Repeated and Stochastic Games

Branislav Bošanský

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

branislav.bosansky@agents.fel.cvut.cz

April 1, 2019

Repeated Games

Repeated Games

Repeated Games are the simplest type of a dynamic game that evolves over time.

Repeated Games

Repeated Games are the simplest type of a dynamic game that evolves over time.

As such we can treat them as an extensive-form game (the finitely repeated case), or a stochastic game (the infinitely repeated case). However, such representations are very inefficient.

Repeated Games

Repeated Games are the simplest type of a dynamic game that evolves over time.

As such we can treat them as an extensive-form game (the finitely repeated case), or a stochastic game (the infinitely repeated case). However, such representations are very inefficient.

Repeated games can thus be seen as an example of a compact representation.

Repeated Games

Repeated Games are the simplest type of a dynamic game that evolves over time.

As such we can treat them as an extensive-form game (the finitely repeated case), or a stochastic game (the infinitely repeated case). However, such representations are very inefficient.

Repeated games can thus be seen as an example of a compact representation.

	C	D
C	$(1, 1)$	$(-1, 2)$
D	$(2, -1)$	$(0, 0)$

Repeated Games

Repeated Games are the simplest type of a dynamic game that evolves over time.

As such we can treat them as an extensive-form game (the finitely repeated case), or a stochastic game (the infinitely repeated case). However, such representations are very inefficient.

Repeated games can thus be seen as an example of a compact representation.

	C	D
C	$(1,1)$	$(-1,2)$
D	$(2,-1)$	$(0,0)$

Natural question: Is a NE of a single game the same as in the (in)finitely repeated game?

Repeated Games

Repeated Games

Definition

Let $G' = (\mathcal{N}, \mathcal{A}, u)$ be a normal-form game. An **infinitely repeated game** with discounted payoff is an extensive-form game with simultaneous moves $G^{\infty} = (\mathcal{N}, \mathcal{H}, \mathcal{A}, g, \delta)$, where

Repeated Games

Definition

Let $G' = (\mathcal{N}, \mathcal{A}, u)$ be a normal-form game. An **infinitely repeated game** with discounted payoff is an extensive-form game with simultaneous moves $G^\infty = (\mathcal{N}, \mathcal{H}, \mathcal{A}, g, \delta)$, where

- $\mathcal{H} = \{\emptyset\} \cup \bigcup_{t=1}^{\infty} A^t \cup A^\infty$

Repeated Games

Definition

Let $G' = (\mathcal{N}, \mathcal{A}, u)$ be a normal-form game. An **infinitely repeated game** with discounted payoff is an extensive-form game with simultaneous moves $G^\infty = (\mathcal{N}, \mathcal{H}, \mathcal{A}, g, \delta)$, where

- $\mathcal{H} = \{\emptyset\} \cup \bigcup_{t=1}^{\infty} A^t \cup A^\infty$
- $\mathcal{S}_i : \mathcal{H} \to \mathcal{A}_i$

Repeated Games

Definition

Let $G' = (\mathcal{N}, \mathcal{A}, u)$ be a normal-form game. An **infinitely repeated game** with discounted payoff is an extensive-form game with simultaneous moves $G^\infty = (\mathcal{N}, \mathcal{H}, \mathcal{A}, g, \delta)$, where

- $\mathcal{H} = \{\emptyset\} \cup \bigcup_{t=1}^{\infty} A^t \cup A^\infty$
- $\mathcal{S}_i : \mathcal{H} \to \mathcal{A}_i$
- $g_i(s_i, s_{-i}) = (1-\delta) \sum_{t=1}^{\infty} \delta^t \mathbb{E}_{a_i \sim s_i, a_{-i} \sim s_{-i}} (u_i(a_i, a_{-i}))$

Repeated Games

Definition

Let $G' = (\mathcal{N}, \mathcal{A}, u)$ be a normal-form game. An **infinitely repeated game** with discounted payoff is an extensive-form game with simultaneous moves $G^\infty = (\mathcal{N}, \mathcal{H}, \mathcal{A}, g, \delta)$, where

