

Philippe MONGIN

"Introduction to decision theory and social ethics"

California Institute of Technology, Division of Humanities and Social Sciences,
Spring 2003, SES/PI 169

(Selected Topics in Science, and Ethics, and Society)

These lectures will present selective topics in the three areas of individual decision theory, social choice theory, and social ethics, more or less in that order. The emphasis will be on the conflicting senses of rationality in decision-making, at both the individual and collective level, and the various interpretations of distributive justice when it comes to social ethics. The theories above are oriented towards resolving philosophical issues and providing foundations for more applied fields, like decision-aid or welfare economics. These theories are abstract in character and involve a good deal of mathematics, though mostly at the elementary level. Knowledge of high-school mathematics or/and ease with symbolic methods such as those of logic are prerequisites for taking this course; otherwise, it is self-contained. The topics are described below with reading lists. Items with * provide an entry to the main issues. Items with # are advanced (either mathematically or philosophically). Reading between classes is strongly encouraged.

1. The meaning of utility, preference and choice

Historical notes on the utility concept in the 19th and 20th century : utilitarians, marginalists, and neo-classicals. Cardinal vs. ordinal utility. The behavioral or choice-based conception of utility ("revealed preference theory"). Current conceptions of preference, utility and choice; connections with individual well-being. What conceptions seem appropriate for normative inquiries like the present ones. Outline of the lectures.

J. Broome (1991), "Utility", *Economics and Philosophy*, 7, 1-12.

T. Ellingsen (1994), "Cardinal Utility : A History of Hedonimetry", in M. Allais and O. Hagen (eds), *Cardinalism*, Kluwer, 1994, p. 105-165.

#J. Griffin (1986), *Well-Being*, Clarendon, esp. ch. 1-3.

*D.M. Hausman and M.S. McPherson, *Economic Analysis and Moral Philosophy*, Cambridge University Press, ch.6.

*P. Mongin and C. d'Aspremont (1998), "Utility Theory and Ethics", in S. Barbera, P. Hammond, C. Seidl (eds), *Handbook of Utility Theory*, Kluwer, 1, p. 371-481; section 2.

*A.K. Sen (1973), "Behaviour and the Concept of Preference", *Economica*, 40, 241-259 ; also in A.K. Sen, *Choice, Welfare and Measurement*, Blackwell , 1982, ch.2.

A.K. Sen (1980-81), "Plural Utility", *Proceedings of the Aristotelian Society*, 81, 193-215.

2. Rationality and the optimizing model of decision

Rationality broader than instrumental rationality, and instrumental rationality broader than optimization. Rationality is both a normative and predictive concept. A basic example of the optimizing conception of rational decision: the neo-classical model of the consumer. Debreu's theorem as an example of representations theorems in decision theory. Normative arguments about transitive and complete preferences. An outline of the formalism of choice functions.

K.J. Arrow (1959), "Rational Choice Functions and Orderings", *Economica*, 26, 121-127.

#G. Debreu (1959), *Theory of Value*, Yale University Press, ch. 4, esp. 4.6.

P. Fishburn (1970), "The Irrationality of Transitivity in Social Choice", *Behavioral Science*, 15, 119-123.

J.C. Harsanyi (1976), *Essays on Ethics, Social Behavior and Scientific Explanation*, Reidel, ch. 6.

A. Mas-Colell, M.D. Whinston and J.R. Green (1995), *Microeconomic Theory*, Oxford, ch.1 and 3B.

*P. Mongin (2000), "Does Optimization Imply Rationality?", *Synthese*, 124, 73-111 (except for section 3).

A.K. Sen (1970), *Collective Choice and Social Welfare*, North Holland, ch. 1*.

*H. Varian (2000), *Microeconomic Analysis*, Norton, ch. 3.

3 . Arrow's impossibility theorem and a glance at social choice theory

Arrow's impossibility theorem, and its interpretations in terms of either collective decision-making, or the social welfare concept, or individual decision-making with "multiple selves". Normative assessments of the five conditions, depending on which interpretation is adopted. The welfare economics implications of the theorem, i.e., interpersonal comparisons are unavoidable. Social choice theory explores this and other ways out.

K.J. Arrow (1963), *Social Choice and Individual Values*, Yale University Press (1st ed. 1951).

