Computational Game Theory, Fall 2009	December 4

Lecture 10 - Part 2: Recursive games and Nash Equilibrium

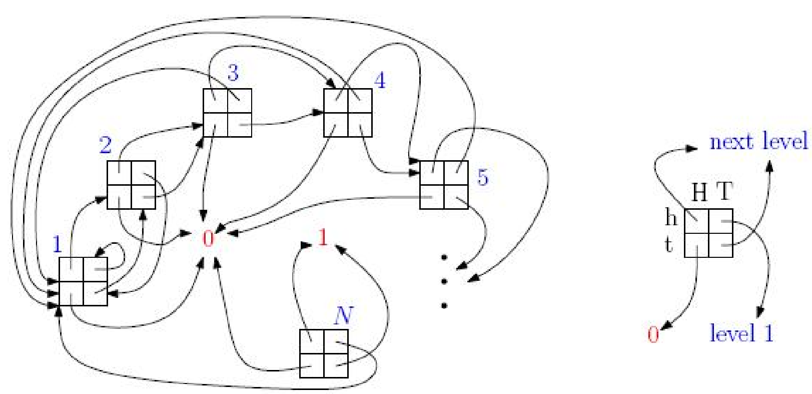
Lecturer: Peter Bro Miltersen — Scribe: Clément Maria

1 Special case 2

The second special case of undiscounted discounted stochastic games we shall look at is the case of non-zero rewards only at absorbing states ("terminals"), but without assuming perfect information. These games were dubbed *recursive games* by Everett in 1957. For this case, Everett showed that the value can be approximated using behavior strategies. That is, $\sup_{x\in\tilde{S}_1}\inf_{y\in S_2} u_1(x,y) = \inf_{y\in\tilde{S}_2}\sup_{x\in S_1} u_1(x,y)$ with $\tilde{S}_i$ being the set of behavior strategies of player i and S_i being the set of general strategies.

However, the "sup" is not a "max", as can be seen in this classical example, due to Everett: Player II hides a penny and Player I must guess if it is heads up or tails up. If Player I guesses correctly, he wins the penny from Player II and the game ends. If Player I incorrectly guesses that the penny is tails up, the game ends with payoff zero to both players. If he incorrectly guesses that it is heads up, the game repeats. If the play continues forever, the payoff to both players is zero. It is easy to see that this game can be modelled as a recursive game with two absorbing states and one non-absorbing state. Given a parameter $\epsilon > 0$, the behavior strategy for Player II of guessing heads up with probability $1-\epsilon$ and tails up with probability ϵ is easily seen to guarantee Player I a probability of $1-\epsilon$ of winning the penny. However, it is easy to see that there is no strategy (neither behavior nor general) for Player I that secures him the penny with probability 1. Therefore, the value of the game is 1 but strategies of Player I can only approximate this value, not achieve it.

The next example is a variant of the above game. The point of this example is to show that to approximate the value of a recursive game, Player I may have to use a behavior strategy that uses extremely small probabilities. The game is called *Purgatory* with N terraces:



In this version, we play with N positions. In each position, the player II chooses between Head and Tail and player I has to guess which one II chose. If I guesses right, we go to next position $(n+1)$ or, if we are in the last position N, we go to the terminal with reward $+1$ (which is equivalent to player I to take the coin). If he is wrong we have two cases :

- if I chooses head whereas II chose tail, we play next the positon number 1 (we come back to the "beginning")
- if it's the contrary, we go to the terminal with reward 0 (which is equivalent to player II to keep the coin)

The game can be simply described as follows: Play the original game described by Everett, but modify the rules so that Player I has to guess correctly N times in a row to win the penny. We can see that if Player I plays all the time h, player II can respond by always hiding tails. The outcome is infinite play, i.e. payoff 0 for Player I. As in the orginal game, Player I should take the "risky" choice of guessing tails occassionally, but in this version of the game, he has to carefully balance the probabilities with which he takes the risky choice. Quantitively, it can be shown that for all $\epsilon > 0$, Player I can ensure a probability of winning of $1 - \epsilon$ if he in position k plays h with probability approximately $1 - \epsilon^{2^{N-k}}$ and t with probability $\epsilon^{2^{N-k}}$. So the value of the game is 1, It can also be shown no strategy that plays tails with probability much higher than $\epsilon^{2^{N-1}}$ in position 1 achieves this guarantee.

Computational problem Given a recursive game G and $\epsilon > 0$, compute a behavior strategy x that achieves the value within ϵ (an ϵ-maximin strategy).

The Purgatory example shows that there if we use standard (binary or decimal) representation of behavior probabilities there can be no polynomial time algorithm for this problem! Simply because the number of digits required to write down the number $\epsilon^{2^{N-k}}$ is exponential.

Open Problem : Find a good succinct non-standard representation of behavior strategies that permits a polynomial size representation of ϵ-maximin strategies.

