OR 3: Chapter 14 - Stochastic games

Recap

In the previous chapter:

- We considered games of incomplete information;
- Discussed some basic utility theory;
- Considered the principal agent game.

In this chapter we will take a look at a more general type of random game.

Stochastic games

Definition

A stochastic game is defined by:

- X a set of states with a stage game defined for each state;
- A set of strategies $S_i(x)$ for each player for each state $x \in X$;
- A set of rewards dependant on the state and the actions of the other players: $u_i(x, s_1, s_2)$;
- A set of probabilities of transitioning to a future state: $\pi(x'|x, s_1, s_2)$;
- Each stage game is played at a set of discrete times t.

We will make some simplifying assumptions in this course:

1. The length of the game is not known (infinite horizon);
2. The rewards and transition probabilities are not dependent;
3. We will only consider strategies called **Markov strategies**.

Definition

A strategy is call a **Markov strategy** if the behaviour dictated is not time dependent.

Example

Consider the following game with $X = \{x, y\}$:

- $S_1(x) = \{a, b\}$ and $S_2(x) = \{c, d\}$;
- $S_1(y) = \{m\}$ and $S_2(x) = \{n\}$;

We have the stage game corresponding to state x:

$$\begin{pmatrix} (8,4) & (5,3) \\ (1,5) & (2,6) \end{pmatrix}$$

The stage game corresponding to state y:

$$\begin{pmatrix} (0,0) \end{pmatrix}$$

The transition probabilities corresponding to state x:

$$\begin{pmatrix} (.8,2) & (1,0) \\ (1,0) & (1,0) \end{pmatrix}$$

The transition probabilities corresponding to state y:

$$\begin{pmatrix} (0,0) \end{pmatrix}$$

Here is a concise way of representing all this:

We see that the Nash equilibrium for stage game corresponding to x is (a, c) however as soon as the players play that strategy profile they will go to state y which is an absorbing state at which players gain no further utility.

To calculate utilities for players in infinite horizon stochastic games we use a discount rate. Thus without loss of generality if the game is in state x and we

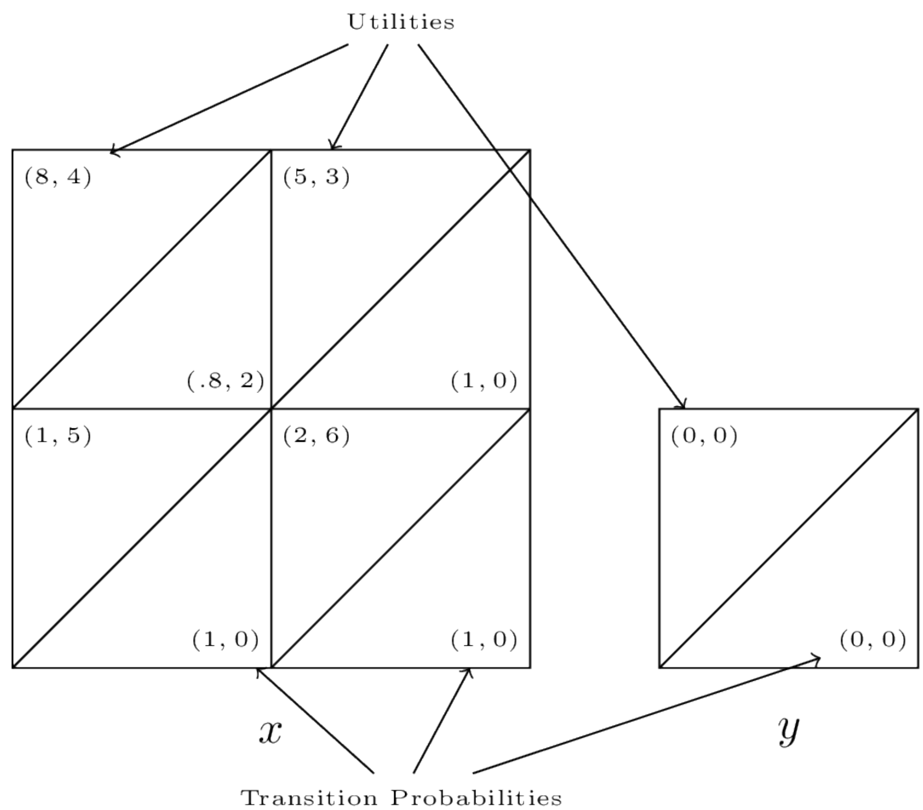
Utilities
(8, 4)
(.8, 2)
(5, 3)
(1, 0)
(1, 5)
(1, 0)
(2, 6)
(1, 0)
(0, 0)
(0, 0)
x
y
Transition Probabilities

Figure 1:

assume that both players are playing σ_i^* then player 1 would be attempting to maximise future payoffs:

$$U_1(r,s) = \left(u_1(x,r,s) + \delta \sum_{x' \in X} \pi(x'|x,r,s)U_1^*(x')\right)$$

where U_1^* denotes the expected utility to player 1 when both players are playing the Nash strategy profile.

Thus a Nash equilibrium satisfies:

$$U_1^*(x) = \max_{r \in S_1(x)} (u_i(x,r,s^*) + \delta \sum_{x' \in X} \pi(x'|x,r,s^*)U_1^*(x')$$
$$U_2^*(x) = \max_{s \in S_2(x)} (u_i(x,r^*,s) + \delta \sum_{x' \in X} \pi(x'|x,r^*,s)U_1^*(x')$$

Solving these equations is not straightforward. We will take a look at one approach by solving the example we have above.

Finding equilibria in stochastic games

Let us find a Nash equilibrium for the game considered above with $\delta = 2/3$.

State x gives no value to either player so we only need to consider state x. Let the future gains to player 1 in state x be u, and the future gains to player 2 in state x be v. Thus the players are facing the following game:

$$\begin{pmatrix} (8+\frac{1}{3}v, 4+\frac{1}{3}u) & (5+\frac{2}{3}v, 3+\frac{2}{3}u) \\ (1+\frac{2}{3}v, 5+\frac{2}{3}u) & (2+\frac{2}{3}v, 6+\frac{2}{3}u) \end{pmatrix}$$

We consider each strategy pair and state the condition for Nash equilibrium:

1. (a,c): $v \leq 21$ and $u \leq 3$.
2. (a,d): $u \geq 3$.
3. (b,c): $v \geq 21$ and $u \leq 3$.
4. (b,d): $5 \geq 2$.

Now consider the implications of each of those profiles being an equilibrium:

1. $8 + v/3 = v \Rightarrow v = 12$ and $4 + u/3 = u \Rightarrow u = 6$ which contradicts the corresponding inequality.

2. $3 + 2u/3 = u \Rightarrow u = 9$.
3. $1 + 2v/3 = v \Rightarrow v = 3$ and $5 + 2u/3 = u \Rightarrow u = 15$ which contradicts the corresponding inequality.
4. The inequality cannot hold.

Thus the unique Markov strategy Nash equilibria is (a, d) **which is not the stage Nash equilibria!**