



**HAL**  
open science

## A Logical Modeling of the Yōkai Board Game

Dominique Longin, Emiliano Lorini, Frédéric Maris

► **To cite this version:**

Dominique Longin, Emiliano Lorini, Frédéric Maris. A Logical Modeling of the Yōkai Board Game. [Research Report] IRIT - Institut de recherche en informatique de Toulouse. 2020. hal-02983253v2

**HAL Id: hal-02983253**

**<https://hal.science/hal-02983253v2>**

Submitted on 2 Dec 2020 (v2), last revised 9 Apr 2021 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Institut de Recherche en Informatique de Toulouse  
CNRS - INP - UT3 - UT1 - UT2J

# A Logical Modeling of the Yōkai Board Game

Dominique Longin<sup>†</sup>, Emiliano Lorini<sup>†</sup>, Frédéric  
Maris<sup>‡</sup>

<sup>†</sup>CNRS, IRIT

<sup>‡</sup> Université Paul Sabatier, IRIT

{Firstname.Surname@irit.fr}

Research report, october 29, 2020 (update: November 26, 2020)

## Abstract

We present an epistemic language for representing an artificial player’s beliefs and actions in the context of the Yōkai board game. Yōkai is a cooperative game which requires a combination of Theory of Mind (ToM), temporal and spatial reasoning to be played effectively by an artificial agent. We show that the language properly accounts for these three dimensions and that its satisfiability problem is NP-complete. This opens up the possibility of exploiting SAT techniques for automating reasoning of an artificial player in the context of the Yōkai board-game.

## 1 Introduction

When one wishes to model socio-cognitive agents and, in particular, agents endowed with a Theory of Mind (ToM) who are capable of reasoning about other agents’ beliefs, some of the privileged tools are epistemic logic (EL) [12, 14] and its extensions by informative and communicative extensions such as public and private announcements [13, 19, 3]. The latter belongs to the Dynamic Epistemic Logic (DEL) family [21].

The major disadvantage of EL and DEL is that they have most of the time a high complexity thereby making them not very well-suited for practical applications. In particular, it has been shown that extending multi-agent EL by simple notions of state eliminating public announcement or arrow eliminating private announcement does not increase its PSPACE complexity (see, e.g., [18, 7]). However, the satisfiability problem of full DEL with public, semi-private and private communicative actions was shown to be NEXPTIME-complete [1]. The situation is even worse in the context of epistemic planning: it is known that epistemic planning in public announcement logic (PAL) is decidable, while it becomes undecidable in full DEL, due to the fact that the epistemic model may grow as a consequence of a private announcement [6].

In [16, 15], a variant of epistemic logic with a semantics exploiting belief bases is introduced. It distinguishes explicit belief from implicit belief. The former is modeled as a fact in an agent’s belief base, while the latter is modeled as a fact that is deducible from the agent’s explicit beliefs. The main advantages of the belief base semantics for epistemic logic compared to the standard possible world semantics based on multi-relational structures (so-called Kripke models) are (i) its compactness, and (ii) its closeness to the way artificial cognitively-inspired agents are traditionally modeled in AI and in the area of knowledge representation and reasoning (KR) by adopting a database perspective [20]. In [17], it is shown that this variant of epistemic logic provides a valuable abstraction for modeling multi-robot scenarios in which each robot is endowed with a ToM whereby being able to ascribe epistemic states to the other robots and to reason about them.<sup>1</sup>

---

<sup>1</sup>See also [5, 9] for a DEL-based approach to modeling and implementing ToM on social robots.

In this paper, we leverage the belief base semantics for epistemic logic to model interaction in the context of the cooperative board-game Yōkai.<sup>2</sup> We consider its two-player variant in which an artificial agent has to collaborate with a human agent to win it and to obtain the best score as possible. Yōkai is an interesting testbed for artificial agents, as it covers a lot of epistemic and strategic reasoning aspects as well as planning and belief revision aspects. The idea of testing the performance of artificial agents in the context of cooperative board-games in which ToM reasoning plays a role is not new. Some works exist about modeling and implementing artificial players for the card game Hanabi [4, 11, 10]. Yōkai adds to the ToM dimension, which is central in Hanabi, the temporal and spatial dimension. First of all, in Yōkai a player’s performance relies on her/its capacity to remember the cards she/it and the other player have seen in the past. Secondly, the players must move cards in a shared space and there are spatial restrictions on card movements that should be taken into account by the players. More generally, the interesting feature of Yōkai, from the point view of KR, is the combination of epistemic, temporal and spatial reasoning that is required to completely apprehend all the game facets and dimensions.

