Beyond the Normal and Extensive Forms

Branislav Bošanský and Michal Pěchouček

Artificial Intelligence Center,
Department of Computer Science,
Faculty of Electrical Engineering,
Czech Technical University in Prague

branislav.bosansky@agents.fel.cvut.cz

November 15, 2016

Previously ... on multi-agent systems.

1. Sequence-Form Representations
2. Solving Extensive-Form Games

Extensive-Form Games

Let's assume that we want to play some normal-form game twice.
For example, rock-paper-scissors:

	R	**P**	**S**
R	$(0,0)$	$(-1,1)$	$(1,-1)$
P	$(1,-1)$	$(0,0)$	$(-1,1)$
S	$(-1,1)$	$(1,-1)$	$(0,0)$

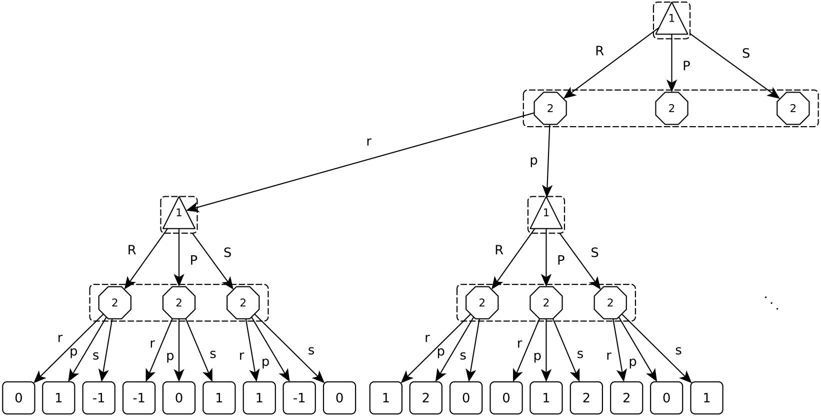
Question

How can we model such games?

We can model the game as an extensive-form game.

Pros: we already know how to solve such a game.
Cons: it is unnecessarily large.

RPS Played Twice as an Extensive-Form Games



We can use a model specific for repeated games.

Finitely Repeated Games

Definition

In repeated games we assume that a normal-form game, termed the *stage game*, is played repeatedly. If the number of repetitions (or rounds) is finite, we talk about *finitely repeated games*.

Question

How can we solve finitely repeated games?

We can use backward induction.

Why does this work if we have an extensive-form game with imperfect information?

Infinitely Repeated Games

Definition

Assume that a *stage game* is played repeatedly. If the number of repetitions (or rounds) is infinite, we talk about *infinitely repeated games*.

We cannot use extensive-form games as a underlying model. There are no leafs to assign utility values to. We need to define other utility measures:

Definition

Given an infinite sequence of payoffs $r_i^{(1)}, r_i^{(2)}, \ldots$ for player i, the *average reward* of i is

$$\lim_{k \to \infty} \frac{\sum_{j=1}^{k} r_i^{(j)}}{k}$$

Infinitely Repeated Games

Definition

Given an infinite sequence of payoffs $r_i^{(1)}, r_i^{(2)}, \ldots$ for player i, and a discount factor β with $0 \leq \beta \leq 1$, the *future discounted reward* is

$$\sum_{j=1}^{\infty} \beta^j r_i^{(j)}$$

Why do we use discount factor?

- a player cares more about immediate rewards
- a repeated game can terminate after each round with probability $1 - \beta$

Strategies in Repeated Games

How can we represent the strategies in infinitely repeated games? (the game tree is infinite)

- *a stationary strategy* – a randomized strategy that is played in each stage game

Is this enough? Consider a repeated prisoners dilemma – what is the most famous strategy in repeated prisoners dilemma?

Tit-for-tat: the player starts by cooperating and thereafter chooses in round $j+1$ the action chosen by the other player in round j.

We can have more complex strategies consisting of states/machines.

