow and werlang (1992), Epstein and Wang (1994), Chateauneuf, Kast and Lapied (1994)) ; under risk a direct application of Yaari's model to insurance is examined : Denneberg (1990), who proposes a convincing premium principle.

In concluding remarks, in the particular case of monetary payments -i.e. of a consequence set $\mathscr{C} = \mathbb{R}$-, we quote some results concerning the general C.E.U. model, when the constant marginal utility assumption is removed.

1. ELLSBERG PARADOX INTUITION AND COMONOTONICITY

1.1. Ellsberg Paradox Intuition

Let us consider the following version of Ellsberg paradox (Ellsberg (1961)).

Subjects are informed that a ball will be drawn at random from an urn that contains 90 balls : 30 are red (R), and each other ball is either black (B) or yellow (Y). Subjects are requested to express their preferences between betting on R (act f_1) or betting on B (act f_2) and also between betting on $R \cup Y$ (act f_3) or betting on $B \cup Y$ (act f_4). Table 1 below summarizes corresponding outcomes :

Table 1 :

	30 Red	60 Black	60 Yellow
f_1	\$100	\$0	\$0
f_2	\$0	\$100	\$0
f_3	\$100	\$0	\$100
f_4	\$0	\$100	\$100

Typical preferences are $f_1 \succ f_2$ and $f_4 \succ f_3$[1] ,hence the Sure-Thing Principle is violated, since f_3 (respectively f_4) is obtained from f_1 (respectively f_2) by merely changing the common outcome \$0 under Y into a common outcome \$100 under Y. Therefore many subjects behave in a "paradoxical" way, in the sense that they are not subjective expected utility maximizers.

As noticed by Machina and Schmeidler (1992) nor such subjects are probabilistically sophisticated : this means that they do not ascribe subjective probabilities p_R, p_B, p_Y to states of nature (i.e. elementary events R, B, Y) and then use first order stochastic dominance[2] - a widely accepted rule for partially ordering random variables-. Otherwise $f_1 \succ f_2$ would entail $p_R > p_B$ and $f_4 \succ f_3$ would entail $p_B + p_Y > p_R + p_Y$: a contradiction.

Let us interpret now the Ellsberg paradox :

1.1.1. Interpretation 1. Let us interpret bets as acts (i.e. mappings) $f : S = \{R,B,Y\} \longrightarrow \mathbb{R}$, and let g be the act defined by $g(R) = g(B) = 0$, $g(Y) = 100$. Clearly $f_3 = f_1 + g$ and $f_4 = f_2 + g$. That $f_1 \succ f_2$ may be interpreted by the fact that there is no uncertainty for f_1 in the sense that Pr {winning \$100 through f_1} is known to be $\frac{1}{3}$, whereas on the

---

[1] $f_1 \succ f_2$ means f_1 is strictly preferred to f_2, the same applies to $f_4 \succ f_3$

---

[2] Recall that if $(S, \mathcal{a}, P)$ is a probabilized space, and if X and Y are $\mathcal{a}$ measurable real-valued functions on S -i.e. real random variables for short-then the first order stochastic dominance rule stipulates that if $\forall t \in \mathbb{R}$, $P(X \geq t) \geq P(Y \geq t)$ then X should be weakly prefered to Y, the preference becoming strict if $P(X \geq t_0) > P(Y \geq t_0)$ for some $t_0 \in \mathbb{R}$.

contrary there is some uncertainty for f_2 since $\Pr\{f_2 = 100\}$ belongs to $\left[0,\frac{2}{3}\right]$; hence aversion to uncertainty may explain $f_1 \succ f_2$.

Adding g to f_1 entails some increase of uncertainty since now $\Pr\{f_3 = 100\} \in \left[\frac{1}{3},1\right]$, on the contrary there is some reduction of uncertainty when adding g to f_2 since $\Pr\{f_4 = 100\}$ is precisely known equal to $\frac{2}{3}$; therefore aversion to uncertainty may explain that now $f_4 \succ f_3$. This can explain the preference reversal.

We come now to the notion of comonotonicity.

Definition 1. Let S be a set of states of nature, and f, g be two acts i.e. two elements of $V = \mathbb{R}^S$, f and g will be said comonotonic if $\forall s,t \in S$: $(f(s) - f(t))(g(s) - g(t)) \geq 0$, i.e. if "f and g have the same sense of variation."