The main novelty of our approach to modeling artificial board-game players is the use of SAT-based techniques. Specifically, the language we present for representing the artificial player’s beliefs about the static and dynamic aspects of the game as well as about the human player’s beliefs has the same complexity as SAT and can be polynomially translated into a propositional logic language. This opens up the possibility of exploiting SAT techniques for automating reasoning of the artificial player in the context of the Yōkai board-game.

The paper is organized as follows. In Section 2, we explain the rules of Yōkai and clarify the representation and reasoning requirements that are necessary for the artificial player to be able to play the game in a clever way. In Section 3, we introduce the specification language for modeling the artificial player’s actions and beliefs about the game properties and about the human player’s beliefs. It is a timed language for explicit and implicit belief with a semantics exploiting belief bases. The main novelty compared to the epistemic language presented in [16] is the temporal component: the artificial player modeled in the language has knowledge about the current time of the game and beliefs about current and past beliefs of the human player. Section 4 is devoted to the encoding of the game in the language. We first model the static aspects, namely, the artificial agent’s beliefs about the rules and properties of the game. Then, we focus on the dynamic aspects by modeling the artificial player’s actions and how they are used for planning a sequence of moves at a given round of the game.

## 2 Game description

In this section, we explain the rule of the Yōkai game. As pointed out in the introduction, we only consider the two-player variant of the game with an

---

<sup>2</sup><https://boardgamegeek.com/boardgame/269146/ykai>

artificial and a human player.

**Colored cards and actions.** There are 4 types of card, red, yellow, green and blue cards, and 4 cards for each color. This gives us a total amount of 16 cards. The cards are placed face down in a grid of size 4x4 in a random way, so that the players cannot see their colors. The goal of the game is to gather together the cards of the same color. Figure 1a illustrates a winning configuration of the game.

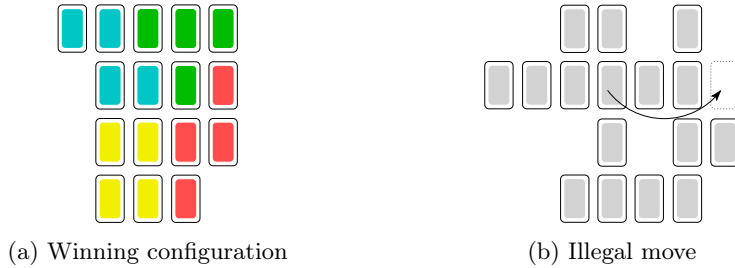


Figure 1: Examples of game configuration

The players play sequentially in each round of the game. At each round, each player has to play the following 4 actions in the order she/it prefers:

1. to look at two cards privately (2 actions),
2. to move one card from its current position to a new position adjacent to another card (that is, linked to the latter by one of its sides) and without separating the cards into two disjointed groups,
3. either to activate a hint from the set of available hints (see next paragraph) or to disclose an information by marking one card with an active hint.

Figure 1b portrays an example of illegal move. Indeed, the move will make the first and second cards on the left bottom corner of the grid separated from the rest of the cards.

**Hints.** Three types of hint are available: 1-color, 2-color and 3-color hints. Specifically, 1-color, 2-color and 3-color hints indicate, respectively, a single color, two colors and three colors in the set {red, green, blue, yellow}. There are 14 different hints available, but only seven hints are randomly selected at the beginning of the game: two occurrences of 1-color hint, three occurrences of 2-color hint and two occurrences of 3-color hint.

A hint can be used after being activated, and only available hints can be activated. Moreover, it can be used only once throughout the game. The state of a hint is unique: it is either available, activated, or marking. In particular, after a hint has been activated and used to mark a card, it is no longer available or activated. For example, suppose a player just looked at one card and discovers

it is a red card. She can mark it with an active 2-color hint of type red/green to inform the other player that it is either a red card or a green card.

A hint is properly used if one of its colors matches with the color of the card it marks. For example, if a player uses an active 2-color card of type green/blue to mark a blue card, then the hint is properly used. Conversely, if the player uses it to mark a red card, then the hint is improperly used. Note that unless a mistake has been made, hints must be used properly.

Finally, when an active hint marks a card, that card can no longer be moved or observed.

