COMP3702/COMP7702
Artificial Intelligence

Module 5: Reasoning about other agents

Dr Archie Chapman

Semester 2, 2020

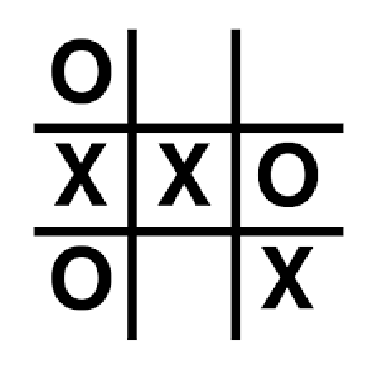
The University of Queensland
School of Information Technology and Electrical Engineering

Week 12: Logistics

- **RiPPLE round 4 closes today!**
- Assignment 4 is open. Due date **extended** to Monday Oct 16
- Marking Assignment 2 nearly complete.
- An Assignment 2 solution walk-through video is now ready for you to view, under Assessment ¿ Assignment 2 (with thanks to Nick)
- Exam Preparation materials can be found on Blackboard, under *Learning Resources*

Module 5: Reasoning about other agents

Two examples of strategic uncertainty



Single agent or multiple agents

Many domains are characterised by multiple agents rather than a single agent
$\Rightarrow$ Multiagent systems

What is a multi-agent system?

- Multi-agent systems is a computational system consisting of two or more interacting intelligent agents.
- Multi-agent systems are used to model and solve complex problems that are difficult or impossible for an individual agent or a monolithic system to solve.
- Agents can be cooperative, competitive or somewhere in between.
- Agents that are strategic can't be modelled as nature.

Multi-agent system and game theory

What is game theory?

- Game theory is an umbrella term for the study of mathematical models of strategic interaction among rational agents.
- The name "game" theory is traced to its origins analysing zero-sum games such as chess and card games, in which each participant's gains or losses are exactly balanced by those of the other participants.
- It is now used to model all manner of scenarios, from cooperation and coordination, to conflict and deceit.

Key idea: Game theory provides a model for reasoning over other agent's action and behaviours in AI, and is considered a foundation of multiagent system design and analysis.

Multi-agent framework

- Each agent can have its own utility.
- Agents select actions autonomously.
- Agents can have different information.
- The outcome can depend on the actions of all of the agents.
- Each agent's value depends on the outcome.

Classical game theory: Normal form vs extensive form

A key distinction in games regards the timing of when information about one agent becomes available to other agents in the environment.

Extensive form: Agents/players take turns to act.

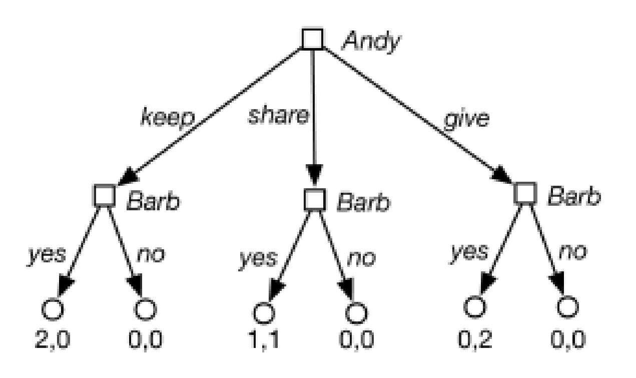
- Game is modelled by a game graph:
  - Game states are nodes
  - Each action is represented a edge
- Two cases:
  - **Perfect information**, can see other moves, akin to fully-observable systems
    E.g. Noughts and crosses, chess, go.
  - **Imperfect information**, like partial observability, cannot fully see other agents' moves
    E.g. Battleships, many card games

Normal form: Simultaneous moves, players cannot observe other's action choice:

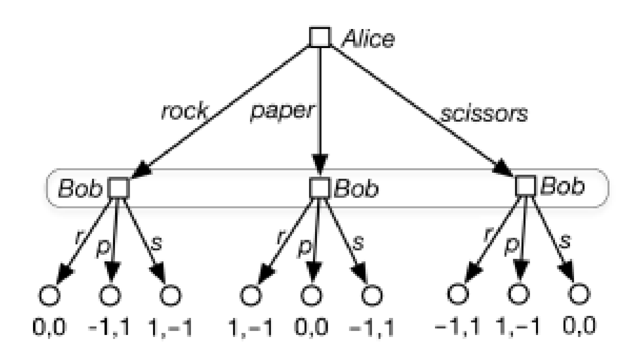
- Typically represented by a payoff matrix
- E.g. Rock-paper-scissors

Extensive form of a sharing game with perfect information

Players alternately take moves at each layer; actions are represented by edges; payoffs indicted by tuples at leaf nodes.