- $\mathcal{H} = \{\emptyset\} \cup \bigcup_{t=1}^{\infty} A^t \cup A^\infty$
- $\mathcal{S}_i : \mathcal{H} \to \mathcal{A}_i$
- $g_i(s_i, s_{-i}) = (1-\delta) \sum_{t=1}^{\infty} \delta^t \mathbb{E}_{a_i \sim s_i, a_{-i} \sim s_{-i}}(u_i(a_i, a_{-i}))$
- $\delta \in (0, 1)$ is the discount factor

Repeated Games

Repeated Games

We can define alternative utility functions in repeated games based on payoff vectors v_i^t for each:

Repeated Games

We can define alternative utility functions in repeated games based on payoff vectors v_i^t for each:

- overtaking payoff: $\lim_{T\to\infty}\sum_{t=1}^{T} v_i^t$

Repeated Games

We can define alternative utility functions in repeated games based on payoff vectors v_i^t for each:

- overtaking payoff: $\lim_{T\to\infty} \sum_{t=1}^{T} v_i^t$
- average payoff (or limit mean payoff): $\lim_{T\to\infty} \sum_{t=1}^{T} v_i^t / T$

Repeated Games

We can define alternative utility functions in repeated games based on payoff vectors v_i^t for each:

- overtaking payoff: $\lim_{T\to\infty} \sum_{t=1}^{T} v_i^t$
- average payoff (or limit mean payoff): $\lim_{T\to\infty} \sum_{t=1}^{T} v_i^t / T$

Definition

Player i's min-max payoff is

$$\underline{v_i} = \min_{s_{-i}} \max_{s_i} g_i(s_i, s_{-i})$$

Repeated Games

We can define alternative utility functions in repeated games based on payoff vectors v_i^t for each:

- overtaking payoff: $\lim_{T\to\infty} \sum_{t=1}^{T} v_i^t$
- average payoff (or limit mean payoff): $\lim_{T\to\infty} \sum_{t=1}^{T} v_i^t / T$

Definition

Player i's min-max payoff is

$$\underline{v_i} = \min_{s_{-i}} \max_{s_i} g_i(s_i, s_{-i})$$

A strategy s is *individually rational* if $g_i(s) \geq \underline{v_i}$

Repeated Games

Repeated Games

Theorem (Nash Folk Theorem)

If v_i is a feasible and an individually rational payoff, then there exists a discount factor $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$, there is a Nash equilibrium of G with payoff v_i.

Theorem (Nash Folk Theorem)

If v_i is a feasible and an individually rational payoff, then there exists a discount factor $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$, there is a Nash equilibrium of G with payoff v_i.

Proof.

If v_i is feasible then there exist a strategy s such that $g_i(s) = v_i$ and let m_{-i} be the minmax strategy of other players to reach value $\underline{v_i}$ for player i. Let consider the following strategy:

Repeated Games

Theorem (Nash Folk Theorem)

If v_i is a feasible and an individually rational payoff, then there exists a discount factor $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$, there is a Nash equilibrium of G with payoff v_i.

Proof.

If v_i is feasible then there exist a strategy s such that $g_i(s) = v_i$ and let m_{-i} be the minmax strategy of other players to reach value $\underline{v_i}$ for player i. Let consider the following strategy:

1. play according to s_i as long as no one deviates

Repeated Games

Theorem (Nash Folk Theorem)

If v_i is a feasible and an individually rational payoff, then there exists a discount factor $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$, there is a Nash equilibrium of G with payoff v_i.

Proof.

If v_i is feasible then there exist a strategy s such that $g_i(s) = v_i$ and let m_{-i} be the minmax strategy of other players to reach value $\underline{v_i}$ for player i. Let consider the following strategy:

1. play according to s_i as long as no one deviates
2. let $\overline{v_i}$ be the maximum value player i can get by a deviation in step t

Repeated Games

Theorem (Nash Folk Theorem)

If v_i is a feasible and an individually rational payoff, then there exists a discount factor $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$, there is a Nash equilibrium of G with payoff v_i.