**End of the game.** When a player thinks that the goal state is achieved, she/it can decide to stop the game. Otherwise, the game ends when all hints have been used by the players. At the end of the game, the cards are turned face up. If the goal state is achieved, the players win the game. Otherwise, they lose it. In case of win, the final score is calculated as follows:  $score = x_1 - x_2 + 2x_3 + 5x_4$  where  $x_1$  is the number of properly used hints,  $x_2$  is the number of improperly used hints,  $x_3$  is the number of activated hints that were not used by the players, and  $x_4$  is the number of non-activated hints. The final score ranges in the interval  $[-7, 35]$ .

**Requirements.** As emphasized in the introduction, in order to be able to reason about the static and dynamic aspects of the game, a player must have beliefs about:

- the other player’s actual beliefs (*ToM reasoning*);
- the current positions of the cards and the executable card movements, given the current spatial configuration of the game (*spatial reasoning*);
- the color of the cards she/it observed in the past as well as the other player’s past observations (*temporal reasoning*).

### 3 A timed language for explicit and implicit belief

This section presents a two-agent timed variant of the language and the semantics of the logic of explicit and implicit belief presented in [16]. The two agents are the artificial agent (or machine)  $m$  and the human user  $h$ . Agents  $m$  and  $h$  are treated asymmetrically. Our language allows us to represent (i)  $h$ ’s explicit beliefs at different points in a game sequence, and (ii)  $m$ ’s actual explicit and implicit beliefs, namely,  $m$ ’s explicit and implicit beliefs at the current time point of the game sequence. Following [16], explicit beliefs are defined to be beliefs in an agent’s belief base, while implicit beliefs are those beliefs that are derivable from the agent’s explicit beliefs.

We first present the static language in which agent  $\mathbf{m}$ 's beliefs do not change. Then, we present a dynamic extension in which agent  $\mathbf{m}$ 's belief base can be expanded by new information.

### 3.1 Static language

Assume a countably infinite set of atomic propositions  $ATM$ . We define the language in two steps.

We first define the language  $\mathcal{L}_0(ATM)$  by the following grammar in BNF:

$$\alpha ::= p^t \mid \Delta_{\mathbf{h}}^t \alpha \mid now^{\geq t} \mid \neg \alpha \mid \alpha_1 \wedge \alpha_2 \mid \Delta_{\mathbf{m}} \alpha,$$

where  $p$  ranges over  $ATM$ .  $\mathcal{L}_0(ATM)$  is the language for representing agent  $\mathbf{h}$ 's timed explicit beliefs and agent  $\mathbf{m}$ 's actual explicit beliefs. Specifically, the formula  $\Delta_{\mathbf{h}}^t \alpha$  is read “agent  $\mathbf{h}$  explicitly believes that  $\alpha$  at time  $t$ ”, whilst  $\Delta_{\mathbf{m}}$  is read “agent  $\mathbf{m}$  actually has the explicit belief that  $\alpha$ ”. Atomic propositions are assumed to be timed:  $p^t$  is read “atomic proposition  $p$  is true at time  $t$ ”. Finally, formula  $now^{\geq t}$  provides information about the current time point. It is read “the actual time of the game play is at least  $t$ ”.

Then, we define  $\mathcal{L}_0^T(ATM)$  to be the subset  $\mathcal{L}_0(ATM)$  including only timed formulas, that is:

$$\begin{aligned} \mathcal{L}_0^T(ATM) = & \{p^t : p \in ATM \text{ and } t \in \mathbb{N}\} \cup \\ & \{\Delta_{\mathbf{h}}^t \alpha : \alpha \in \mathcal{L}_0(ATM) \text{ and } t \in \mathbb{N}\} \cup \\ & \{now^{\geq t} : t \in \mathbb{N}\}. \end{aligned}$$

Elements of  $\mathcal{L}_0^T(ATM)$  are denoted by  $x, y, \dots$

The language  $\mathcal{L}(ATM)$  extends the language  $\mathcal{L}_0(ATM)$  by a modal operator of implicit belief for agent  $\mathbf{m}$  and is defined by the following grammar:

$$\varphi ::= \alpha \mid \neg \varphi \mid \varphi_1 \wedge \varphi_2 \mid \Box_{\mathbf{m}} \alpha,$$

