RATIONAL CHOICE LECTURE 1

1. Rational choice theory like any science deals not with reality but with an idealized model of it: ideally rational agents, their beliefs and desires. The ultimate aim of decision theory and game theory is to say something about how human agents (a) do in fact and (b) should make choices in the light of given beliefs and desires. Here we'll focus on (a), which is the main source of applications in social science.

2. The basic assumption behind (a) is that rational persons act on their beliefs and desires. But this seems to make prediction impossible: we cannot observe your beliefs and desires directly but only what you do; so how can we systematically predict how people will act—how, for instance, they will respond to a change in the tax system, or the imposition of a new law? Savage (1972) showed was that on certain assumptions about agents, we can in extract quite detailed information about these mental states from bits of their behaviour, from which we can then predict other bits of behaviour.

3. The basic idea is that each agent has a set of **preferences**. We can think of these as a relation amongst possible acts. I'll write **P** $\succ$ **Q** to say that the agent prefers P to Q. The point of these is that they can be extracted more or less straightforwardly from behaviour: **you prefer P to Q if you always choose P when Q is the only alternative**. (See Sen 1977 for critical comments on this.) We assume that the relation $\succ$ is (a) negatively transitive; (b) asymmetric. (a) and (b) imply that $\succ$ is both transitive and acyclic; given a preference for more money over less, these properties have an economic justification. (For discussion of this style of justification see Anand 1993 ch. 4 and Binmore 2009: 17-19.)

4. The mental matrix from which preferences emerge cannot simply be beliefs and desires. For it may be that A takes an umbrella and B does not, even though neither A nor B (fully) *believes* that it is going to rain, and both A and B *desire* not to get wet. The necessary refinement replaces beliefs and desires with measurements of how *strongly* you believe and want something i.e. subjective probabilities and utility. Your subjective probability in a proposition P, written Pr (P), is a function that takes a value between 0 and 1 inclusive, where 0 corresponds to complete rejection and 1 corresponds to total certainty. Your utility for a proposition, U (P), takes any real value, with U (P) > U (Q) just in case P represents a better outcome (for you) than Q.

5. But we cannot directly observe credences and utilities but only actual choices, and these seem to leave one's mental state systematically underdetermined. For instance, that I choose not to take my umbrella might mean that I am confident of dry weather and don't want to get wet, *or* that I am confident of rain but *like* getting wet. How can we tell?

6. What Ramsey (1926 [1990]) saw, and Savage showed systematically, was that given *enough* facts about people's behaviour, we can simultaneously pin down subjective probabilities *and* their utilities. For a simple example of this, consider the following three bets:

	Rain tomorrow	**No rain tomorrow**
Bet 1	\$1	A kick
Bet 2	A kick	\$1
Bet 3	A kick	A kick

Suppose that in a two-way choice you'd (a) take Bet 2 over Bet 3 and (b) take Bet 1 over Bet 2. Fact (a) shows that you prefer getting \$1 to getting a kick, so U (\$1) > U (kick). Fact (b) shows that you think rain tomorrow is more likely than no rain, so Pr (Rain) > Pr (No rain). So we have managed to say something quite definite about your mental state on the basis of entirely observable facts about your behaviour.

7. More generally, Savage showed that if your preferences satisfy certain conditions including negative transitivity and asymmetry, then there is a unique subjective probability function, and a (nearly) unique utility function such that your behaviour *maximizes expected utility*, in the sense that for any options A and B we have:

$$(7.1)\ A \succ B \text{ iff } \Sigma_{S \in \mathbf{S}} \Pr(S)\, U(A \wedge S) > \Sigma_{S \in \mathbf{S}} \Pr(S)\, U(B \wedge S)$$

(7.1) is the fundamental doctrine of standard decision theory.

8. First, the terms A and B denote possible actions that you can take. These are the row headings that appear in the left-hand columns of a decision table e.g. as at no. 6. Second, **S** denotes the set of possible states of the world that matters to you. For instance, in the Table in no. 6, **S** = {Rain tomorrow, No rain tomorrow}. So each $S \in \mathbf{S}$ is one possible such state of the world, or 'state of nature'. **S** must be chosen so that for each act A that you can perform, and for any $S \in \mathbf{S}$, the outcome $A \wedge S$ captures everything that matters to you about the problem that you are facing. So the content of equation (7.1) is that you evaluate an option A by taking a weighted average over the utilities of the outcomes in which A might eventuate, where the weights are your subjective probabilities for the sates of nature which, together with A, yield that utility. This will become clearer if we look at an example.

9. Suppose that I offer you a bet on the next toss of a coin: you win \$1 if it lands heads and lose \$1 if it lands tails. We can represent your choice as follows:

RATIONAL CHOICE LECTURE 1

	S_1: Heads	S_2: Tails
A: Bet	\$1	-\$1
B: No bet	\$0	\$0

The options are A and B. The states of nature are **S** = {Heads, Tails}. Suppose that your utility for \$n is just n, so that

$U(A \wedge S_1) = U(\$1) = 1$
$U(A \wedge S_2) = U(-\$1) = -1$
$U(B \wedge S_1) = U(\$0) = 0$
$U(B \wedge S_2) = U(\$0) = 0$

So by (7.1) we have:

(9.1) $A \succ B$ iff Pr (Heads) – Pr (Tails) > 0

So you will take the bet iff you are more confident that the coin will land heads than that it will land tails. More generally, the idea is that people do in fact behave as expected utility maximizers.

Exercises

1. You are about to go out and you need to choose whether to take an umbrella. Your subjective probability that it will rain is 0.6, and 0.4 that it will not rain. The outcomes are {Rain ∧ umbrella, No rain ∧ umbrella, Rain ∧ no umbrella, No rain ∧ no umbrella} and their utilities are 0.5, 0.5, Y and 1 respectively. What is the maximum value of Y at which you prefer to take the umbrella?

2. You are about to go out and you need to choose whether to take an umbrella. Your subjective probability that it will rain is p, and 1-p that it will not rain. The outcomes are {Rain ∧ umbrella, No rain ∧ umbrella, Rain ∧ no umbrella, No rain ∧ no umbrella} and their utilities are 0.5, 0.5, -2 and 1 respectively. What is the minimum value of p at which you prefer to take the umbrella?

References

Anand, P. 1993. *Foundations of Rational Choice under Risk*. Oxford: OUP.
Binmore, K. 2009. *Rational Decisions*. Princeton: Princeton, UP.
Ramsey, F. P. 1926 [1990]. Truth and probability. In his *Philosophical Papers*, ed. D. H. Mellor. Cambridge: CUP 1990: 52-94.
Savage, L. 1972. *Foundations of Statistics*. New York: Dover.
Sen, A. 1977. Rational fools: a critique of the behavioural foundations of economic theory. *Philosophy and Public Affairs* 6: 317-44.