Strategies in Repeated Games

Definition

A payoff profile $r = (r_1, r_2, \ldots, r_n)$ is *enforceable* if $\forall i \in \mathcal{N}$, $r_i \geq v_i$.

where v_i is a minmax value for player i

$$v_i = \min_{s_{-i} \in \mathcal{S}_{-i}} \max_{s_i \in \mathcal{S}_i} u_i(s_{-i}, s_i)$$

Definition

A payoff profile $r = (r_1, r_2, \ldots, r_n)$ is *feasible* if there exist rational, nonnegative values α_a such that for all i, we can express r_i as $\sum_{a \in \mathcal{A}} \alpha_a u_i(a)$, with $\sum_{a \in \mathcal{A}} \alpha_a = 1$.

Nash Strategies in Repeated Games

Theorem (Folk Theorem)

Consider any n-player normal-form game G and any payoff profile $r = (r_1, r_2, \ldots, r_n)$.

1. *If r is the payoff profile for any Nash equilibrium s of the infinitely repeated G with average rewards, then for each player i, r_i is enforceable.*
2. *If r is both feasible and enforceable, then r is the payoff profile for some Nash equilibrium of the infinitely repeated G with average rewards.*

Stochastic Games

Let's generalize the repeated games. We do not have to play the same normal-form game repeatedly. We can play different normal-form games (possibly for infinitely long time).

Definition (Stochastic game)

A *stochastic game* is a tuple $(Q, \mathcal{N}, \mathcal{A}, \mathcal{P}, \mathcal{R})$, where:

- Q is a finite set of games
- $\mathcal{N}$ is a finite set of players
- $\mathcal{A}$ is a finite set of actions, $\mathcal{A}_i$ are actions available to player i
- $\mathcal{P}$ is a transition function $\mathcal{P} : Q \times \mathcal{A} \times Q :\to [0, 1]$, where $\mathcal{P}(q, a, q')$ is a probability of reaching game q' after a joint action a is played in game q
- $\mathcal{R}$ is a set of reward functions $r_i : Q \times \mathcal{A} \to \mathbb{R}$

Stochastic Games

Similarly to repeated games we can have several different rewards (or objectives):

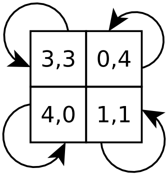
- discounted
- average
- reachability/safety

In reachability objectives a player wants to visit certain games infinitely often.

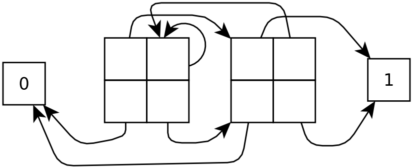
Related to reaching some target state (for example attacking a target) in a game without a pre-determined horizon.

Stochastic Games - Examples

Repeated prisoners dilemma:



Dante's purgatory:



Equilibria in Stochastic Games

Definition (History)

Let $h_t = (q_0, a_0, q_1, a_1, \ldots, a_{t1}, q_t)$ denote a history of t stages of a stochastic game, and let H_t be the set of all possible histories of this length.

Definition (Behavioral strategy)

A behavioral strategy $s_i(h_t, a_{i_j})$ returns the probability of playing action a_{i_j} for history h_t.

Definition (Markov strategy)

A Markov strategy s_i is a behavioral strategy in which $s_i(h_t, a_{i_j}) = s_i(h'_t, a_{i_j})$ if $q_t = q'_t$, where q_t and q'_t are the final games of h_t and h'_t, respectively.

Equilibria in Stochastic Games

Definition (Stationary strategy)

A stationary strategy s_i is a Markov strategy in which $s_i(h_{t_1}, a_{i_j}) = s_i(h_{t_2}, a_{i_j})$ if $q_t = q'_t$, where q_t and q'_t are the final games of h_{t_1} and h_{t_2}, respectively.

