Non-von Neumann-Morgenstern expected utility maximization models of choice from behavioural game theory

James Wright
Department of Computer Science
University of British Columbia
Vancouver, B.C., Canada, V6T 1Z4
jrwright@cs.ubc.ca

September 23, 2008

Abstract

I survey a number of papers that describe models of choice from behavioural game theory. These models are alternatives to expected utility maximization, the standard game-theoretic model of choice [Von Neumann and Morgenstern, 1944].

1 Introduction

Behavioural game theory aims to determine and model the ways in which people systematically deviate from the traditional game theoretic model in which all agents are expected utility maximizers, and have common knowledge of this fact. This survey concentrates on models of decision that assume that agents' preferences over outcomes can be represented as a utility function of some sort, albeit possibly one whose domain is not total wealth (e.g. in prospect theory and other models involving loss aversion [Kahneman and Tversky, 1979; Tversky and Kahneman, 1992; Shalev, 2000]).

The rest of the paper is divided into two main sections. In section 2, I describe models that explain deviations from the standard model by assuming that agents maximize something other than expected utility of final wealth. In section 3, I describe models that explain deviations from the standard model by assuming that players do not necessarily ascribe full rationality to their opponents, although they may still maximize their own utility subject to that belief.

2 Non-expected utility

Under standard decision theory, agents act so as to maximize their expected utility [Von Neumann and Morgenstern, 1944]. Empirically, however, people are not expected-utility maximizers. *Prospect theory* [Kahneman and Tversky, 1979] models two of the major systematic deviations from EU maximization that are consistently observed in people: *loss aversion* (also known as *reference-dependent valuations*), where losses relative to some reference point are overweighted compared to gains; and *non-linear decision weights*, where prospects are not directly weighted by their probability of occurring, but rather by some non-linear *decision weight* instead.

Camerer [1998] surveys models of non-expected utility that better describe human decision-making than the standard model. In particular he claims that prospect theory, *non-additive probability*, and *(quasi-)hyperbolic time discounting* are better descriptive models[1] than expected-utility maximization, probability, and exponential time discounting

[1]Note that each of these models is a generalization of the corresponding standard version; e.g., quasi-hyperbolic time discounting with $\beta = 1$ corresponds to exponential time discounting.

(although obviously the latter are better normatively). Note that there is an apparent discrepancy between studies that explain aggregate behaviour, where non-linearity consistently appears, versus studies that fit individual behaviour [Hey and Orme, 1994], where sizable minorities of the subjects appeared to be expected utility maximizers. This may be either because the experiment did not include enough low (below .10) probability events, or it may be that large segments of the population are indeed expected-utility maximizers; but a single "representative agent" for such populations will still exhibit non-linearity.

2.1 Regret minimization

In his review of [Wald, 1950], Savage [1951] introduced the *minimax regret* decision rule. The advantage of minimax regret is that it is usable even by agents who have no prior belief about the relative likelihood of states of the world.

Hyafil and Boutilier [2004] propose the *minimax regret* criterion for games of strictly incomplete information, where agents have a common prior over which type profiles are possible, but their uncertainty over those profiles is unquantified (i.e., they have no probability distribution over the possible type profiles). Hyafil and Boutilier [2004] prove that every finite game has a (possibly mixed) *minimax-regret equilibrium*, where every agent is playing the max-regret-minimizing strategy given the strategy profile of the others. They also demonstrate how to perform automated mechanism design in this setting, under the assumption that the mechanism designer is also a regret-minimizer. This paper is good example of defining and exploring a non-expected utility decision criterion.

Hyafil and Boutilier [2006] present an algorithm that uses a minimax-regret decision criterion in conjunction with partial revelation auctions. The idea is that full revelation may be too expensive (in terms of computation and/or communication), so the auctioneer instead queries the bidders to narrow the space of possible type profiles, and then chooses the allocation and payments to minimize efficiency and payment regret with respect to the possible types. This allows a mechanism to guarantee that it will be δ-efficient and ϵ-incentive-compatible, within arbitrary bounds, usually with a vastly-reduced need to reduce type uncertainty. In practice, if δ and ϵ are sufficiently small, then δ-incentive-compatible is the same as fully incentive-compatible.

Halpern and Pass [2008] propose the *iterated regret minimization* criterion. Under the assumption that agents minimize regret with respect to their belief that other agents are also regret-minimizers, they first remove all non-regret-minimizing strategies from consideration for each agent; then, considering only those strategies that have survived for each agent, they perform another round; and so on until no further strategies can be eliminated. On many games (e.g. Traveller's Dilemma) the iterated regret minimization solution gives results extremely close to those of actual experiments. However, on others (e.g., the coordination-game example given in example 3.15 of the paper), the results are actually worse than those of Nash equilibrium. This solution concept does not seem entirely convincing to me, since the k-th iteration of deletion relies on a kth-level belief in the other agents' being a regret-minimizer.

