T-61.6020 Reinforcement Learning – Theory and Applications

Könönen, V. (2004). Multiagent Reinforcement Learning in Markov Games: Asymmetric and Symmetric Approaches

Chapter 3: Game Theory

Joni Pajarinen, Joni.Pajarinen@tkk.fi

28.2.2006

Game theory

- Basic game theoretical models
- Solution concepts for games

Basic game theoretical models

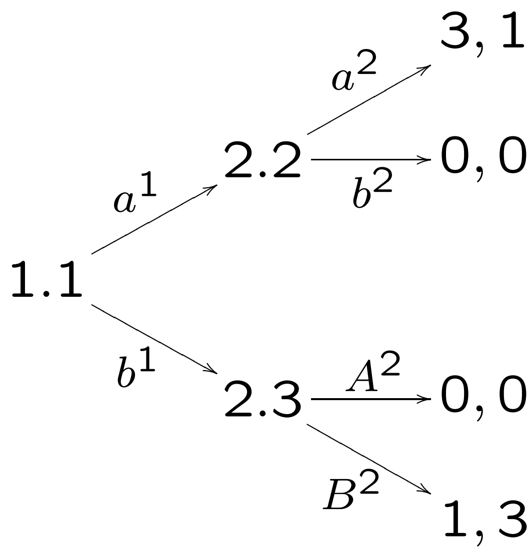
- Categorization of games
- Extensive form
- Strategic form (matrix games)
- Correspondence of extensive form and the strategic form

Categorization of games

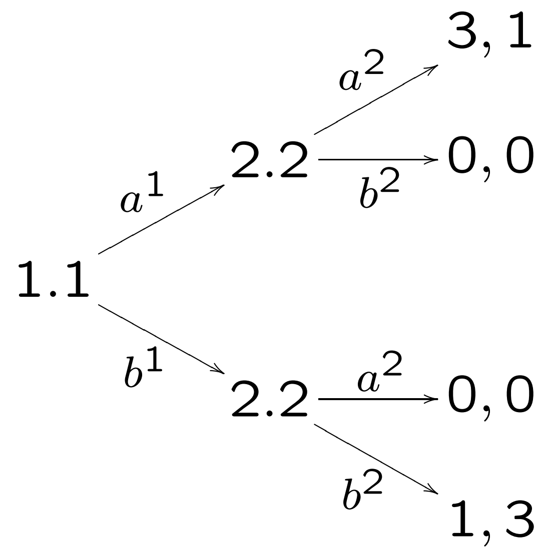
- Game theory extension of decision theory?
- Game payoff structures
  - Zero-sum
  - Team game

Extensive form

- Presented as a game tree
- Includes temporal relationships


3, 1
a2
2.2
0, 0
b2
a1
1.1
b1
2.3
A2
0, 0
B2
1, 3

$Game\,1$


3, 1
a2
2.2
0, 0
b2
a1
1.1
b1
2.2
a2
0, 0
b2
1, 3

$Game\,2$

Strategic form

- Presented as a matrix

- Example:

	a^2	b^2	c^2
a^1	0, 1	-2, -1	-1.5, 1.5
b^1	-1, -2	-1, 0	-3, -1
c^1	1, 0	-2, -1	-2, 0.5

- A pure strategy is one strategy of player i's all possible strategies: $a^i \epsilon A^i$

- A mixed strategy is a probability distribution over all possible strategies: $\sigma^i \ \epsilon \ \Delta(A^i)$

Correspondence of the extensive form and the strategic form

- Extensive form can be converted to strategic form: loss of temporal information

- Payoff equivalence leads to the purely reduced normal form

- Removal of randomly redundant strategies leads to the fully reduced normal form

- For multiagent games the player is added to the information state: Behavior strategy profile

Solution concepts for games

- Solution by elimination of dominated strategies
- Stackelberg equilibrium
- Correlated equilibrium
- Nash equilibrium
- MaxMin and MaxMax solutions
- Stackelberg and SubGame Perfectness

Solution by elimination of dominated strategies

- Strongly dominated strategy:
$r^i(a^1,...,a^{i-1},a^i,a^{i+1},...,a^N) < r^i(a^1,...,a^{i-1},\sigma^i,a^{i+1},...,a^N)$

- Weakly dominated strategy:
$r^i(a^1,...,a^{i-1},a^i,a^{i+1},...,a^N) \leq r^i(a^1,...,a^{i-1},\sigma^i,a^{i+1},...,a^N)$

