Richer Representations: Stochastic Games

Stochastic Games

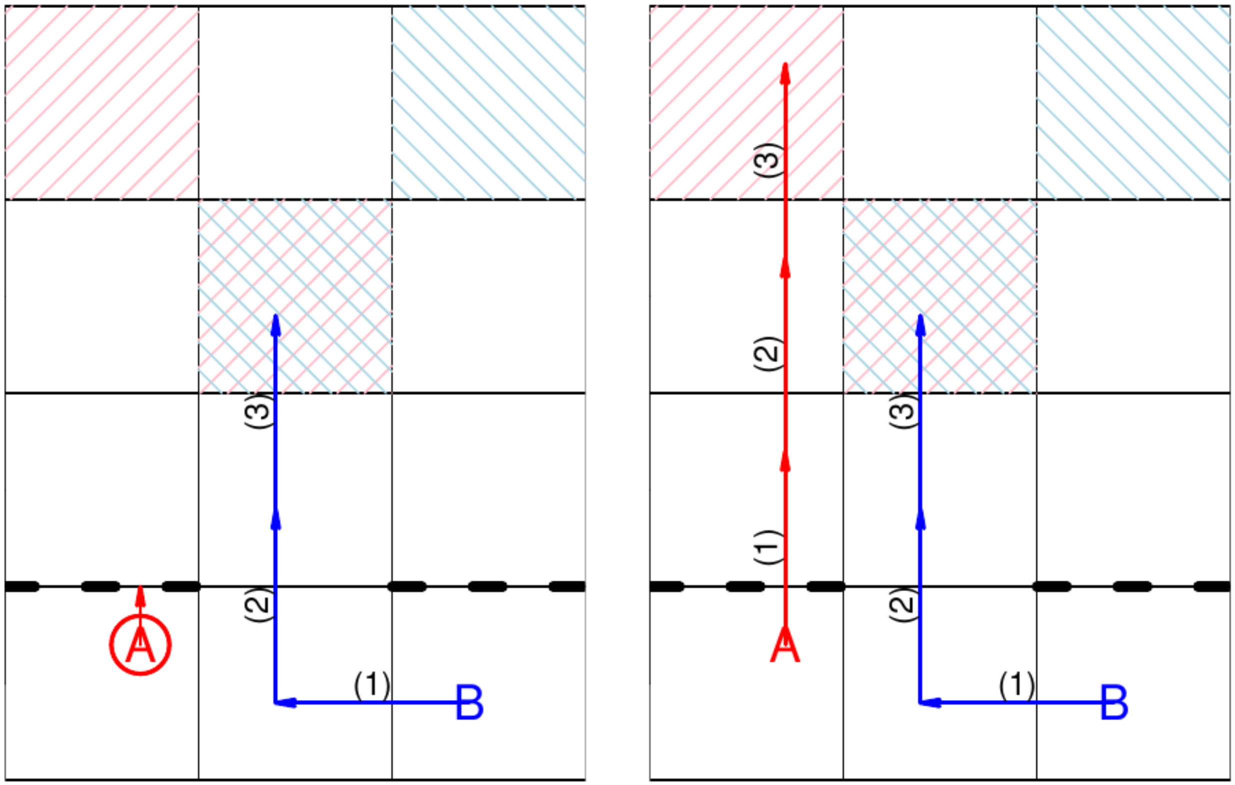
- Repeated game, where players play potentially different games at each iteration
- Transition probability to go to new game based on the chosen action profile
- Rewards can be defined based on the single-stage rewards similar to repeated games

Stochastic Games

Definition 6.2.1 (Stochastic game) *A* stochastic game *(also known as a* Markov game*) is a tuple* (Q, N, A, P, r)*, where:*

- Q *is a finite set of games;*
- N *is a finite set of* n *players;*
- $A = A_1 \times \cdots \times A_n$*, where* A_i *is a finite set of actions available to player* i*;*
- $P : Q \times A \times Q \mapsto [0, 1]$ *is the transition probability function;* $P(q, a, \hat{q})$ *is the probability of transitioning from state* q *to state* $\hat{q}$ *after action profile* a*; and*
- $R = r_1, \ldots, r_n$*, where* $r_i : Q \times A \mapsto \mathbb{R}$ *is a real-valued payoff function for player* i*.*