Definition

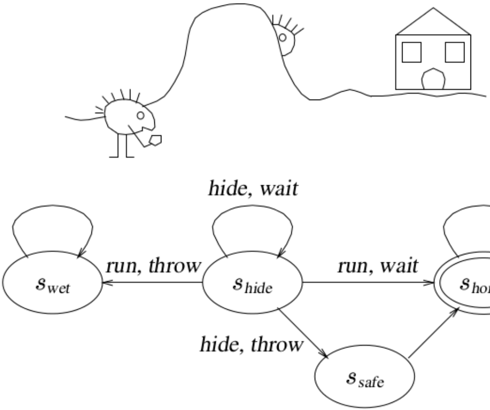
A strategy profile is called a *Markov perfect equilibrium* if it consists of only Markov strategies, and is a Nash equilibrium.

Theorem

Every n-player, general-sum, discounted-reward stochastic game has a Markov perfect equilibrium.

Equilibria in Stochastic Games

For other rewards, Markov perfect equilibrium does not have to exist.



Approximating Optimal Strategies in Stochastic Games

Standard algorithms from Markov Decision Processes, value and strategy iteration, translates to stochastic games.

Algorithm 1. Value Iteration

$t := 0$

$\tilde{v}^0 := (0, \ldots, 0, 1)$ // *the vector* $\tilde{v}^0$ *is indexed* $0, 1, \ldots, N, N+1$

while true do

  $t := t + 1$

  $\tilde{v}^t_0 := 0$

  $\tilde{v}^t_{N+1} := 1$

  for $i \in \{1, 2, \ldots, N\}$ **do**

    $\tilde{v}^t_i := \mathrm{val}(A_i(\tilde{v}^{t-1}))$

Approximating Optimal Strategies in Stochastic Games

Algorithm 2. Strategy Iteration

$t := 1$
$x^1 :=$ the strategy for Player I playing uniformly at each position
while true do
    $y^t :=$ an optimal *best reply* by Player II to x^t
    for $i \in \{0, 1, 2, \ldots, N, N+1\}$ **do**
        $v_i^t := \mu_i(x^t, y^t)$
    $t := t + 1$
    for $i \in \{1, 2, \ldots, N\}$ **do**
        if $\text{val}(A_i(v^{t-1})) > v_i^{t-1}$ **then**
            $x_i^t := \text{maximin}(A_i(v^{t-1}))$
        else
            $x_i^t := x_i^{t-1}$

Succinct Representations

compact representation of the game with $n = |\mathcal{N}|$ players

we want to reduce the input from $|\mathcal{S}|^{|\mathcal{N}|}$ to $|\mathcal{S}|^d$, where $d \ll |\mathcal{N}|$

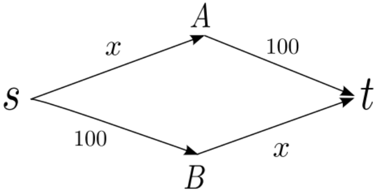
examples of succinct representations :

- congestion games (network congestion games, ...)
- polymatrix games (zero-sum polymatrix games)
- graphical games (action graph games)

Atomic Congestion Games

We have n players, set of edges E, strategies for each player are *paths* in the network ($\mathcal{S}$), and there is a congestion function $c_e : \{0, 1, \ldots, n\} \to \mathbb{Z}^+$. When all players choose their strategy path $s_i \in \mathcal{S}_i$ we have the load of edge e, $\ell_s(e) = |\{s_i : e \in s_i\}|$ and $u_i = -\sum_{e \in s_i} c_e(\ell_s(e))$.

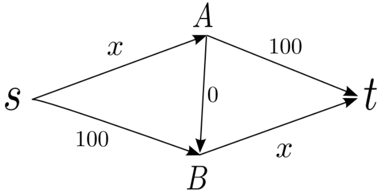
Braess' paradox



100 drivers that want to go from s to t.
What is Nash equilibrium?

Atomic Congestion Games

Now consider that we introduce a new edge between A and B, such that $c_{(A,B)}(x) = 0, \ \forall x \in \ell_{(A,B)}$.



What is Nash equilibrium?

Atomic Congestion Games

Theorem

Every atomic congestion game has a pure Nash equilibrium.

We can find it by an algorithm where players iteratively switch to their pure best response. This holds for generalizations:

- weighted congestion games
- all games known as *potential games*

For some subclasses, it is polynomial to find a pure NE (e.g., for symmetric network congestion games due to min-cost flow).