Logic-based reasoning about strategic abilities of socially interacting rational agents

Lecture 2: Logics for temporal strategic reasoning with incomplete and imperfect information

Valentin Goranko
Stockholm University

2nd Logic for the AI Spring Summer School
Lago di Como, September 2-5, 2024

Outline

- Modelling strategic reasoning with incomplete and imperfect information.
- ATL with incomplete/imperfect information.
- Epistemic extensions of ATL.
- An application to multi-agent planning.
- Concluding remarks.

Strategic reasoning with incomplete information

Incomplete information vs imperfect information

The decision making and abilities of strategically reasoning players crucially depend on the knowledge they possess about the game/system, other players' abilities and goals, etc.

So far I have considered structures of *complete and (almost) perfect information*. In reality this is seldom the case.

Thus, the question arises:
what can players achieve in a game/MAS if they are not completely informed about its structure and the current play?

Note the distinction:

– incomplete information: about the game structure, rules, players' possible actions, etc.
– imperfect information: about the play (current state, players' actions, etc

Incomplete information usually implies imperfect information.

On the other hand, any game of incomplete information can be regarded as a game of imperfect information that starts with Nature choosing a game in a way that players are not perfectly informed about. (Harsanyi's reduction)

Hereafter, both terms will be used almost interchangeably.

Knowledge and strategic abilities of agents

Consider:

- The agent **a** has a strategy to eventually achieve the goal γ
- The agent **a** knows that she has a strategy to eventually achieve γ.
- The agent **a** knows a strategy to achieve γ

Clearly, these are different, and only the last version implies practical ability.

Knowledge and strategies of coalitions

For coalitions, things become even more complicated. Compare:

- The coalition A has a joint strategy to achieve γ.
- Every agent in the coalition A knows that the coalition has a joint strategy to achieve γ.
- It is a common knowledge in the coalition A that it has a joint strategy to achieve γ.
- Every agent in A knows a joint strategy to achieve γ.
- A joint strategy to achieve γ is a common knowledge in A.

Can any of these guarantee that a coalition of rational players can achieve the goal γ?

The coordinated attack problem revisited

Two armies, positioned at the opposite sides of a castle intend to attack the common enemy in the castle.

They can only succeed if they attack together and simultaneously.

The armies have two choices: to attack at dawn or to attack at dusk tomorrow.

In order to coordinate the attack, the army generals must exchange messages via messenger. However, he can be captured by the enemy on his way there, or on his way back, or ...

Thus, it can be proved that coordination (i.e., common knowledge of the time of the planned attack) in this situation is impossible.

Describing the coordinated attack problem in ATL

The armies: A_1, A_2; 'coordination': C; 'victory': V.

It is known that $\neg\langle\langle A_1, A_2\rangle\rangle\mathcal{F} C$ holds.

But: $\neg\langle\langle A_1, A_2\rangle\rangle\mathcal{F} C \rightarrow \neg\langle\langle A_1, A_2\rangle\rangle\mathcal{F} V$ should be assumed valid.

These together lead to the conclusion that $\neg\langle\langle A_1, A_2\rangle\rangle\mathcal{F} V$.

However, intuitively the coalition $\{A_1, A_2\}$ does have a strategy to win, e.g., by both armies attacking simultaneously at 8am.

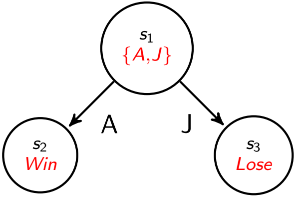
So, is $\langle\langle A_1, A_2\rangle\rangle\mathcal{F} V$ true, after all?

It depends on whether they can coordinate.

The ace and joker game

Consider the following game:

Two cards, **Ace** and **Joker**, lie face down. The player must choose one. The Ace wins, the Joker loses.



Does the player have a strategy to win the game?

Does the player know that she has a strategy to win the game?

Does the player know a strategy to win the game?