Stoye [2007] provides axiomatizations of several minimax regret decision criteria. He pays particular attention to a rule that minimizes expected regret over sets of prior beliefs (whether those beliefs are behavioural, i.e. endogenous, or a feature of the environment, i.e. exogenous). This rule is very similar to the maxmin rule of [Gilboa and Schmeidler, 1989], and indeed several results are imported from [Gilboa and Schmeidler, 1989]. A link is described between a preference relation formulation of the decision rule and a choice-correspondance formulation.

2.2 Loss aversion

Kahneman and Tversky [1979] model loss-aversion by assuming that utilities are translated into reference-dependent valuations. The *loss aversion coefficient* represents how much more or less losses (i.e., utilities below the reference point) are weighted compared to gains (utilities above the reference point). Related is the *reflection effect*, where people tend to be risk-averse in gains and risk-seeking in losses. In other words, the value function is concave in gains and convex in losses, giving it an S shape.

Tversky and Kahneman [1992] estimate the following functional form and parameters for monetary prospects (when the reference point is 0):

$$v(x) = \begin{cases} x^{\alpha} & \text{for } x \geq 0 \\ (-\lambda)(-x)^{\alpha} & \text{for } x < 0 \end{cases}$$

where

$$\alpha = 0.88$$
$$\lambda = 2.25$$

Shalev [2000] introduces the concept of the *loss aversion equilibrium*. This is an equilibrium where each agent is assigned his own loss-aversion coefficient. A myopic loss-aversion equilibrium is one where each agent evaluates his utilities using his starting or "root" reference point; a non-myopic loss-aversion equilibrium is one where each utility is valued relative to the new reference point at each information set. Shalev uses a linear loss-aversion valuation function, which makes it impossible to represent differing risk attitudes between losses and gains. Shalev [2002] also investigates loss-aversion in the context of bargaining, finding a unique "stable and self-supporting" solution to bargaining problems where either or both agents are loss-averse.

2.3 Non-linear decision weights

Kahneman and Tversky [1979] propose a model in which each probability is converted to a decision weight independently. Under this model, stochastically dominated alternatives may be preferred, which does not match the experimental data; in the model of Kahneman and Tversky [1979] this is worked around through an initial "editing" phase, where agents edit prospects into a more tractable form and apply heuristics such as detection of dominated alternatives.

Prospect-theoretic agents under-weight moderate- or high-probability events that are not certain, and over-weight extremely low-probability events.

Cumulative prospect theory [Tversky and Kahneman, 1992] solves this problem by applying a non-linear transformation to the cumulative distribution rather than to each probability directly. Under this formulation, a stochastically-dominated option will never be preferred. (On the other hand, this makes it more difficult to account for the occasions when stochastic dominance *is* violated).

The CPT decision weight function is piecewise around the reference point. In other words, decision weights over losses are calculated differently from decision weights over gains.

2.4 Ambiguity and non-additive probabilities

A non-additive probability measure P (sometimes referred to as a *capacity*) exhibits uncertainty aversion if it satisfies

$$P(A) + P(B) \leq P(A \cup B) + P(A \cap B). \quad (1)$$

In particular, $P(A) + P(\overline{A})$ may be less than 1.

The *core* of a non-additive probability measure P is the set of all additive probability measures π such that $\pi(A) \geq P(A)$ for all events A. If P exhibits uncertainty aversion, then the core is non-empty.

Dow and Werlang [1994] propose a solution concept for 2-player strategic games that they call *Nash equilibrium under uncertainty*, which is exactly analogous to Nash equilibrium, except that the mixed strategies may be represented by non-additive probability measures. They demonstrate that agents that maximize expected utility relative to non-additive, uncertainty-averse probability measures will choose the maxmin choice even for very small levels of uncertainty aversion. This is very similar to Kreps and Milgrom [1982]'s δ-craziness model, where a small uncertainty that the other agent might play tit-for-tat is enough to produce cooperation in the finitely-repeated Prisoners' Dilemma rather than the backward-induction result. However, where δ-craziness requires that the "crazy" strategy be exogenously specified, NEUU automatically chooses the strategy to be "afraid of" based on what would be most harmful.

Eichberger and Kelsey [2000] extend this solution to arbitrary n-player strategic games; they call this new concept *equilibrium under uncertainty*. They define two parameters for a capacity ν: The *degree of* *confidence* $\gamma(\nu)$ and the *degree of ambiguity* $\lambda(\nu)$, where

$$\gamma(\nu) = \max_{A \subseteq S_{-i}} \nu(A) + \nu(S_{-i} \backslash A)$$
$$\lambda(\nu) = 1 - \min_{A \subseteq S_{-i}} \nu(A) + \nu(S_{-i} \backslash A).$$

Note that $\lambda(\nu) = 0$ implies that ν is additive, but $\gamma(\nu) = 1$ does not.