It is easy to understand that if f_1 and g were comonotonic, and if f_2 and g were comonotonic too, then g would be neither a hedge against f_1 nor against f_2, hence it would be natural to require (or to observe) that direction of preferences be retained after adding g both to f_1 and to f_2 -no asymmetric reduction of uncertainty would result-. In the previous example g is neither comonotonic with f_1 nor with f_2, this can explain asymmetric reduction of uncertainty will result by adding g to f_1 and to f_2, hence the observed preference reversal.

We will come again to this way of interpretation in the sequel. This first interpretation is in the spirit of the simplified version of Schmeidler's model, which will be examined in section 2. We now come to interpretation 2, in the spirit of Schmeidler's model (1982, 1989).

1.1.2. Interpretation 2. Here Ellsberg paradox is interpreted in Anscombe and Aumann's framework (1963). Now uncertainty concerns the composition of the urn. The set S of states of nature is now composed of sixty one possible states of nature $S = \{0,1,\ldots,k,\ldots,60\}$ where k stands for the elementary event : "the number of black balls is k".

Bets (i.e. proposals) f are now considered as horse lotteries $X_f : S \longrightarrow \mathcal{Y}$ i.e. mappings from S to $\mathcal{Y}$ where $\mathcal{Y}$ is the set of -roulette- lotteries on $\mathcal{C} = \{0,100\}$.

Thus $X_f(k) = \frac{30}{90}\,\delta_{f(R)} + \frac{k}{90}\,\delta_{f(B)} + \frac{60-k}{90}\,\delta_{f(Y)}$ is the lottery faced by the D.M. (i.e. Decision maker) if his bet is f and if the number of black balls is k, namely in such a case with probability $\frac{30}{90}$ he will earn \$f(R), with probability $\frac{k}{90}$ \$f(B), and with probability $\frac{60-k}{90}$ \$f(Y).

Suppose as implicitely assumed by Schmeidler as by Anscombe and Aumann, that our D.M. is an expected utility maximizer with Von Neumann Morgenstern's utility function (1947) u (assumed without loss of generality to satisfy $u(0) = 0$, $u(100) = 1$). Denoting by $X_i(k)$ $i = 1,2,3,4$ the value $u(X_{f_i}(k))$ of the resulting lottery if the true state of nature is k, leads $\forall k \in S$ to : $X_1(k) = \frac{100}{3}$, $X_2(k) = \frac{100\ k}{90}$, $X_3(k) = X_1(k) + X(k)$, $X_4(k) = X_2(k) + X(k)$ where $X(k) = \frac{100}{90}(60-k)$

It is immediate that X is comonotonic with X_1, but not with X_2. More precisely $\forall k,\ell \in S$ one gets $(X_2(k) - X_2(\ell))(X(k) - X(\ell) \leq 0$, with strict inequality if $k \neq \ell$; hence X is a hedge against X_2 but not against X_1.

Adding X to X_2 smoothes values (i.e. reduces uncertainty) thus here $X_4(k)$ = constant = $\frac{100 \times 60}{90}$ $\forall k$ while $X_2(k) - X_2(\ell) = \frac{100}{90}(k-\ell)$. On the contrary adding X to X_1 does not smooth values : initially $X_1(k)$ = constant = $\frac{100}{3}$, but $X_3(k) - X_3(\ell) = \frac{100}{90}(\ell-k)$.

This can explain, since usually decision makers are uncertainty averse, the observation of typical preferences $f_1 \succ f_2$ and $f_4 \succ f_3$ in Ellsberg's experiment.

To end this paragraph notice that of course in Anscombe-Aumann framework, the expected utility model under uncertainty cannot explain previous preferences. Actually suppose the D.M. assigns probabilities p_k to events k, and acts according to the expected utility model of Anscombe-Aumann, i.e. weakly prefers f to g if and only if $\Sigma\, p_k\, u(X_f(k)) \geq \Sigma\, p_k\, u(X_g(k))$ where $u(100) = 100$ and $u(0) = 0$.

Hence $f_1 \succ f_2$ would give $\frac{100}{3} > \frac{100}{90}(\Sigma\, p_k.k)$ i.e. $30 > \Sigma\, p_k.k$

and $f_4 \succ f_3$ would give $\frac{100}{90} \times 60 > \frac{100}{90}(90 - \Sigma\, p_k.k)$ i.e. $30 < \Sigma\, p_k.k$, a contradiction.