where  $\alpha$  ranges over  $\mathcal{L}_0(ATM)$ . For notational convenience we write  $\mathcal{L}_0$  instead of  $\mathcal{L}_0(ATM)$ ,  $\mathcal{L}_0^T$  instead of  $\mathcal{L}_0^T(ATM)$  and  $\mathcal{L}$  instead of  $\mathcal{L}(ATM)$ , when the context is unambiguous. The formula  $\Box_{\mathbf{m}} \alpha$  is read “agent  $\mathbf{m}$  actually has the implicit belief that  $\alpha$ ”. The other Boolean constructions  $\top$ ,  $\perp$ ,  $\vee$ ,  $\rightarrow$  and  $\leftrightarrow$  are defined in the standard way. Notice that only formulas from the sublanguage  $\mathcal{L}_0$  can be in the scope of the implicit belief operator  $\Box_{\mathbf{m}}$ . Therefore, nesting of this operator is not allowed (e.g.,  $\Box_{\mathbf{m}} \neg \Box_{\mathbf{m}} p^t$  is not a well-formed formula). As we will show at the end of the section, this syntactic restriction on our language is useful to make the complexity of its satisfiability problem the same as the complexity of SAT.

The interpretation of the language  $\mathcal{L}$  exploits the notion of belief base. While the notions of possible world and epistemic alternative are primitive in the standard Kripke semantics for epistemic logic [12], they are defined from the primitive concept of belief base in our semantics.

**Definition 1 (State)** A state is a tuple  $S = (B, V)$  where (i)  $B \subseteq \mathcal{L}_0$  is agent  $\mathbf{m}$ 's belief base (or, agent  $\mathbf{m}$ 's subjective view of the actual situation), (ii)  $V \subseteq \mathcal{L}_0^T$  is the actual situation, and such that, for every  $t, t' \in \mathbb{N}$ ,

$$\text{now}^{\geq 0} \in V, \quad (1)$$

$$\text{if } \text{now}^{\geq t} \in V \text{ and } t' \leq t \text{ then } \text{now}^{\geq t'} \in V, \quad (2)$$

$$\text{now}^{\geq t} \in V \text{ iff } \text{now}^{\geq t} \in B, \text{ and} \quad (3)$$

$$\text{now}^{\geq t} \notin V \text{ iff } \neg \text{now}^{\geq t} \in B. \quad (4)$$

The set of all states is denoted by  $\mathbf{S}$ .

Conditions (1) and (2) in the previous definition guarantees time consistency, namely, that the current time should be at least 0 and that if the current time is at least  $t$  and  $t' \leq t$ , then it should be at least  $t'$ . Conditions (3) and (4) capture agent  $\mathbf{m}$ 's time-knowledge, namely, the assumption that  $\mathbf{m}$  has complete information about the current time.

The sublanguage  $\mathcal{L}_0(ATM)$  is interpreted w.r.t. states as follows:

**Definition 2 (Satisfaction)** Let  $S = (B, V) \in \mathbf{S}$ . Then:

$$\begin{aligned} S \models x &\iff x \in V, \\ S \models \neg\alpha &\iff S \not\models \alpha, \\ S \models \alpha_1 \wedge \alpha_2 &\iff S \models \alpha_1 \text{ and } S \models \alpha_2, \\ S \models \Delta_{\mathbf{m}}\alpha &\iff \alpha \in B. \end{aligned}$$

Observe in particular the set-theoretic interpretation of the explicit belief operator for agent  $\mathbf{m}$ : agent  $\mathbf{m}$  actually has the explicit belief that  $\alpha$  if and only if  $\alpha$  is included in her actual belief base. This highlights the asymmetry between agent  $\mathbf{m}$  and agent  $\mathbf{h}$  in our semantics. We adopt agent  $\mathbf{m}$ 's *internal* perspective, that is, the point of view of its belief base.<sup>3</sup> On the contrary, agent  $\mathbf{h}$ 's explicit beliefs are modeled from an *external* point of view and semantically interpreted in the same way as the other timed formulas in  $\mathcal{L}_0^T(ATM)$ .

A multi-agent belief model (MAB) is defined to be a state supplemented with a set of states, called *context*. The latter includes all states that are compatible with agent  $\mathbf{m}$ 's background knowledge.

**Definition 3 (Model)** A model is a pair  $(S, Cxt)$ , where  $S \in \mathbf{S}$  and  $Cxt \subseteq \mathbf{S}$ . The class of all models is denoted by  $\mathbf{M}$ .

Note that we do not impose that  $S \in Cxt$ . When  $Cxt = \mathbf{S}$  then  $(S, Cxt)$  is said to be *complete*, since  $\mathbf{S}$  is conceivable as the complete (or universal) context which contains all possible states.