A *simple capacity* is the "contraction" of an additive probability π, where

$$\nu(E) = \gamma \cdot \pi(E) \quad \forall \text{ events } E.$$

Note that for a simple capacity, $\gamma(\nu) = 1 - \lambda(\nu)$, although this is not true of general capacities. Using simple capacities as a tool, Eichberger and Kelsey [2000] prove that an equilibrium under uncertainty exists for all games and all profiles γ of confidence.

For all games, there is a value $\epsilon > 0$ such that any profile of confidence γ where $\gamma_i \leq \epsilon \, \forall i \in N$ gives rise to an equilibrium under uncertainty that coincides with a profile of maxmin strategies. For all *two-player* games, EUU coincides with NE as $\lambda_i \to 0 \forall i \in N$; this is not true for general n-player games, although it is true for n-players games where each player's beliefs are *independent* and *consistent*. In other words, when players are confident in their beliefs, they will tend to play NE strategies, whereas when players have highly ambiguous beliefs, they will "play it safe" by playing maxmin strategies. Gilboa and Schmeidler [1989] give an axiomatic characterization of a maxmin expected utility decision rule for agents who have a convex set of priors, rather than a single prior. According to [Dow and Werlang, 1994], in many (although not all?) cases, applying this rule over the core C of a non-additive uncertainty-averse probability P is equivalent to directly maximizing the expected utility with respect to P.

Lo [1999] presents a new solution concept for extensive-form games in the presence of ambiguity where agents' beliefs are represented as sets of additive probabilities. A straightforward extension of Nash equilibrium is insufficient for this setting, because MEU preferences are not dynamically consistent under either of the straightforward forms of belief update.[2] Lo's model assumes that each probability in a given agent's belief set has the same support. However, unlike standard Nash equilibrium, two agents i and j are not required to have identical beliefs about any third agent k.

Bewley [2002] sketches a decision theory that represents ambiguity (which he refers to as *Knightian uncertainty*) through a compact set of probability measures Δ. His model includes a notion of the *status quo*, and assumes that the decision maker will choose an action only if it preferred to the status quo under *all* the probability measures in Δ. This amounts to violating the completeness axiom, since if two acts are not strictly preferred to the status quo, they may be incomparable.

Ghirardato et al. [2004] provides a formal description of representing ambiguous preferences under a general set of axioms that characterize Maxmin Expected Utility [Gilboa and Schmeidler, 1989], Choquet Expected Utility[3] [Schmeidler, 1989], and Subjective Expected Utility maxmimization [Savage, 1972]. The axioms are essentially those of [Gilboa and Schmeidler, 1989], except without the Uncertainty Aversion axiom. This paper is very concerned with distinguishing between a decision maker's *perception* of ambiguity (which can be measured by the size of her set of "possible" probability measures) and her *reaction* to ambiguity (i.e., is she ambiguity averse or ambiguity seeking?)

3 Opponent modelling

McKelvey and Palfrey [1992] explains discrepancies between actual results and those predicted by BI in experiments on the centipede game by suggesting that some percentage of the population are actually *altruists* whose utility comes from the total amount of money awarded rather than just their own share. If it is common knowledge that some agents are al-

[2] Sets of probabilities can be updated either by performing Bayesian updating on each measure in the set, or by removing every measure that did not give the event being conditioned on maximal probability.

[3] Only *some* CEU models are consistent with the axioms of [Gilboa and Schmeidler, 1989], whereas *all* CEU models are consistent with those of [Ghirardato et al., 2004].

truists, then immediately playing TAKE is no longer a dominant strategy for "selfish" non-altruists.

Essential elements of their model included: error in actions (i.e., with ϵ probability the agents will make a random move); learning rate (i.e., errors in actions decreases over time); and errors in beliefs (i.e., heterogeneous beliefs about the distribution of altruists).

Stahl and Wilson [1995] propose and experimentally test a richer model with 5 boundedly-rational "archtypical" agents plus a "rational expectations" archetype. Each agent archetype has a probability of making an "action error" (where they act randomly instead of in the way that they "intended" to), plus several other paramters describing their (possibly-incorrect) beliefs about the proportion of each of the other archetypes in the population. The experimental evidence rejects the hypothesis that the rational expectations archetype is present, but is consistent with the boundedly rational types. Costa-Gomes et al. [2001] perform an extremely similar study on 9 archetypes (4 non-strategic and 5 strategic), with a stronger focus on information search behaviours.