Example – “Turkey”



Stochastic Games – Strategies and Equilibria

Stochastic Games

Definition 6.2.1 (Stochastic game) *A* stochastic game *(also known as a* Markov game*) is a tuple* (Q, N, A, P, r)*, where:*

- Q *is a finite set of games;*
- N *is a finite set of* n *players;*
- $A = A_1 \times \cdots \times A_n$*, where* A_i *is a finite set of actions available to player* i*;*
- $P : Q \times A \times Q \mapsto [0, 1]$ *is the transition probability function;* $P(q, a, \hat{q})$ *is the probability of transitioning from state* q *to state* $\hat{q}$ *after action profile* a*; and*
- $R = r_1, \ldots, r_n$*, where* $r_i : Q \times A \mapsto \mathbb{R}$ *is a real-valued payoff function for player* i*.*

Strategies

- Let $h_t = (q^0, a^0, q^1, a^1, \ldots, a^{t-1}, q^t)$ denote a history of *t* stages of a stochastic game. And let H_t be set of all histories of length *t*
- The set of deterministic strategies: $\prod_{t, H_t} A_i$
- Mixed strategy: Any mixture over deterministic strategies

Restricted Classes of Strategies

Definition 6.2.2 (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 6.2.3 (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 states of h_t and h'_t, respectively.*

Definition 6.2.4 (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 $q_{t_1} = q'_{t_2}$, where q_{t_1} and q'_{t_2} are the final states of h_{t_1} and h'_{t_2}, respectively.*

Equilibria

- With discounted reward, every stochastic game has a Nash equilibrium
- *Markov Perfect Equilibrium* – consists only of Markov strategies, and is a Nash equilibrium regardless of the start state (analogue to subgame-perfect equilibrium)

Equilibria

- With discounted reward, every stochastic game has a Nash equilibrium
- *Markov Perfect Equilibrium* – consists only of Markov strategies, and is a Nash equilibrium regardless of the start state (analogue to subgame-perfect equilibrium)

Theorem 6.2.5 *Every n-player, general-sum, discounted-reward stochastic game has a Markov perfect equilibrium.*

Stochastic Games – Computing Equilibria

Stochastic Games

Definition 6.2.1 (Stochastic game) *A* stochastic game *(also known as a* Markov game*) is a tuple* (Q, N, A, P, r)*, where:*

- Q *is a finite set of games;*
- N *is a finite set of* n *players;*
- $A = A_1 \times \cdots \times A_n$*, where* A_i *is a finite set of actions available to player* i*;*
- $P : Q \times A \times Q \mapsto [0, 1]$ *is the transition probability function;* $P(q, a, \hat{q})$ *is the probability of transitioning from state* q *to state* $\hat{q}$ *after action profile* a*; and*
- $R = r_1, \ldots, r_n$*, where* $r_i : Q \times A \mapsto \mathbb{R}$ *is a real-valued payoff function for player* i*.*

Computing Nash Equilibria

- Q. Is there a polynomial-time procedure to compute a Nash equilibrium of a stochastic game?
  - There is an LP formulation for solving (single-player) MDPs
- There is an LP formulation of a single player controls the transitions:

Definition 6.2.8 (Single-controller stochastic game) *A stochastic game is* single-controller *if there exists a player i such that $\forall q, q' \in Q, \forall a \in A,\ P(q, a, q') = P(q, a', q')$ if $a_i = a'_i$.*

Computing Nash Equilibria

- There is also LP formulation to find equilibrium if:
  - state and action profile have independent effects on the reward achieved by each agent
  - transition function only depends on action profile

Definition 6.2.9 (SR-SIT stochastic game) *A stochastic game is* separable reward state independent transition *(SR-SIT) if the following two conditions hold:*

- *there exist functions* α, γ *such that* $\forall i, q \in Q, \forall a \in A$ *it is the case that* $r_i(q, a) = \alpha(q) + \gamma(a)$*; and*
- $\forall q, q', q'' \in Q, \forall a \in A$ *it is the case that* $P(q, a, q'') = P(q', a, q'')$.