Proof.

If v_i is feasible then there exist a strategy s such that $g_i(s) = v_i$ and let m_{-i} be the minmax strategy of other players to reach value $\underline{v_i}$ for player i. Let consider the following strategy:

1. play according to s_i as long as no one deviates
2. let $\overline{v_i}$ be the maximum value player i can get by a deviation in step t

$$(1-\delta)[v_i + \delta v_i + \ldots + \delta^t \overline{v_i} + \delta^{t+1} \underline{v_i} + \ldots] \leq$$
$$\leq (1-\delta)[v_i + \delta v_i + \ldots + \delta^t v_i + \quad \delta^{t+1} v_i + \ldots]$$

Repeated Games

(Proof cont.)

By setting $\underline{\delta}$ sufficiently large approaching 1 the above inequality holds. $\square$

Repeated Games

(Proof cont.)

By setting $\underline{\delta}$ sufficiently large approaching 1 the above inequality holds. $\square$

The Nash folk theorem says that essentially anything goes as a Nash equilibrium payoff in a discounted repeated game.

Repeated Games

(Proof cont.)

By setting $\underline{\delta}$ sufficiently large approaching 1 the above inequality holds. $\square$

The Nash folk theorem says that essentially anything goes as a Nash equilibrium payoff in a discounted repeated game.

The players threat by playing *grim trigger* strategies.

Stochastic Games

Stochastic Games

Let's generalize the repeated games.

Stochastic Games

Let's generalize the repeated games. We do not have to play the same normal-form game repeatedly.

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).

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:

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

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

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

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

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 $\mathcal{P} : Q \times \mathcal{A} \times Q :\rightarrow [0, 1]$, 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 $r_i : Q \times \mathcal{A} \rightarrow \mathbb{R}$

Stochastic Games

Stochastic Games

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

Stochastic Games

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

- discounted

Stochastic Games

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

- discounted
- average

Stochastic Games

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

- discounted
- average
- reachability/safety

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.

Stochastic Games

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

- discounted
- average
- reachability/safety

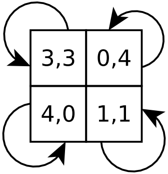
In reachability objectives a player wants to visit certain games infinitely often.

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

Stochastic Games - Examples

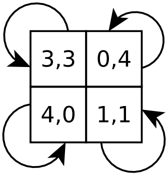
Stochastic Games - Examples

Repeated prisoners dilemma:

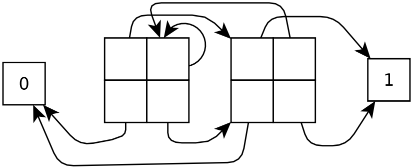


Stochastic Games - Examples

Repeated prisoners dilemma:



Dante's purgatory:



Equilibria in Stochastic Games

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.

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.

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

Equilibria in Stochastic Games

Definition

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

Equilibria in Stochastic Games

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

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.

Theorem

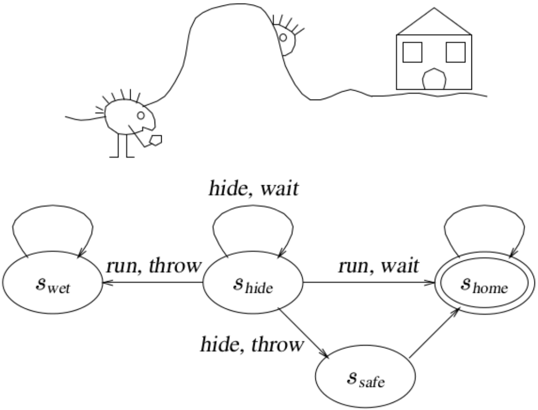
Problem of computing an optimal strategy in simple (turn-taking) stochastic games, where pure stationary strategies are known to be optimal, is in PLS.

Equilibria in Stochastic Games

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

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, translate to stochastic games.

[1]Pseudocode from [3].