A simpler, single-parameter *Cognitive Hierarchy* model is experimentally investigated by Camerer et al. [2004]. "Level 0" agents choose their actions randomly; "Level 1" agents best-respond against Level-0 agents; and Level-k agents best-respond against a population of agents of level $< k$. The model assumes that agents are aware of the correct ratios between agents of lower level, but assume that no other agent reasons at level k or higher. They fit their data on p-beauty contests and similar games using a Poisson distribution. They find that τ is almost always between 1 and 2, and suggest 1.5 is a good place to start for predictions. This single-parameter Poisson model fits the data very nearly as well as a much more complicated 7-parameter model.

3.1 Finite state automata

Freund et al. [1995] discusses algorithms for playing against a specific model of boundedly rational agents represented by boolean formulas, under the specific assumption that no action is irrevocable. Gilboa and Samet [1989] shows that an unboundedly rational agent always has a dominant strategy against a boundedly rational automaton; these strategies often favour the bounded agent. But *knowledge of* (i.e., not just belief in) the bounded agent's boundedness is crucial; it's not good enough for one agent to fake it.

3.2 Heuristics

[Gigerenzer et al., 1999] is a collection of papers on "fast and frugal" heuristics. These heuristics (including single-reason decision making, the recognition heuristic, and others) are argued to be *ecologically rational* in that they provide a good tradeoff between deliberation cost and effectiveness.

References

Bewley, T. (2002). Knightian decision theory. Part I. *Decisions in Economics and Finance*, 25(2):79–110.

Camerer, C. (1998). Bounded Rationality in Individual Decision Making. *Experimental Economics*, 1(2):163–183.

Camerer, C., Ho, T., and Chong, J. (2004). A Cognitive Hierarchy Model of Games. *Quarterly Journal of Economics*, 119(3):861–898.

Costa-Gomes, M., Crawford, V., and Broseta, B. (2001). Cognition and Behavior in Normal-Form Games: An Experimental Study. *Econometrica*, 69(5):1193–1235.

Dow, J. and Werlang, S. (1994). Nash Equilibrium under Knightian Uncertainty: Breaking Down Backward Induction. *Journal of Economic Theory*, 64(2):305–324.

Eichberger, J. and Kelsey, D. (2000). Non-Additive Beliefs and Strategic Equilibria. *Games and Economic Behavior*, 30(2):183–215.

Freund, Y., Kearns, M., Mansour, Y., Ron, D., Rubinfeld, R., and Schapire, R. (1995). Efficient algorithms for learning to play repeated games against computationally bounded adversaries. *Proceedings* *of the Thirty Sixth Annual Symposium on Foundations of Computer Science*, pages 332–341.

Ghirardato, P., Maccheroni, F., and Marinacci, M. (2004). Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory*, 118(2):133–173.

Gigerenzer, G., Todd, P., et al. (1999). *Simple heuristics that make us smart.* Oxford University Press New York.

Gilboa, I. and Samet, D. (1989). Bounded versus unbounded rationality: The tyranny of the weak. *Games and Economic Behavior*, 1(3):213–221.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153.

Halpern, J. and Pass, R. (2008). Iterated regret minimization: A more realistic solution concept. In *World Congress of Game Theory08.*

Hey, J. and Orme, C. (1994). Investigating Generalizations of Expected Utility Theory Using Experimental Data. *Econometrica*, 62(6):1291–1326.

Hyafil, N. and Boutilier, C. (2004). Regret minimizing equilibria and mechanisms for games with strict type uncertainty. *Proceedings of the 20th conference on Uncertainty in artificial intelligence*, pages 268–277.

Hyafil, N. and Boutilier, C. (2006). Regret-based Incremental Partial Revelation Mechanisms. *PROCEEDINGS OF THE NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE*, 21(1):672.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Kreps, D. and Milgrom, P. (1982). Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory*, 27:245–272.

Lo, K. (1999). Extensive Form Games with Uncertainty Averse Players. *Games and Economic Behavior*, 28(2):256–270.

McKelvey, R. and Palfrey, T. (1992). An Experimental Study of the Centipede Game. *Econometrica*, 60(4):803–836.

Savage, L. (1951). The Theory of Statistical Decision. *Journal of the American Statistical Association*, 46(253):55–67.

Savage, L. (1972). *The Foundations of Statistics.* Dover Publications.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587.

Shalev, J. (2000). Loss aversion equilibrium. *International Journal of Game Theory*, 29(2):269–287.

Shalev, J. (2002). Loss Aversion and Bargaining. *Theory and Decision*, 52(3):201–232.

Stahl, D. and Wilson, P. (1995). On Players' Models of Other Players: Theory and Experimental Evidence. *Games and Economic Behavior*, 10(1):218–254.

Stoye, J. (2007). Axioms for Minimax Regret Choice Correspondences. Technical report, Northwestern University.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323.

Von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior.* Wiley.

Wald, A. (1950). Statistical Decision Functions. *New York.*