Matrix games

Given a matrix M the first player chooses a row i and at the same time and independently the second player chooses a column j. The second player pays to the first one the amount $M[i, j]$.

Example. *Matching pennies.*

Stochastic games

$S = \{1, \ldots, n\}$ the set of states,
$\forall s \in S$, $A^1(s), A^2(s)$ are actions of player 1 and 2 respectively at the state s.

In s, player 1 chooses $a^1 \in A^1(s)$ while player 2 chooses simultaneously and independently $a^2 \in A^2(s)$.

Then the immediate reward of player 1 is

$$r(s, a^1, a^2)$$

while the immediate reward of player 2 is $-r(s, a^1, a^2)$ (antagonistic game).

Therefore

$$r : (s, a^1, a^2) \mapsto \mathbb{R}$$

is the immediate reward function.

In other words with each state s we associate a matrix game with the rows indexed by actions of player 1 and columns indexed by the actions of the player 2.

The probability that $s' \in S$ will be the new state when players 1 and 2 choose a^1 and a^2 respectively in s is

$$p(s' \mid s, a^1, a^2)$$

Thus $p(s' \mid s, a^1, a^2)$ is the conditional probability of going to s' from s upon the choice of a^1 and a^2.

Obviously

$$\sum_{s' \in S} p(s' \mid s, a^1, a^2) = 1$$

Rewards

Suppose that $p = (s_1, a_1^1, a_1^2), (s_2, a_2^1, a_2^2), \ldots, (s_i, a_i^1, a_i^2), \ldots$ be an infinite play.

The reward of player 1 can be calculated in different ways.

Discounted games

$$\sum_{i \geq 1} \beta^i r(s_i, a_i^1, a_i^2)$$

$0 < \beta < 1$ is a fixed discount factor.

Mean payoff games

$$\limsup_{n \to \infty} \frac{\sum_{i=1}^{n} r(s_i, a_i^1, a_i^2)}{n}$$

Gambling systems

$$\limsup_{n\to\infty} r(s_i, a_i^1, a_i^2)$$

Parity games

$r(s, a^1, a^2) = r(s) \in \mathbb{N}$ and

$$\text{reward} = \begin{cases} 1 & \text{if} \quad \limsup_i r(s_i) \quad \text{is odd,} \\ 0 & \text{if} \quad \limsup_i r(s_i) \quad \text{is even.} \end{cases}$$

Strategies

A strategy for player 1 gives in the current state s a probability distribution over the set $A^1(s)$ of his actions.

In general this can depend on the whole past history of the play.

A strategy f is *stationary* (positional, memoryless) if it depends only of the current state, for example such a strategy f for player 1 gives the probability $f(s, a^1)$ of choosing the action $a^1 \in A^1(s)$ in the state s.

$$\sum_{a^1 \in A^1(s)} f(s, a^1) = 1 \quad \text{for each state } s.$$

A strategy f is *pure* (deterministic) if f chooses exactly one action with probability 1.

If we fix the initial state s and the strategies f and g for 1 and 2 then we get a probability measure over the set of plays that begin at s. All previous reward functions become (Borel) measurable random variables and we assume that the take the expectation of such a variable as the payoff of player 1. In other words, the task of player 1 is apply to maximize the expectation of the reward function (i.e. his mean gain).

Obviously the task of 2 consists in minimizing the payoff.

If we note $E_{fg}[reward]$ the the expectation of the reward function for fixed f and g then always

$$\underline{v} = \sup_f \inf_g E_{fg}[reward] \leq \inf_g \sup_f E_{fg}[reward] = \overline{v}$$

1 has a strategy assuring him the payoff of at least $\underline{v}$.

2 has a strategy assuring him the loss not exceeding $\overline{v}$.

The game has a **value** if $\underline{v} = \overline{v}$.

Results

Shapley 1953. Discounted games have a value. Each player has a stationary optimal strategy. Game value can be irrational even for rational data.

Mertens Neyman 1981. Mean payoff games have a value. Moreover $v(s) = \lim_{\beta \to 1}(1-\beta)v_\beta(s)$ where $v_\beta(s)$ is the value of the discounted game with the discount factor β. Optimal strategies may need infinite memory.

Maitra, Sudderth 1994. Gambling systems have a value, even for infinite state space. (This implies easily that mean payoff games have a value).

Luca de Alfaro, Rupak Majumdar STOC'01, Parity stochastic games have a value. Optimal strategies may need an infinite memory. Values can be irrational even for rational data (even for the reachability game!)

Example

State 1

(1/2,1/2,0)	(0,0,1)
(0,0,1)	(1/4,3/4,0)

State 2

(0,1,0)

State 3

(0,0,1)

We have three states. In the state 1 player 1 has two actions (corresponding to the choice of the line) and player 2 has also two actions corresponding to the choice of the column. In two other states each player has only one action.

Each triple (p_1, p_2, p_3) gives the probability of going to the state 1, 2 and 3 respectively. Thus for example if players 1 and 2 choose the first action in the state 1 then with probability $1/2$ either we stay at state 1 or with the same probability $1/2$ we go to the state 2

Note that the states 2 and 3 are absorbing.