C. Blackorby, D. Donaldson and J. Weymark (1984), "Social Choice Theory with Interpersonal Utility Comparisons: A Diagrammatic Introduction", *International Economic Review*, 25, 327-356.

P. J. Hammond (1991), "Interpersonal Comparisons of Utility", in J. Elster and J.E. Roemer (eds), *Interpersonal Comparisons of Well-Being*, Cambridge University Press, 200-254.

#S. Hurley (1989), *Natural Reasons*, Oxford University Press, ch. 12.

*A.K. Sen (1970), *Collective Choice and Social Welfare*, North Holland, ch. 2, 3, 4.

A.K. Sen (1982), *Choice, Welfare and Measurement*, Cambridge, ch. 11-12.

*A.K. Sen (1999), "The Possibility of Social Choice", *American Economic Review*, 79, 349-378.

4. Risk and expected utility theory

The notion of a risky prospect, or a lottery, contrasted with those of a sure option and of an uncertain prospect. What it means mathematically for a prospect to be riskier than another. The expected utility of a lottery, and the von Neumann-Morgenstern (VNM) axioms and representation theorem. Basic applications of the framework, especially to the measurement of attitudes toward risk. Can VNM utility be given a cardinal interpretation?

P.C. Fishburn (1970), *Utility theory for for Decision-Making*, Wiley, ch.8.

#P.C. Fishburn (1982), *The Foundations of Expected Utility*, Reidel.

C. Gollier (2001), *The Economics of Risk and Time*, MIT Press, ch. 2 and 3.

*D.M. Kreps (1988), *Notes on the Theory of Choice*, Westview, ch. 4-5-6.

#M. Rothschild and J. Stiglitz (1970), "Increasing Risk, I: A Definition", *Journal of Economic Theory*, 3, 66-84.

5. Expected utility : a normative assessment

Allais's paradox and further experimental evidence against expected utility theory, i.e., the common consequence, common ratio, and evaluation effects. How normatively compelling are these paradoxes? A parade from expected utility theory: the dynamic consistency argument. More troubling paradoxes: Kahneman and Tversky's reflection effect and the framing of decisions. The preference reversal phenomenon.

*D. Kahneman and A. Tversky (1979), "Prospect Theory : An Analysis of Decision Under risk", *Econometrica*, 47, 263-291.

E.F. MacClennen (1983), "Sure-Thing Doubts", in B. Stigum and F. Wenstop (eds), *Foundations of Utility and Risk Theory with Applications*, Reidel, 1983. Also in P. Gärdenfors and N.E. Sahlin (eds), *Decision, Probability, and Utility*, Cambridge University Press, 1986, ch. 10.

*M.J. Machina (1983), "Generalized Expected Utility Analysis and the Nature of Observed Violations of the Independence Axiom", in B. Stigum and F. Wenstop (eds), *Foundations of Utility and Risk Theory with Applications*, Reidel, 1983. Also in P. Gärdenfors and N.E. Sahlin (eds), ch. 12.

M.J. Machina (1987), "Choice Under Uncertainty. Problems Solved and Unsolved", *Journal of Economic Perspectives*, 1, 121-154.

M.J. Machina (1991), "Dynamic Consistency and Non-Expected Utility", in M. Bacharach and S. Hurley (eds), *Foundations of Decision Theory*, Blackwell, 39-91.

*A. Tversky and R. Thaler (1990), "Preference Reversals", *Journal of Economic Perspectives*, 4, 201-211.

6. Subjective probability and subjective expected utility theory

Subjective vs. objective probability. 20th century Bayesianism, or how subjective probability became entangled with subjective expected utility theory. Savage's axioms and representation theorem. Normative assessment of the axioms, especially the sure-thing principle and the event-independence axioms. A hint of the more general models, with emphasis on the problem of uniquely defining subjective probability.

E. Karni (1996), "Probabilities and Beliefs", *Journal of Risk and Uncertainty*, 13.

*D.M. Kreps (1988), *Notes on the Theory of Choice*, Westview, ch. 4, 7 and 8.

H. Kyburg (1968), "Bets and Beliefs". Reprinted in B. Stigum and F. Wenstop (eds), ch.6.

#L. Savage, (1972), *The Foundations of Statistics*, Wiley (1st ed. 1954), ch. 2-5.