Extensive Form of an imperfect-information Game



Bob cannot distinguish the nodes in an information set.

Rock-Paper-Scissors in normal form

Players choose their action simultaneously

		Bob		
		rock	*paper*	*scissors*
Alice	*rock*	$0, 0$	$-1, 1$	$1, -1$
	paper	$1, -1$	$0, 0$	$-1, 1$
	scissors	$-1, 1$	$1, -1$	0,0

Fully Observable + Multiple Agents = Perfect Information Games

If agents act sequentially and can observe the state before acting:
$\Longrightarrow$ Perfect Information Games.

Can do dynamic programming or search: Each agent maximises for itself.

- When we apply this of these to two-person zero-sum games (pure conflict)
  $\Longrightarrow$ minimax search (very similar to AND-OR trees).
  We will think about this, but most real games are too big to carry out minimax search.
- When we apply this to general-sum games
  $\Longrightarrow$ Subgame perfect equilibrium solution.This can get very complicated to solve.

Topics beyond this course:

Multi-agent MDPs: A value function for each agent: Each agent maximises its own value function.
Multi-agent reinforcement learning: Each agent has its own Q function, rewards coupled through the environment.

We will focus on normal form games

The strategic form of a game or normal form game $\langle I, \{A_i, u_i\}_{i \in I} \rangle$:

- a finite set I of agents, $\{1, \ldots, n\}$

and for each agent $i \in I$:

- a set of actions A_i
    - A **pure strategy profile** a is a tuple $(a_1, \ldots, a_n)$, in which agent i carries out a_i.
    - A **mixed strategy profile** σ is a tuple of distributions over joint actions $(\sigma_1, \ldots, \sigma_n)$, in which agent i draws an action from a lottery $\sigma_i \in \Delta(a_i)$ ($\Delta(a_i)$ is the simplex over A_i).
- a utility function $u_i(\sigma)$ for strategy profile σ, which gives the expected utility for agent i when all agents follow their component of σ.

Under mild conditions on A_i and u_i, there always exists a mixed strategy profile in which no agent benefits from adjusting their strategy $\implies$ Nash equilibrium solution

Simultaneous moves and normal form games

Soccer penalty kick. The kicker can kick to his left or right. The goalkeeper can jump to her left or right.

		goalie	
		left	right
kicker	left	0.6	0.2
	right	0.3	0.9

Matrix shows probability of a goal

Expected utility of lotteries over actions

		goalie	
		left	right
kicker	left	0.6	0.2
	right	0.3	0.9

Matrix shows probability of a goal

Suppose that the kicker decides to kick to his right with probability p_k, and that the goalkeeper decides to jump to her right with probability p_g.

What is the probability of a goal?

Break: Reminder — SECaT surveys are available at:

`https://eval.uq.edu.au/`

Expected utility of lotteries over actions

		goalie	
		left	right
kicker	left	0.6	0.2
	right	0.3	0.9

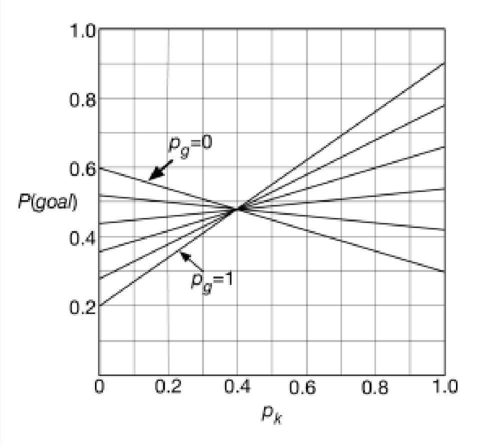
Matrix shows probability of a goal

Suppose that the kicker decides to kick to his right with probability p_k, and that the goalkeeper decides to dive to her right with probability p_g

What is the probability of a goal?

$$P(goal) = 0.9p_k p_g + 0.3p_k(1 - p_g) + 0.2(1 - p_k)p_g + 0.6(1 - p_k)(1 - p_g)$$

Expected utility of lotteries over actions



This figure shows the probability of a goal as a function of p_k The different lines correspond to different values of $p_g = \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$

There is something special about $p_k = 0.4$. At this value, the probability of a goal is 0.48, independent of the value of p_g.