The aim of player 1 is to reach state 2, which is a particular case of the parity game. The value of this game for state 1 is $\frac{-3+\sqrt{6}}{5}$ (i.e. this is the optimal probability that player 1 can obtain to reach state 2 against the best strategy of player 2).

Perfect information games

$S = S_1 \cup S_2$. In the state $s \in S_i$ only player i has an action to choose, we say that he controls s.

In other words, $A(s)$ is the set of actions of the player i that controls s, $r(s,a)$ immediate reward of 1, $p(s' \mid s,a)$ conditional probability of going to s' upon the choice of a.

[Vrieze 1987] Discounted perfect information games can be solved by a finite number of iterations of a linear programming problem.

Perfect information parity games.

Single player stochastic games - Markov Decision Process

These are games with only one player, $S = S_1$.

Reach literature, see for example the textbook of Puterman, Markov Decision Processes.

Parity MDP were considered in the thesis of Luca de Alfaro. See also Courcoubetis, Yannakakis ICALP 1990.

Deterministic games

Perfect information games such that $p(s' \mid s, a) \in \{0, 1\}$.

Parity games.

Mean payoff deterministic games. Mycielski 1979. Optimal strategies are stationary (and pure). Pseudopolynomial algorithm, Paterson, Zwick.

Perfect information parity games

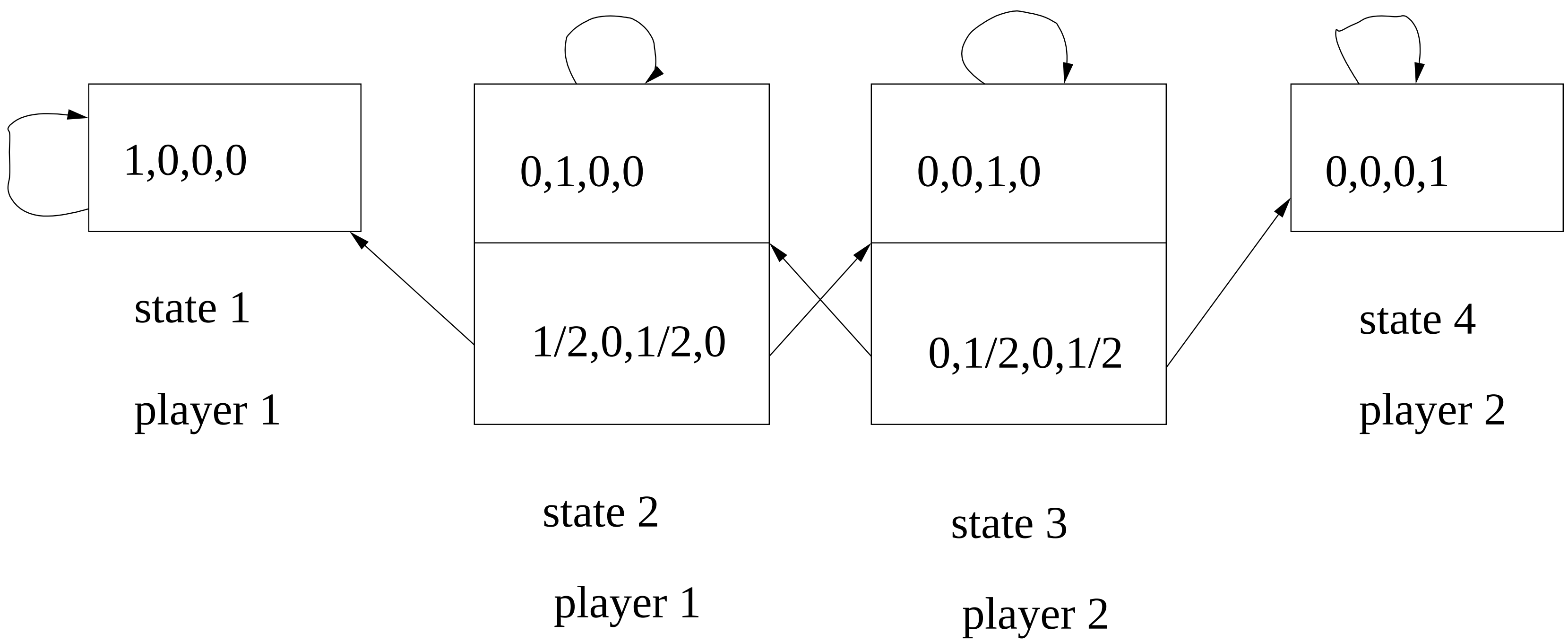
Perfect information reachability games - there are optimal pure stationary strategies for both players, Anne Condon 1992. Quadratic programming. Polynomial algorithm ?

General case.

Theorem. *Both players have optimal pure stationary strategies for perfect information parity games. If the data are rational then the game values are rational.*

The theorem above has several proofs:
A.K. McIver, C.C. Morgan, LPAR 2002, full version 2003
Chatterejee, Jurdziński, Henzinger, SODA 2004.
Zielonka. FOSSACS. 2004.

Example of de Alfaro, Majumdar



Perfect information parity game where the aim of player 1 is to remain inside of $\{s_1, s_3\}$ from some moment onwards.