7. Collective Bayesianism and modern utilitarianism

Today's utilitarianism is often Bayesian. Harsanyi's aggregation theorem, and how to improve on it. Sen's objection and a reply to it. The more severe objection created by differing subjective probabilities in a collective context: a sample of impossibility theorems, one for each brand of Bayesianism. *Ex ante* vs. *ex post* schools of social preference. Major arguments for each school. What is left of utilitarianism in the end?

P.D. Hammond (1982), "Utilitarianism, Uncertainty and Information", in A.K. Sen and B. Williams (eds), *Utilitarianism and Beyond*, Cambridge University Press, ch. 4.

J.C.Harsanyi (1976), *Essays on Ethics, Social Behavior and Scientific Explanation*, Reidel, ch. 2.

*J.C. Harsanyi (1977), "Morality and the Theory of Rational Behavior", *Social Research*. Reprinted in A.K. Sen and B. Williams (eds), ch. 2.

P. Mongin (2002), "Impartiality, Utilitarian Ethics, and Collective Bayesianism", *Cahiers du Laboratoire d'économétrie* 2002-030.

*P. Mongin and C. d'Aspremont (1998), "Utility Theory and Ethics", in S. Barbera, P. Hammond, C. Seidl (eds), *Handbook of Utility Theory*, Kluwer, 1, p. 371-481; section 5.

J. Weymark (1991), "A Reconsideration of the Sen-Harsanyi Debate on Utilitarianism", in J. Elster and J.E. Roemer (eds), *Interpersonal Comparisons of Well-Being*, Cambridge University Press, 255-320.

8. The impartiality approach to social ethics

The notion of an ethically impartial observer and its various decision-theoretic reconstructions. Two rules of preference under the "veil of ignorance", i.e., Rawls's leximin vs. Harsanyi's mean utilitarianism. A genuine Bayesian reconstruction, and where it stops. Hint of the recent modellings of impartiality. The clash between utilitarian and egalitarian social ethics more broadly considered.

R.M. Hare (1976), "Ethical Theory and Utilitarianism", in H.D. Lewis (ed.), *Contemporary British Philosophy*, George Allen and Unwin, p. 113-131.

J.C.Harsanyi, J. (1976), *Essays on Ethics, Social Behavior and Scientific Explanation*, Reidel, ch. 1, 2 and 5.

#E. Karni (1998), "Impartiality : Definition and Representation", *Econometrica*, 66, 1405-1415.

#P. Mongin (2000), "The Impartial Observer Theorem of Social Ethics", *Economics and Philosophy*, 17, 147-179.

*P. Mongin and C. d'Aspremont (1998), "Utility Theory and Ethics", in S. Barbera, P. Hammond, C. Seidl (eds), *Handbook of Utility Theory*, Kluwer, 1, p. 371-481; section 6.

J. Rawls (1971), *A Theory of Justice*, Harvard University Press, ch. 3.

*A.K. Sen (1970), *Collective Choice and Social Welfare*, North Holland, ch. 9.

A.K. Sen ((1974), "Rawls versus Bentham : An Axiomatic Examination of the Pure Distribution Problem", *Theory and Decision*, 4, 301-309.

9. Bargaining and social ethics

Nash's bargaining model and its later variants. What normative sense can be made of them, and whether they should be related to an underlying non-cooperative game. Gauthier's application of bargaining theory to the original position. More on the reconstruction of social norms in terms of self-interest. Do the value concepts of cooperative game theory have anything to contribute to social ethics?

L . Bovens (1987), "On Arguments from Self-Interest for the Nash Solution and the Kalai Egalitarian Solution to the Bargaining Problem", *Theory and Decision*, 23, p. 231-260.

K. Binmore (1994), *Game Theory and the Social Contract*, 1, Cambridge (Mass.), MIT Press.

#D.Gauthier (1986), *Morals by Agreement*, Clarendon Press, ch. 1 and 5.

#H. Moulin (1988), *Axioms of Cooperative Decision-Making*, Cambridge University Press, ch. 3.

J.F. Nash (1950), "The Bargaining Problem", *Econometrica*, 18, 155-162.

*A. K. Sen (1970), *Collective Choice and Social Welfare*, North Holland, ch. 8-8*.

*Various authors in P. Vallentyne (ed.), *Contractarianism and Rational Choice*, Cambridge University Press, 1991, ch.1,2,3.