The equivalent plot for p_g is similar, with all of the lines meeting at $p_g = 0.3$. Again, when $p_g = 0.3$ the probability of a goal is 0.48.

Strategic Dominance

Strategic dominance

A simple way to reason over which strategies to use in a game is by *strategic dominance* analysis.

An agent's strategy **dominates** another strategy if it is always a better choice, irrespective of how the other agents in the game play; that is, if the following holds $\forall \sigma_{-i} \in \Delta_{-i}$, $\bar{\sigma}_i \neq \underline{\sigma}_i$:

$$u_i(\bar{\sigma}_i, \sigma_{-i}) > u_i(\underline{\sigma}_i, \sigma_{-i}) \tag{1}$$

then $\bar{\sigma}_i$ dominates $\underline{\sigma}_i$.

Similarly, a strategy, $\underline{\sigma}_i$, is **dominated** if another strategy is always a better choice, independent of the other agents' actions:

$$u_i(\underline{\sigma}_i, \sigma_{-i}) < u_i(\sigma_i, \sigma_{-i}) \quad \forall \sigma_{-i} \in \Delta_{-i}, \; \underline{\sigma}_i \neq \sigma_i \tag{2}$$

Both concepts above rely on strict inequalities, but they can be weakened to consider weakly-dominating or weakly-dominated strategies, respectively, using weak inequalities.

Iterative elimination of weakly-dominated strategies

Iterative elimination of weakly-dominated strategies can be applied to any game.

This algorithmic procedure starts by choosing an agent removing all weakly-dominated strategies from its action space.

The idea is that a weakly dominated strategy will never be played, so it should not feature in any of the agents' strategic reasoning.

The elimination step is repeated for the next agent, and so on; and then the entire process is repeated for all agents until no further strategies can be eliminated.

At the end, the remaining joint strategy space gives a reduced form of the game which can be much easier to analyse or reason over, and in some cases, it can even return unique solution to the game (although this is certainly not guaranteed in general).

Question: This has strong similarities to which algorithm for CSP?

Dominant strategy equilibrium

A second approach is to find dominant strategies and look for a **dominant strategy equilibrium** (DSE).

A DSE is a strategy profile in which each agent's strategy weakly dominates all other strategies:

$$u_i(\sigma_i^{\mathrm{DSE}}, \sigma_{-i}) \geq u_i(\sigma_i, \sigma_{-i}) \quad \forall \sigma_i \neq \sigma_i^{\mathrm{DSE}},\ \forall i \in I \tag{3}$$

That is, an agent's strategy least as good as all other strategies, irrespective of the strategies of all other players.

DSE is a strong notion of equilibrium, and is not guaranteed exist; it appears only in limited contexts.

Nash equilibrium

Best response

The notion of Nash equilibrium relies on a concept of an agent's *best response* to other the strategies of other agents.

A best response for an agent i to the strategies σ_{-i} of the other agents is a strategy that has maximal utility for that agent.

That is, σ_i is a **best response** to σ_{-i} if, for all other strategies σ_i' for agent i:

$$u_i(\sigma_i, \sigma_{-i}) \geq u_i(\sigma_i', \sigma_{-i})$$

Nash equilibrium

A strategy profile σ^* is a **Nash equilibrium** if strategy σ_i^* is a best response to σ_{-i}^* for every agent i:

$$u_i(\sigma^*) - u_i(\sigma_i, \sigma_{-i}^*) \geq 0 \quad \forall\ \sigma_i \in \Delta(A_i),\ \forall\ i \in I.$$

That is, a Nash equilibrium is a strategy profile such that no agent can do better by unilaterally deviating from that profile.

Nash equilibrium

Key results:

1. One of the great results of game theory, proved by Nash [1950], is that every finite game has at least one Nash equilibrium.
2. A second great result (in the same paper) regards symmetric games. These is a game in which all players have the same actions and symmetric payoffs given each individual's action. Every finite symmetric game has at a symmetric Nash equilibrium, in which actions are played with the same probability by all players.
3. A zero sum game is a game in which the payoffs for all players in each outcome sum to zero. Another useful result of game theory is that in every finite two-player zero-sum game, every Nash equilibrium is equivalent to a mixed-strategy minimax outcome.
4. Putting 2 and 3 together, in the equilibrium of a two-player, symmetric, zero-sum game, each player must receive a payoff of 0, and two-player, symmetric, zero sum games always have equilibria in symmetric strategies.

