A Subset Construction for Epistemic Reasoning over Multi-Agent Games

Edvin Lundberg Dilian Gurov

KTH Royal Institute of Technology, Stockholm

1 Introduction

In Game Theory, *imperfect information* refers to a player's uncertainty about the current state of the game. Over 30 years ago, Reif [2] proposed a strategy-preserving construction (similar to the standard language-preserving *subset construction* in Automata Theory that converts a non-deterministic finite automaton to a deterministic one) to reduce two-player games with imperfect information to two-player games with perfect information, albeit with a worst-case exponential increase of the number of states. We are interested in generalisations of this construction to games of multiple agents.

Reif's construction is concerned with the knowledge of the agent (the one we are siding with in a two-player game). The knowledge is the set of game states that the agent considers possible. For two-player games with observable ω-regular objectives, this is sufficient for the existence of a strategy-preserving construction. In the case of multiple agents, however, it can be necessary for an agent to consider also what other agents consider as possible states, or even the knowledge about knowledge of the other agents, etc.

We propose a generalisation of Reif's construction for multi-agent games. The construction does not necessarily eliminate imperfect information, but it may decrease it by epistemic reasoning. The construction can be applied iteratively, capturing deeper and deeper nesting of knowledge. So, after the first application, the knowledge of an agent consists of the states it considers possible. After two applications, it also captures what the agent considers as the possible knowledge of the other agents. Thus, the construction can be applied iteratively to capture the *nested knowledge* of the agents. This knowledge can be important since, for instance, a memoryless winning strategy may exist at a certain epistemic level of depth, but not at the lower levels.

2 Multi-Agent Games

A multi-agent game is a game where a coalition of agents collaborates against an opponent *nature*. We denote the n agents of the coalition as agent 0, agent 1, etc. Following [1], we refer to a tuple of elements $(x_i)_{i<n} = (x_0, x_1, ..., x_{n-1})$, one for each agent, as a *profile*. For instance, an *action profile* $(\sigma_i)_{i<n}$ is a tuple consisting of an action for each agent.

Definition 1 (iCGS)**.** A concurrent game structure (of imperfect information) is a tuple $G = \langle L, l_0, \Sigma, \Delta, (\sim_i)_{i<n} \rangle$, where:

- L is a finite set of *states* (or *labels*),
- l_0 is the *initial state*,
- Σ is an alphabet of *actions*,
- $\Delta \subseteq L \times \Sigma^n \times L$ is a transition relation, with transitions labelled by action profiles. Additionally, there must always be an outgoing transition from each state (reachable from the initial state) to ensure infinite plays,
- For each agent i, $\sim_i \subseteq L \times L$ is an *indistinguishability* equivalence relation over L. The elements of the quotient set $\mathcal{O}_i = L/{\sim_i}$ are referred to as the *observations* of agent i. We require that if $l \sim_i l'$, then the enabled actions for agent i should be the same in both states.

The game evolves by each agent of the coalition simultaneously and independently choosing an available action. *Nature* then resolves non-determinism and each player makes an observation describing a subset of the states that the game can be in. The *plays* of a game G are defined as the infinite sequences of states and action profiles respecting the rules of the game, that is:

$$\mathsf{Plays}(G) = \{ \ \ell_0 a_1 \ell_1 a_2 ... \mid \ell_0 = l_0 \wedge \forall i \geq 0 : (\ell_i, a_{i+1}, \ell_{i+1}) \in \Delta \ \}$$

An *objective* ϕ in a game G is expressed as a subset of $\mathsf{Plays}(G)$. *Histories* are finite prefixes of plays ending in a state. As we lift plays and histories to record observations instead of states, we define these *locally*, that is from the perspective of an agent. Given a play $\ell_0 a_1 \ell_1 a_2 \ldots$, the corresponding *local play* for agent i is $[\ell_0]_{\sim_i} a_{1,i} [\ell_1]_{\sim_i} a_{2,i} \ldots$, where $a_{1,i}$ is the i-th element of the action profile a_1. Let $\mathsf{Prefs}_i(G)$ denote the set of *local histories* of agent i, that is, all finite prefixes of its local plays. A (uniform) *strategy* for agent i is a function $\alpha : \mathsf{Prefs}_i(G) \to \Sigma$. In other words, an agent can only base its choice of action on its local perspective of the game. A strategy is *memoryless* if it only depends on the last observation in the history. A *joint strategy* $\prod = (\alpha_0, \alpha_1, ..., \alpha_{n-1})$ is a strategy profile assigning each agent of the coalition an individual strategy. A joint strategy is *winning* w.r.t. a given objective ϕ if all plays of G that respect the strategy (called the *outcomes* of the strategy) are in ϕ.