1.2. Relaxing independence conditions into comonotonic independence conditions.

1.2.1. Schmeidler's model and Choquet expected utility for uncertainty. Let S be a finite set of states of nature, and let $\mathcal{A} = 2^S$ be the events. Define the acts as the set V of horse lotteries i.e. of mappings from S to $\mathcal{Y}$ where $\mathcal{Y}$ is the set of roulette lotteries on a consequence set $\mathcal{C}$, say $\mathcal{C} = \mathbb{R}$, i.e. $\mathcal{Y}$ is the set of probability distributions over $\mathcal{C}$ with finite support.

$\succsim$ will be the preference relation of a D.M. over the acts V where as usually $X \succsim Y$ means X is -weakly- preferred to Y, $X \succ Y$ means X is strictly preferred to Y and $X \sim Y$ means the D.M. is indifferent between X and Y.

Definition 2. Acts X and Y are said to be comonotonic if for no s and t in S, $X(s) \succ X(t)$ and $Y(t) \succ Y(s)$.

Independence condition (Anscombe and Aumann). For all X,Y and Z in V and for all α in (0,1) : $X \succ Y$ implies $\alpha X + (1-\alpha)Z \succ \alpha Y + (1-\alpha)Z$.

Comonotonic independence (Schmeidler). For all pairwise[3] comonotonic acts X,Y and Z in V and for all α in (0,1) : $X \succ Y$ implies $\alpha X + (1-\alpha)Z \succ \alpha Y + (1-\alpha)Z$.

Considerations of section 1.1.2. explain why by relaxing the independence condition into comonotonic independence, Schmeilder has obtained a model in which the special preferences of the above example become admissible. Let us precise that with the help of comonotonic independence and a few usual simple axioms, Schmeidler proved that the preference relation $\succsim$ on V is represented through a Choquet integral with respect to a unique capacity v (instead of a unique probability P), that is for all X and Y in V : $X \succsim Y$ iff $\int_S u(X(.))dv \geq \int_S u(Y(.))dv$, where u is a VNM -Von Neumann Morgenstern- utility function on the set $\mathcal{Y}$ of roulette lotteries.

More precisely

Definition 3. A (normalized) capacity v on $(S,\mathcal{a}=2^S)$ is a monotone set function ($A,B \in \mathcal{a}$, $A \subseteq B \Rightarrow v(A) \leq v(B)$) such that $v(\emptyset) = 0$, $v(S) = 1$.

Definition 4. Choquet integral $\int_S u(X)dv$ is defined by :

$$\int_S u(X)dv = \int_{-\infty}^{0} (v(u(X) \geq t) -1)dt + \int_{0}^{+\infty} v(u(X) \geq t)dt$$

where $v(u(X) \geq t)$ stands for $v(s \in S, u(X(s) \geq t)$.

For $X = \sum_{i=1}^{n} y_i A_i^*$, $y_i \in \mathcal{Y}$, $y_1 \precsim \ldots y_i \precsim \ldots \precsim y_n$, $A_i \in \mathcal{a}$, (A_i) partition of S, A_i^* characteristic function of A_i i.e. $A_i^*(s) = 1$ if $s \in A_i$, 0 otherwise, one gets $u(X) = \sum_{i=1}^{n} a_i A_i^*$ with $a_1 = u(y_1) \leq \ldots \leq a_i = u(y_i) \leq \ldots \leq a_n = u_i(y_n)$, and Choquet integral writes :

$$\int_S u(X)dv = a_1 + (a_2-a_1)v(u(X) \geq a_2) + \ldots + (a_{i+1}-a_i)v(u(X) \geq a_{i+1}) + \ldots + (a_n-a_{n-1})\, v(u(X) \geq a_n).$$

Notice that $\int_S u(X)dv$ is nothing else that $E_P(u(X))$ = the mathematical expectation of u(X) with respect to probability P if v proves to be equal to a probability measure P.

$\int_S u(X)dv$ can be interpreted as : the D.M. calculates "the value of X" by taking for sure the minimum expected payoff a_1, and adds to this payoff the successive possible additional payoffs $a_{i+1} - a_i$, $1 \leq i \leq n-1$,

[3] Actually it has been proved by several authors that the assumption "pairs X,Z and Y,Z are comonotonic" is enough to get Schmeidler's model.

weighted by his personal estimation $v(u(X) \geq a_{i+1})$ of their occurence. Whith the additional assumptión that the D.M. is uncertainty averse that is $X \succsim Y \Rightarrow \alpha X + (1-\alpha)Y \succsim Y$ (convexity of preferences interpreted as "smoothing potential outcomes makes the D.M. better off") Schmeidler proved that this entails v convex i.e. : $v(A \cup B) + v(A \cap B) \geq v(A) + v(B)$ $\forall A, B \in \mathcal{a}$, and the resulting utility functional (Choquet integral) on (V,) allows to explain the typical preferences of D.M. faced to Ellsberg's urn.