(Nb: Any constant-sum game can be normalised to make it equivalent to a zero-sum game.)

Calculating the Nash equilibrium

To compute a Nash equilibrium for a game in normal form, there are three steps:

1. Eliminate dominated strategies.
2. Determine which remaining actions should be in the *support set*, the set of actions which have non-zero probabilities.
3. Determine the probability for the actions in the support set.

It turns out that the second of these is the most difficult.

However, in games with only two actions (or if iterative elimination of dominated strategies leaves only two actions), we don't have to think about it.

		goalie	
		left	right
kicker	left	0.6	0.2
	right	0.3	0.9

Matrix shows probability of a goal

If the goalkeeper dives right, the probability of a goal is $P(goal|right) = 0.9p_k + 0.2(1 - p_k)$.

If the goalkeeper dives left, the probability of a goal is $P(goal|left) = 0.3p_k + 0.6(1 - p_k)$.

The only time the goalkeeper would randomise is if these are equal; that is, if

$$0.9p_k + 0.2(1 - p_k) = 0.3p_k + 0.6(1 - p_k)$$

Solving for p_k gives 0.4.

		goalie	
		left	right
kicker	left	0.6	0.2
	right	0.3	0.9

Matrix shows probability of a goal

Question: At what probability of p_g is the kicker indifferent to kicking left or right, i.e. induces the kicker to randomise?

Answer: $p_g = 0.3$.

So there is Nash equilibrium at $(p_k, p_g) = (0.4, 0.3)$. See this example at: `https://artint.info/2e/html/ArtInt2e.Ch11.S4.SS1.html`

Let’s play Morra

Morra

In the game of Morra, each player shows either one or two fingers and announces a number between 2 and 4. If a player's number is equal to the sum of the number of fingers shown, then her opponent must pay her that many dollars. The payoff is the net transfer, so that both players earn zero if both or neither guess the correct number of fingers shown.

In this game, each player has 6 strategies:

- she may show one finger and guess 2;
- she may show one finger and guess 3;
- she may show one finger and guess 4; or
- she may show two fingers and guess one of the three numbers.

a) There are two weakly dominated strategies in Morra. What are they?

Morra

a) There are two weakly dominated strategies in Morra. What are they?

1. It never pays to put out one finger and guess that the total number of fingers will be 4, because the other player cannot put out more than two fingers.
2. Likewise, it never pays to put out two fingers and guess that the sum will be 2, because the other player must put down at least one finger.

This leaves the following reduced payoff matrix for a player:

	12	13	23	24
12	0	2	-3	0
13	-2	0	0	3
23	3	0	0	-4
24	0	-3	4	0

Morra

	12	13	23	24
12	0	2	-3	0
13	-2	0	0	3
23	3	0	0	-4
24	0	-3	4	0

b) Imagine that player A can read player B's mind and guess how he plays before he makes his move. What pure strategy should player B use?

Morra

c) Player B consults a textbook and decides to use randomisation to improve his performance in Morra. Ideally, if he can find a best mixed strategy to play, what would be his expected payoff?

Trick question - it depends what his opponent does! Will his opponent adapt?

What happens if Player B know that Player A favours one strategy over another?

What happens if Player B know that Player A randomises uniformly over all strategies?

If player A can still read his mind to see his mixed strategy, how should player B approach this?

Since this is a symmetric two-player zero-sum game we know that an equilibrium must be symmetric. What does this imply about the equilibrium rewards?

Morra

d) One possible mixed strategy is to play show one finger and call “three” with probability 0.6, and to show two fingers and call “three” with probability 0.4 (and play the other strategies with probability 0). Is this a Nash equilibrium strategy? Assume that Player B is risk neutral with respect to the game payoffs.

Good luck!

Attributions and References

Thanks to Dr Alina Bialkowski and Dr Hanna Kurniawati for their materials.

Many of the slides in this course are adapted from David Poole and Alan Mackworth, *Artificial Intelligence: foundations of computational agents*, 2E, CUP, 2017 `http://https://artint.info/`.

Other materials derived from Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, 3E, Prentice Hall, 2009.

All remaining errors are Archie's — please email if you find any: archie.chapman@uq.edu.au