Given a game G and an objective ϕ, the *strategy synthesis problem* is the problem of determining whether there exists a joint winning strategy for the coalition, and construct such a strategy if one exists.

Example 2.1. There has been an accident at a chemical plant in your town. A full glass of very dangerous acid is on the floor. Luckily, the acid is all in the glass. We dare not send in humans to remove the glass, but we have two simple robots. The robots each have a small gripper with which they can lift

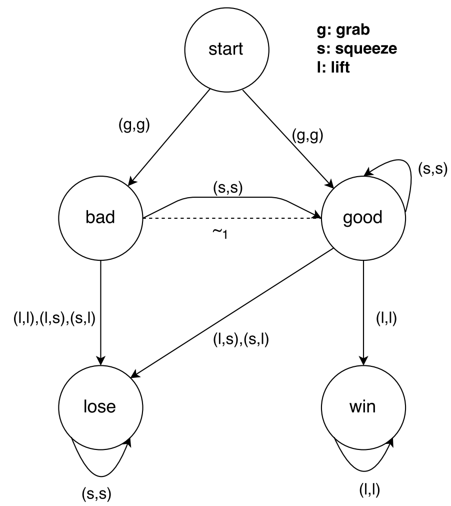

Figure 1: An iCGS G for the chemical accident scenario, in which robot 0 has perfect information, while states *good* and *bad* look the same to robot 1.

objects. The robots are positioned on opposite sides of the glass. The glass is too heavy for one robot to lift, so they have to lift simultaneously. Both robots are equipped with a small optic sensor that can detect if they are touching the glass or not. The unknown is the humidity, which the opponent *nature* controls. If the robots grab the glass and the humidity is too high, the initial grip will be bad. They can squeeze to improve their grip. The first robot (robot 0) has a grip sensor that can detect whether or not they (both) have a good enough grip. Unfortunately, that sensor is broken in robot 1. The robots have synchronised clocks, so they act simultaneously. You need to synthesise a joint uniform strategy that guarantees that the glass is lifted without falling.

Your colleague has produced the iCGS in Figure 1 to model the problem. Note that there is no memoryless uniform winning strategy for this game. The problem is that robot 1 has to take the same action in states *good* and *bad*. If it commits to squeezing, the robots can never lift the glass, while if it commits to lifting, *nature* can always choose to make the game go to the state *bad* initially and then robot 1 will lift the glass, but they will drop it or tip it over. □

3 Epistemic Unfolding

Instead of giving a formal definition here, we shall show how we can transform the game to capture deeper and deeper knowledge of the agents. We call this construction the *knowledge-based subset construction* (KBSC), just like Reif's construction.

Consider the iCGS G^K shown in Figure 2, which is the result of applying our knowledge-based subset construction to the game G in Example 2.1. To

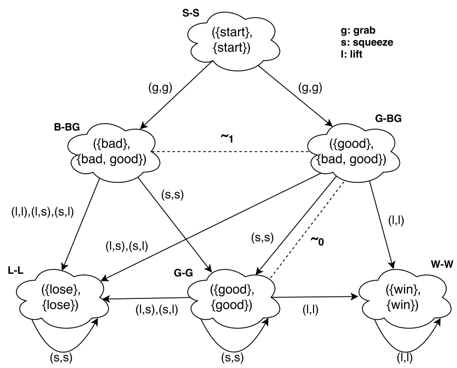

Figure 2: The iCGS G^K. If both agents squeeze, they will both know that they have a good grip. However, since robot 0 is only concerned with its own knowledge, it cannot distinguish between the state where *"they have a good grip, but robot 1 does not know it"* and the state where *"they have a good grip, and robot 1 knows it"*.