1.2.2. Some other axioms related to comonotonicity. In the framework of Schmeidler, developments in section 1.1.2. also explain the following axiom of Wakker (1990).

Pessimism independence

$\succsim$ satisfies pessimism independence if $X, Y, Z \in V$, Y and Z comonotonic $X \succ Y$ implies $\alpha X + (1-\alpha)Z \succ \alpha Y + (1-\alpha)Z$.

The intuitive idea in the words of P.P. Wakker is that "a pessimist dislikes uncertainty, hence the reduction of uncertainty through (eventual) hedging will lead to additional appreciation. An optimist, who expects uncertainty to turn out favorable, will not appreciate the reduction of uncertainty through hedging".

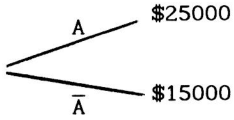
If as in section 1.1.1. acts are mappings from a set S of states of nature (here assumed to finite) to $\mathbb{R}$, the following uncertainty aversion axiom of Chateauneuf (1994 a) is related to interpretation 1.

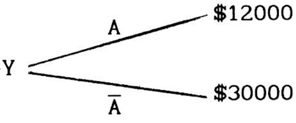
Uncertainty aversion

Let X,Y,Z be acts from $(S, \mathcal{a}=2^S)$ to $\mathbb{R}$, then the preference relation satisfies uncertainty aversion if [Y and Z comonotonic and $X \sim Y$] entails $X+Z \succsim Y+Z$.

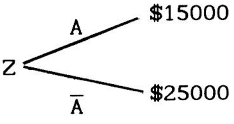
Example 1 illustrates that this axiom might reasonably be fulfilled by D.M. satisfying typical preferences of Ellsberg's experiment.

Example 1

Assume X 

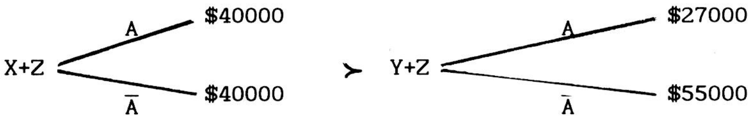
~Y 

that is the DM is indiferent between receiving $25000 if event A occurs and $15000 if A does not occur or receiving $12000 if A occurs and $30000 if A does not occur.

Let Z 

. Z is comonotonic with Y but not with X.

Z is a hedge against X but not against Y hence an uncertainty averse D.M. might exhibit after addition of Z the strict preference :



2. CHOQUET EXPECTED UTILITY AND COMONOTONICITY

Here we intend to present in a unified framework (as in Chateauneuf (1994 a)) the model under uncertainty of Schmeidler (1982,1989) and under risk of Yaari (1987). We must emphasize that this will be done by assuming, as Yaari, the restrictive assumption that our D.M. displays a constant marginal utility of wealth.

2.1 Decision under uncertainty

We consider a decision maker faced with choices among acts X, the set V of such acts consisting of all bounded real-valued a-measurable functions on S (S is a set of states of nature, a a σ.algebra of events i.e. of subsets of S). A natural way to interpret an act X is to view it as a financial asset i.e. a promise of payoffs : the D.M. will receive or pay -depending of the sign of X(s)- an amount of money X(s) if state s occurs.

The D.M. is supposed to face uncertainty, this means that objective probabilistic information concerning the occurence of events is not necessarily avalaible to him.

2.1.1. Schmeidler's model with linear utility for wealth. Let $\succsim$ be the preference relation on V of the agent. First we state three axioms which are usual and natural requirements, whatever the attitude towards uncertainty may be.

A.1. $\succsim$ is a non-trivial weak order (i.e. $\succsim$ is a binary relation on V, transitive, complete (hence reflexive) and non-trivial i.e. there exist $X, Y \in V$ such that $X \succ Y$).

A.2. Continuity with respect to monotone uniform convergence :

(A.2.1.) $[X_n, X, Y \in V,\ X_n \succsim Y,\ X_n \downarrow^u X] \Rightarrow X \succsim Y$

(A.2.2.) $[X_n, X, Y \in V,\ X_n \precsim Y,\ X_n \uparrow_u X] \Rightarrow X \precsim Y$