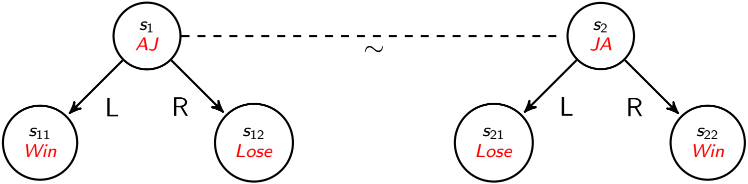
Again, does the player have a strategy to win the game?

It depends on what 'strategy' in the case of incomplete information means.

The game model above is incorrect and misleading!

A concurrent game model with incomplete information modelling the Ace and Joker game

There are two possible initial states, not one! They are AJ and JA.
They lead to two different game trees:



The player cannot distinguish states s_1 and s_2.

Concurrent game models with incomplete information (CGMII):
add an indistinguishability relation on states for each agent.

NB: indistinguishable states for an agent must enable the same actions for that agent.

Concurrent game models with incomplete information

Concurrent game models with incomplete information (CGMII):

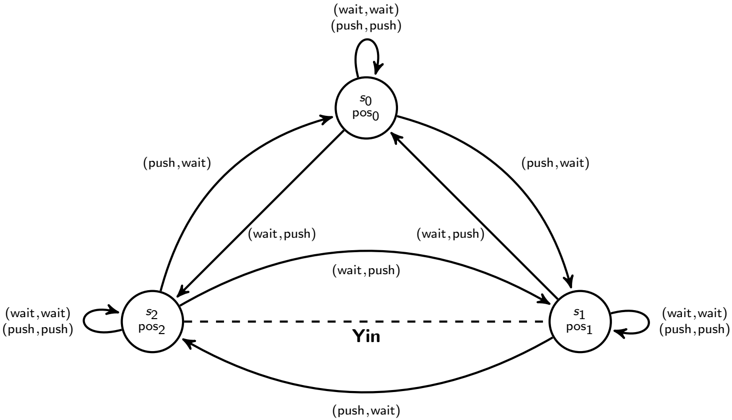
$$\langle \mathbb{A}, S, \{\sim_{\mathbf{i}}\}_{\mathbf{i}\in\mathbb{A}}, \mathsf{Act}, \mathsf{act}, \mathsf{out}, \mathsf{AP}, \mathsf{L} \rangle$$

where:

- $\langle \mathbb{A}, S, \{\sim_{\mathbf{i}}\}_{\mathbf{i}\in\mathbb{A}}, \mathsf{AP}, \mathsf{L} \rangle$ is a multi-agent epistemic model.
- $\langle \mathbb{A}, S, \mathsf{Act}, \mathsf{act}, \mathsf{out}, \mathsf{AP}, \mathsf{L} \rangle$ is a concurrent game model.
- For every $s_1, s_2 \in S$ and $\mathbf{i} \in \mathbb{A}$ such that $s_1 \sim_{\mathbf{i}} s_2$ it holds that

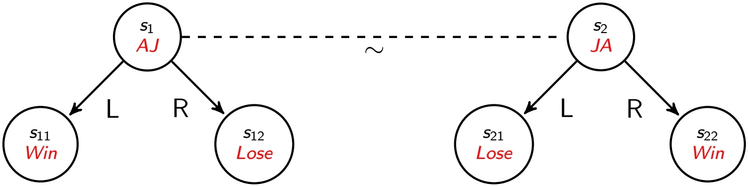
$$\mathsf{act}(\mathbf{i}, s_1) = \mathsf{act}(\mathbf{i}, s_2).$$

Concurrent game models with incomplete information: a variation of the two-robots example



Yin cannot distinguish states s_1 and s_2.

Uniform strategies



The problem with the Ace-Joker game is that the strategy "choose the Ace" is not executable in this game!

For a strategy to be executable by a player with imperfect information, it must be uniform: one that prescribes the same actions at indistinguishable states.

The player does not have a uniform strategy to ensure winning at both s_1 and s_2.