shorten the notation, we label the knowledge states with new names, for example **B-BG** $= \{(\{bad\}, \{bad, good\})\}$, which is the state where robot 0 knows that they have a bad grip, and robot 1 knows that they have a bad or good grip. Can we find a memoryless uniform strategy for both robots in G^K? It turns out that we can not. The reason is that the only way to win is from states **G-BG** and **G-G**, as those are the only states adjacent to the winning state. However, these states are indistinguishable to robot 0, so it has to choose the same action in both of them. Clearly, robot 0 has to lift in these states, otherwise the winning state can never be reached.

We now look at robot 1, who cannot distinguish between states **B-BG** and **G-BG**. If robot 1 lifts in those states, the adversarial *nature* can spoil the strategy by choosing state **B-BG** initially (high humidity). Then the robots will have a bad grip and robot 1 will try to lift and they will fail. The only remaining option for robot 1 is to squeeze in both of these states.

The problem is now that in state **G-BG** we know that robot 0 must lift, and robot 1 must squeeze. Thus, since nature can force the game into that state initially, the robots would loose. (Curiously, if robot 0 did not have the grip detector, there would be a memoryless winning strategy at this point!)

Intuitively, we need to make robot 0 aware of the knowledge of robot 1. Before lifting, robot 0 has to know that robot 1 knows that they have a good grip. Another way of putting this, is that we would like robot 0 to keep track of the possible global knowledge states (that is, the states in G^K that are possible at any given time).

Interestingly, this is precisely what is achieved after another knowledge-based subset construction, this time over the iCGS G^K. The resulting iCGS, denoted G^{2K}, is depicted on Figure 3. The agents now have some awareness of each others knowledge. Can we find a memoryless uniform strategy for both

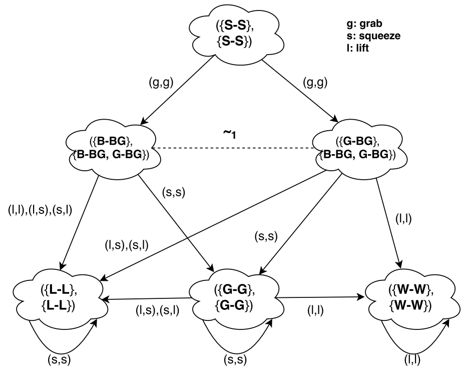

Figure 3: The iCGS G^{2K}. When the robots initially grab the glass, and nature decides that it is slippery, this game will go the the state where robot 0 knows that the game G^K is in state **B-BG**, in other words it knows that it knows that they have a bad grip and it knows that robot 1 does not know whether they have a good or a bad grip. In this state, robot 1 knows that the game is in one of the states in {**B-BG**, **G-BG**}. That is, it knows that it does not know whether they have a good or bad grip, but it knows that robot 0 knows that.

robots in this game? Yes, we can! Intuitively, both robots initially grab the glass. Then, they both squeeze it, regardless of whether the grip is good or bad. The game is then forced into the state where they both know, that they both know, that they have a good grip. This time, robot 0 can distinguish between that state and the one where they already have a good grip but robot 1 does not know whether they have a good or bad grip. They both lift from this state and reach the winning state.

In addition to the above construction, we also have a transformation that *recovers* the memoryless strategies of the game G^{mK}, resulting from m applications of our construction, to the original game G. However, in the original game G the strategy is represented by a finite-state *transducer* that updates the knowledge of the agent given what action it took and the resulting observation. The state of the transducer represents the knowledge of the agent and determines its next action.

References

[1] Dietmar Berwanger, Lukasz Kaiser, and Bernd Puchala. “A perfect-information construction for coordination in games”. In: *LIPIcs-Leibniz International Proceedings in Informatics*. Vol. 13. 2011.

[2] John H. Reif. “The complexity of two-player games of incomplete information”. In: *Journal of Computer and System Sciences* 29.2 (1984), pp. 274–301.