Approximating Optimal Strategies in Stochastic Games

Standard algorithms from Markov Decision Processes, value and strategy iteration, translate 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
    $t := t + 1$
    $\tilde{v}_0^t := 0$
    $\tilde{v}_{N+1}^t := 1$
    for $i \in \{1, 2, \ldots, N\}$ **do**
        $\tilde{v}_i^t := \text{val}(A_i(\tilde{v}^{t-1}))$

[1]

[1] Pseudocode from [3].

Approximating Optimal Strategies in Stochastic Games

Algorithm 2. Strategy Iteration [2]

$t := 1$
$x^1 :=$ 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 $\mathrm{val}(A_i(v^{t-1})) > v_i^{t-1}$ **then**
            $x_i^t := \mathrm{maximin}(A_i(v^{t-1}))$
        else
            $x_i^t := x_i^{t-1}$

[2] Pseudocode from [3].

Stochastic Games with Imperfect Information

Extending stochastic games to imperfect information (known as *partial observability*, hence termed Partially Observable Stochastic Games (POSGs)) is lot more complicated compared to finite EFGs.

Stochastic Games with Imperfect Information

Extending stochastic games to imperfect information (known as *partial observability*, hence termed Partially Observable Stochastic Games (POSGs)) is lot more complicated compared to finite EFGs.

The problem lies with *Nested beliefs*. Consider a two-player game where each player has some private state unobserved by the opponent:

Stochastic Games with Imperfect Information

Extending stochastic games to imperfect information (known as *partial observability*, hence termed Partially Observable Stochastic Games (POSGs)) is lot more complicated compared to finite EFGs.

The problem lies with *Nested beliefs*. Consider a two-player game where each player has some private state unobserved by the opponent:

- A player i has uncertainty about the exact state of the opponent $-i$ – there is a belief (a probability distribution) over possible states.

Stochastic Games with Imperfect Information

Extending stochastic games to imperfect information (known as *partial observability*, hence termed Partially Observable Stochastic Games (POSGs)) is lot more complicated compared to finite EFGs.

The problem lies with *Nested beliefs*. Consider a two-player game where each player has some private state unobserved by the opponent:

- A player i has uncertainty about the exact state of the opponent $-i$ – there is a belief (a probability distribution) over possible states.
- The optimal strategy of player i depends on the strategy of the opponent $-i$ that depends on the belief over possible private states of player i.

Stochastic Games with Imperfect Information

Extending stochastic games to imperfect information (known as *partial observability*, hence termed Partially Observable Stochastic Games (POSGs)) is lot more complicated compared to finite EFGs.

The problem lies with *Nested beliefs*. Consider a two-player game where each player has some private state unobserved by the opponent:

- A player i has uncertainty about the exact state of the opponent $-i$ – there is a belief (a probability distribution) over possible states.
- The optimal strategy of player i depends on the strategy of the opponent $-i$ that depends on the belief over possible private states of player i.
- Each player needs to consider beliefs, belief of beliefs, ... etc.

Partially Observable Stochastic Games (POSGs)

Solving general POSGs is not tractable (even solving related single-player decision problems is often undecidable [2]).

Partially Observable Stochastic Games (POSGs)

Solving general POSGs is not tractable (even solving related single-player decision problems is often undecidable [2]).

We can restrict to subclasses of games with limited partial observability:

Partially Observable Stochastic Games (POSGs)

Solving general POSGs is not tractable (even solving related single-player decision problems is often undecidable [2]).

We can restrict to subclasses of games with limited partial observability:

- One-Sided Partially Observable Stochastic Games [4]

Partially Observable Stochastic Games (POSGs)

Solving general POSGs is not tractable (even solving related single-player decision problems is often undecidable [2]).

We can restrict to subclasses of games with limited partial observability:

- One-Sided Partially Observable Stochastic Games [4]
- Partially Observable Stochastic Games with Public Observations [5]

Partially Observable Stochastic Games (POSGs)

Theoretical results:

Partially Observable Stochastic Games (POSGs)

Theoretical results:

- Value function (function that assigns a belief point value of a (sub)-game) is convex (or convex-concave).