- Outcome of elimination of weakly dominated strategies order dependant

- Solution:

	a^2	b^2	c^2
a^1	2,3	3,0	0,1
b^1	0,0	1,6	4,2

No Solution:

	a^2	b^2
a^1	3, 1	0, 0
b^1	0, 0	1, 3

Stackelberg equilibrium

- Strict ordering of players decisions

- Solution (b^1, b^2) for:

	a^2	b^2	c^2
a^1	1,1	0,0	0,0
b^1	0,0	3,1	0,0
c^1	0,0	0,0	2,1

- Requires unique strategy responses. Example, where strategy selection among equal payoffs must be explicitly defined:

	a^2	b^2
a^1	3, 1	0, 1
b^1	0, 0	1, 2

- Exists always in pure strategies and therefore fast to compute

- Only leader needs model of follower

Correlated equilibrium

- Simultaneous strategy selection

- Mediator that recommends pure strategies from the distribution $\delta \epsilon \Delta(A^1 \times ... \times A^N)$

- δ is a correlated equilibrium if $\sum_{a^{-i} \epsilon A^{-i}} \delta(a^i, a^{-i})(r^i(a^i, a^{-i}) - r^i(b^i, a^{-i})) \geq 0, \forall i \epsilon N, \forall a^i \epsilon A^i, \forall b^i \epsilon A^i$, where N is the number of persons and $A^{-i} = A^1 \times ... \times A^{i-1} \times A^{i+1} \times ... \times A^N$

- The sums consist of equations, which punish if the opponents strategies form a payoff, which is not the maximum that could be achieved with the players proposed strategy and reward if they do -> The equilibrium is an acceptable solution to all players

Nash equilibrium

- Independent mixed strategies $(\sigma^1, ..., \sigma^N)$ form Nash equilibrium if no player wants to change its strategy while knowing other players are obeying the nash equilibrium

- Definition: $\sigma^{\mathsf{i}} \epsilon \Delta(\mathsf{A}^{\mathsf{i}})$
  $r^i(\sigma_*^1, ..., \sigma_*^{i-1}, \sigma^i, \sigma_*^{i+1}, ..., \sigma_*^N) \leq r^i(\sigma_*^1, ..., \sigma_*^N)$

- Every finite game in strategic form has at least one mixed strategy nash equilibrium

- For pure strategies one might not exist

- Nash equilibrium example: pure strategies $([a^1],[a^2])$, $([b^1],[b^2])$ and mixed strategy $(0.75*[a^1]+0.25*[b^1], 0.25*[a^2]+0.75*[b^2])$ for game:

	a^2	b^2
a^1	3, 1	0, 0
b^1	0, 0	1, 3

MaxMin and MaxMax solutions

- Two-person zero-sum games: $r^1(a^1, a^2) = -r^2(a^1, a^2)$

- $\sigma^1 \;\epsilon\; \Delta(A^1)$ is a MaxMin strategy if $\sigma^1 \epsilon \arg\max_{\tau \epsilon \Delta(A^1)} \min_{a^2 \epsilon A^2} r^1(\tau, a^2)$

- MaxMin assumes that opponent tries to minimize player's reward and player tries to maximize it

- Simultaneous decisions -> only mixed strategy MaxMin solution

- Sequential decisions -> pure strategy solution exists (Stackelberg solution). This applies also to MaxMax solutions.

Stackelberg equilibrium and SubGame Perfectness

- A Nash equilibrium has the SGP property if it is a rational solution for all the proper subgames in the extensive form representation

- Achieved by backward induction starting from the leaf nodes and evaluating each proper subgame

- Constructing player hierarchies with SGP and Stackelberg: An example 3 player game, where leader chooses either a^1 or b^1 , which results in subgames 1 or 2 for the followers. If the followers choose the pure Nash equilibrium solution for the subgame, then the leader can easily choose its strategy. If for example the followers choose (a^2, a^3) for subgame 1 and (a^2, b^3) for subgame 2 and those result in payoffs 2 and 1 for the leader, then the leader can choose a^1,which results in (a^1, a^2, a^3) and payoff 2.

Subgame 1

	a^3	b^3
a^2	1, 1	0, 0
b^2	0, 0	1, 1

Subgame 2

	a^3	b^3
a^2	0, 0	1, 1
b^2	1, 1	0, 0