2 Compute Nash equilibrium in general sum games

Let G be a l-player game in strategic form, the strategy spaces $S_1, S_2, .., S_l$ being finite and the payoff functions $u_i : S_1 \times S_2 \times .. \times S_l \to \mathbb{R}$ (not 0-sum).

Definition 1 *A* **Nash Equilibrium** *is a strategy profile* $(x_1, x_2, .., x_l)$ *so that :*

$$x_i = \arg \max_{x'_i \in \Delta(S_i)} u_i(x'_i, x_{-i})$$

where $x_{-i} = (x_1, .., x_{i-1}, x_{i+1}, .., x_l)$.

Example : Chicken In this game, player I and II have two actions : stay and turn. If the two players stay, they collide and die and lose a lot (here 100). If the two of them turn, nothing happens. If one turns and the other stays, the one who turns is the chicken and loose 1, the other earns 1.

	S	T
s	-100 \ -100	1 \ -1
t	-1 \ 1	0 \ 0

There are two pure Nash equilibria in this example:

$$\begin{pmatrix} \text{player I chooses } s \\ \text{player II chooses } T \end{pmatrix} \text{ or } \begin{pmatrix} \text{player I chooses } t \\ \text{player II chooses } S \end{pmatrix}.$$

Note that Player I has a different payoff in these two equibria. In particular, unlike in the zero-sum case, we can not assign a "value" to a general-sum game. In addition to these two pure Nash equilibria, we have a mixed Nash equilibrium. We can If player II plays S with probability p and T with probability $1-p$, we find the expected pay-off for I :

- if I plays S, the expected pay-off is : $-100p + (1-p)$
- if he plays T : $-p$

In a Nash equilibrium where Player I plays both choices with positive probability, these two values must be equal, so we must have $p = 1/100$. And in fact, we can verify that if Player II does play s with probability 1/100 and Player II plays S with probabiltiy 1/100, we do have a Nash equilibrium.

The interpretation of mixed Nash equilibria is the subject of many philosophical considerations in the game theory literature. It should be noted that intuitively, in a Nash equilibrium of a general-sum game, one can *not* argue that a player is playing a mixture to benefit his payoff. Indeed, if we really are in the equilibrium, Player I expects the same payoff no matter what strategy he plays among the ones to which he assigns positive probability. If we are *not* in the equilibrium it is even less clear what playing by these probabilities achieves[1]. So why on earth is Player I playing by a randomized strategy?

Summing up this discussion, a mixed Nash equilibrium describes a "stable situation" but it is not quite clear how and why such a stable situation should come about or even what the situation "means". Aumann has argued that one should interpret the mixed strategy of Player I to acutally "belong" to Player II and model an aspect of the state of mind of Player I: His *belief* about what Player I will do. This interpretation makes a lot of sense, but it is still not so clear why we would want algorithms *computing* mixed Nash equilibria. It should be noted that the lecturer does not know any "industrial applications" of algorithms for computing Nash equilibrium of general sum games[2], while he knows a lot of industrial applications of algorithms for computing maximin strategies of zero-sum games. However, from a "pure" mathematical (and algorithmic!) point of view, mixed Nash equilibria are extremely interesting, so we shall not be so worried about their philosophical interpretation.

Theorem 2 Nash equilibrium, '51[3]*: For any finite game, a (mixed) Nash equilibrium exists.*

Important information for computer science For 2-player games with rational payoffs, there is a Nash equilibrium using only rational probabilities. This is not the case for 3-player games.

Theorem 3 (Brouwer fixed point theorem) *Let $\mathcal{B}_n = \{x \in \mathbb{R}^n \| ||x|| \leq 1\}$ for the usual Euclidean norm $||.||$*
Let $f : \mathcal{B}_n \to \mathcal{B}_n$ be continuous.
Then $\exists x : f(x) = x$.

[1]Unlike the zero-sum case, where the mixed strategy in a Nash equilibrium is one achieving an optimal guarantee that holds, no matter what Player II does

[2]But he would be very interested in being informed about such examples.

[3]Nobel price

Proof (of the Nash equilibrium theorem)
Let $f : \times_{i=1}^{l} \Delta(S_i) \to \times_{i=1}^{l} \Delta(S_i)$. The set $\times_{i=1}^{l} \Delta(S_i)$ is homeomorphic to a ball.
We define $(f(x_1, .., x_l))_{i,j} = \frac{max(x_{ij}+d_{ij},0)}{\Sigma_{j'=1}^{m_i} max(x_{ij'}+d_{ij'},0)}$
with

- x_{ij} is the probability of j choice of player i
- $d_{ij} = u_i(x1, .., x_{i-1}, e_j, ..x_l) - u_i(x_1, .., x_l)$, where e_j is the distribution that puts all probability mass on j.

Applying the Brouwer fixed point theorem, the function f has a fixed point. This is easily seen to be a Nash equilibrium. □