Partially Observable Stochastic Games (POSGs)

Theoretical results:

- Value function (function that assigns a belief point value of a (sub)-game) is convex (or convex-concave).
- We can define dynamic-programming operator – a generalization of Bellman update.

$$[Hv](b) = \min_{\pi_2} \max_{\pi_1} \left(R^{\mathrm{imm}}_{\pi_1,\pi_2} + \gamma \cdot R^{\mathrm{subs}}_{\pi_1,\pi_2}(v) \right)$$

Partially Observable Stochastic Games (POSGs)

Theoretical results:

- Value function (function that assigns a belief point value of a (sub)-game) is convex (or convex-concave).
- We can define dynamic-programming operator – a generalization of Bellman update.

$$[Hv](b) = \min_{\pi_2} \max_{\pi_1} \left(R^{\mathrm{imm}}_{\pi_1,\pi_2} + \gamma \cdot R^{\mathrm{subs}}_{\pi_1,\pi_2}(v) \right)$$

$$R^{\mathrm{imm}}_{\pi_1,\pi_2} = \sum_{s \in \mathcal{S}} \sum_{a \in \mathcal{A}_1} \sum_{a' \in \mathcal{A}_2} b(s) \cdot \pi_1(a) \cdot \pi_2(s, a') \cdot \mathcal{R}(s, a, a')$$

$$R^{\mathrm{subs}}_{\pi_1,\pi_2}(v) = \sum_{a \in \mathcal{A}_1} \sum_{o \in \mathcal{O}} \pi_1(a) \cdot \Pr[o|a, \pi_2] \cdot v(b^{a,o}_{\pi_2})$$

Partially Observable Stochastic Games (POSGs)

We can generalize value-iteration algorithms for POMDPs to POSGs.

Heuristic Search Value Iteration (HSVI):

Data: Game $\langle \mathcal{S}, \mathcal{A}_1, \mathcal{A}_2, \mathcal{O}, \mathcal{T}, \mathcal{R} \rangle$, initial belief b^0, discount factor γ, desired precision $\epsilon > 0$, neighborhood parameter R

Result: Approximate value function $\hat{v}$

Initialize $\hat{v}$
while $\text{gap}(\hat{v}(b^0)) > \epsilon$ **do**
    `Explore` $(b^0, \epsilon, R, 0)$
return $\hat{v}$

procedure `Explore` (b, ϵ, R, t)
    $\pi_2 \leftarrow$ optimal strategy of player 2 in $[H\underline{v}](b)$
    $(a, o) \leftarrow$ select according to forward exploration heuristic
    if $\text{excess}(\hat{v}(b^{a,o}_{\pi_2}), t+1) > 0$ **then**
        `Explore` $(b^{a,o}_{\pi_2}, \epsilon, R, t+1)$
    $\Gamma \leftarrow \Gamma \cup \{L\Gamma(b)\}$
    $\Upsilon \leftarrow \Upsilon \cup \{U\Upsilon(b)\}$ and make $\overline{v}$ $(U-L)$-Lipschitz

Algorithm 1: HSVI algorithm for one-sided POSGs

References I

(besides the books)

[1] M. Osborne and A. Rubinstein, *A course in game theory*.
MIT press, 1994.

[2] O. Madani, S. Hanks, and A. Condon, "On the undecidability of probabilistic planning and infinite-horizon partially observable Markov decision problems," in *AAAI/IAAI*, pp. 541–548, 1999.

[3] K. A. Hansen, R. Ibsen-Jensen, and P. B. Miltersen, "The Complexity of Solving Reachability Games Using Value and Strategy Iteration," in *International Computer Science Symposium in Russia*, pp. 77–90, 2011.

[4] Horák, K., Bošanský, B., and Pěchouček, M. (2017).
Heuristic Search Value Iteration for One-Sided Partially Observable Stochastic Games. In *In Proceedings of AAAI Conference on Artificial Intelligence*, pages 558–564.

References II

[5] Horák, K., Bošanský, B. (2019).
Solving Partially Observable Stochastic Games with Public Observations.
In *In Proceedings of AAAI Conference on Artificial Intelligence*.