RATIONAL CHOICE LECTURE 3

1. Savage's decision theory has one individual agent facing a natural world that ignores his desires or beliefs. *Game theory* seeks to model individual agents facing not nature but *other* agents, whose decisions therefore depend on their opinions about each other (and their opinions about each others' opinions about each other, and…).

2. A game is an interaction between n (≥ 2) *players*. Each player j must choose a *strategy* from some set $\mathbf{A_j}$. There is a set **Z** of possible prizes on which each player has a utility function. The game is then defined as a function from $\mathbf{A_1} \times \mathbf{A_2} \times \ldots \times \mathbf{A_n}$ to $\mathbf{Z}^n$. More informally, a complete specification of everyone's strategy (a strategy profile) determines the utility of the outcome for *each* player.

3. The simplest (*normal* form) representation of a 2-player game in which player 1 ('Alice') has m options and player 2 ('Bob') has n options is an m x n matrix in which ordered pairs specify the utility to each player of the corresponding outcome. E.g. Alice and Bob drive in opposite directions along the same street. All that matters is that they don't collide, which will happen if and only if both drive on the left or both on the right. Then the game (= its payoff matrix) looks like this:

	Bob drives on left	Bob drives on right
Alice drives on left	(1, 1)	(0, 0)
Alice drives on right	(0, 0)	(1, 1)

Table 3.1: Co-ordination

The convention is that the first entry in a cell is the utility for Alice and the second that for Bob. The top left entry in Table 3.1 (1, 1) says that if both Alice and Bob drive on the left then Alice's utility is 1 and so is Bob's, an outcome that both prefer to that on the *bottom* left, where Alice drives on the *right* and Bob on the left, in which case the cars collide and both players get utility 0.

4. Many interpersonal interactions can be thus represented as games. Consider 'Chicken', in which two cars drive towards one another, the loser being the one that swerves first. (If neither swerves there is a collision, this being the worst possible outcome for both.)

	Bob swerves	Bob doesn't swerve
Alice swerves	(2, 2)	(1, 3)
Alice doesn't swerve	(3, 1)	(0, 0)

Table 3.2: Chicken

Each player can only get the best possible outcome by not swerving; but doing so runs the risk of a collision, bottom right. It is possible to understand the Cuban missile crisis and other political standoffs in these terms (Brams 1990 s. 4.3).

5. In Prisoners' Dilemma, Alice and Bob have been arrested and are being interrogated. The police say to each: if you each betray one

another then you'll get a heavy sentence (10 years). If one betrays and the other keeps quiet, the first goes free and the second will be shot. If you both keep quiet you'll both get a light sentence (5 years).

	Bob betrays	Bob keeps quiet
Alice betrays	(-10, -10)	(0, -100)
Alice keeps quiet	(-100, 0)	(-5, -5)

Table 3.3: Prisoner's dilemma

The decision to advertise is a plausible economic interpretation of this game (Corfmann 1994). Advertising gives each firm an advantage whatever the other firm is doing; but if both advertise then *both* are worse off than if neither does (the outcome is 'Pareto inefficient'). Philosophers model addiction (Ainslie and Monterosso 2003) and Hobbes's 'State of Nature' on the PD (Hobbes 2008 Book I; Rawls 1999: 238, also the references at 238 n. 8.)

6. In PD each player is better off betraying the other whatever the other does. This relation between the strategies, that betrayal is at least as good as keeping quiet *whatever* the other players' strategy, is (like the analogous situation in decision theory) called *dominance*: betrayal dominates keeping quiet (for player i). A rational player will never play a dominated strategy (*whatever* the players agree in advance).

7. The most important type of 'solution' to a game is a *Nash equilibrium*. A strategy profile S* (i.e. a function from players to strategies) is a Nash equilibrium if it assigns to each player the best response (for him) given all of the others' strategies in S*. E.g. Table 1 has two equilibria i.e. both drive left or both drive right. In PD there is one equilibrium i.e. both betray. If Nash equilibria are rationally mandatory, both players betray in the PD; thus Hobbes's argument for an absolute sovereign.

8. Some games have *no* 'pure' Nash equilibria. In 'Penalty' – more commonly known as 'Matching Pennies' – Alice is taking a penalty and Bob is goalkeeper. Alice has to kick left or right; Bob has to jump left or right.

	Bob: jump left	Bob: jump right
Alice: kick left	(1, -1)	(-1, 1)
Alice: kick right	(-1, 1)	(1, -1)

Table 3.4: Penalty

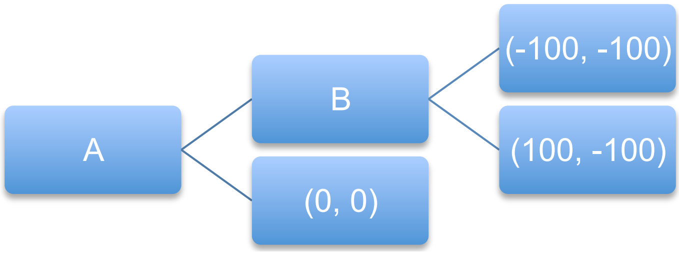
To solve this case we introduce mixed equilibria. As well as the 'pure' strategies that I mentioned we allow the players to choose a random strategy that assigns any probability at all to each 'pure' strategy. Then what Nash showed was that *any* game has at least one Nash equilibrium in these mixed strategies (Nash 1951). For instance, in Table 4 there is a mixed equilibrium in which each player plays each pure strategy with probability 0.5.

9. In these games both players choose simultaneously. But we can also use game theory to represent situations in which one player moves first and the other player sees this before choosing how to move. Thus consider a nuclear standoff in which only Alice's side has the capacity to strike first and Bob must choose whether to retaliate if attacked. We *could* represent the situation like this:

	B: retaliate if struck	B: don't retaliate
A: strike	(-100, -100)	(100, -100)
A: don't strike	(0, 0)	(0, 0)

Table 3.5: First Strike

But we can represent it better in the following *extensive* form:



We can see from both diagrams that if A strikes then B has no reason to strike back. If A knows this then A has an incentive to strike. To prevent this, B might wire the missiles to fire automatically in case of attack. But this only works if A *knows* that B has done this (that was the point of the doomsday machine in *Dr Strangelove*).

Exercises

1. Identify any Nash equilibria in Table 3.2.
2. In the Ultimatum Game, Alice announces a number n between 0 and 100. Bob has to choose whether to accept or reject. If Bob accepts then Alice gets $n and Bob gets $(100-n). If Bob rejects then both players get $0. What are the Nash equilibria for this game?

References

Ainslie, G. and J. R. Monterosso. 2003. Hyperbolic discounting as a factor in addiction. In R. Vucinich and N. Heather (eds), *Choice, Behavioural Economics and Addiction*. Oxford: Pergamon: 35-62.

Brams, S. J. 1990. *Negotiation Games*. London: Routledge.

Corfman, K. P. 1994. The prisoner's dilemma and the role of information in setting advertising budgets. *J. Adv*. 23: 35-48.

Hobbes, T. 2008 [1651]. *Leviathan*. Ed. J. C. A. Gaskin. Oxford: OUP.

Maynard Smith, J. 1982. *Evolution and the Theory of Games*. Cambridge: CUP.

Nash, J. 1951. Non-cooperative games. *Ann. Math*. 54: 286-95.

Rawls, J. 1999 [1971]. *A Theory of Justice*. 2d ed. Cambridge, Mass.: Harvard UP.