<sup>3</sup>See [2] for an in-depth logical analysis of the internal perspective on modeling knowledge and belief.



**Definition 4 (Epistemic alternatives)** We define  $\mathcal{R}$  to be the binary relation on the set  $\mathbf{S}$  such that, for all  $S = (B, V), S' = (B', V') \in \mathbf{S}$ :

$$S\mathcal{R}S' \quad \text{if and only if} \quad \forall \alpha \in B : S' \models \alpha.$$

$S\mathcal{R}S'$  means that  $S'$  is an epistemic alternative for the artificial agent  $\mathbf{m}$  at  $S$ . So  $\mathbf{m}$ 's set of epistemic alternatives at  $S$ , noted  $\mathcal{R}(S) = \{S' \in \mathbf{S} : S\mathcal{R}S'\}$ , includes exactly those states that satisfy  $\mathbf{m}$ 's explicit beliefs.

Definition 5 extends Definition 2 to the full language  $\mathcal{L}$ . Its formulas are interpreted with respect to models. We omit Boolean cases that are defined in the usual way.

**Definition 5 (Satisfaction, cont.)** Let  $(S, Cxt) \in \mathbf{M}$ . Then:

$$\begin{aligned} (S, Cxt) \models \alpha &\iff S \models \alpha; \\ (S, Cxt) \models \Box_{\mathbf{m}}\varphi &\iff \forall S' \in Cxt, \text{ if } S\mathcal{R}S' \text{ then } (S', Cxt) \models \varphi. \end{aligned}$$

A formula  $\varphi \in \mathcal{L}$  is valid in the class  $\mathbf{M}$ , noted  $\models_{\mathbf{M}} \varphi$ , if and only if  $(S, Cxt) \models \varphi$  for every  $(S, Cxt) \in \mathbf{M}$ ; it is satisfiable in  $\mathbf{M}$  if and only if  $\neg\varphi$  is not valid in  $\mathbf{M}$ . As the following theorem indicates, the satisfiability problem for  $\mathcal{L}(ATM)$  has the same complexity as SAT.

**Theorem 1** *Checking satisfiability of  $\mathcal{L}(ATM)$  formulas in the class  $\mathbf{M}$  is an NP-complete problem.*

SKETCH OF PROOF. Hardness is clear since  $\mathcal{L}(ATM)$  extends the propositional logic language. As for membership, we can find a polysize satisfiability preserving translation from  $\mathcal{L}(ATM)$  to propositional logic. The translation is divided in three steps. First, we transform the input formula in  $\mathcal{L}(ATM)$  into negated normal form (NNF). Secondly, we translate the formula in NNF into a restricted mono-modal language with no nesting of the modal operator. Thirdly, we translate the latter into a propositional logic language in a way similar to the standard translation of modal logic into FOL. We take care of translating a finite theory including axioms corresponding to the four constraints of Definition 1. The axioms have the following form:  $now^{\geq 0}, now^{\geq t} \rightarrow now^{\geq t'}$  for  $t' \leq t$ ,  $now^{\geq t} \leftrightarrow \Delta_{\mathbf{m}} now^{\geq t}$  and  $\neg now^{\geq t} \leftrightarrow \Delta_{\mathbf{m}} \neg now^{\geq t}$ . The theory is finite since we only need to consider instances of the axioms whose symbols occur in the input formula. For example, if  $t' \leq t$  and both  $now^{\geq t}$  and  $now^{\geq t'}$  occur in the input formula, then  $now^{\geq t} \rightarrow now^{\geq t'}$  should be included in the theory, otherwise not. ■

### 3.2 Dynamic extension

Let us now move from a static to a dynamic perspective by presenting an extension of the language  $\mathcal{L}(ATM)$  with belief expansion operators. Specifically, we introduce the following language  $\mathcal{L}^+(ATM)$ :

$$\varphi ::= \alpha \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \Box_{\mathbf{m}}\alpha \mid [+_{\mathbf{m}}^t\alpha]\varphi,$$

where  $\alpha$  ranges over  $\mathcal{L}_0$  and  $t$  ranges over  $\mathbb{N}$ . The formula  $[+_m^t\alpha]\varphi$  is read “ $\varphi$  holds after agent  $m$  has privately learnt that  $\alpha$  and that the current time is at least  $t$ ” or simply “ $\varphi$  holds after agent  $m$  has privately learnt that  $\alpha$  at time at least  $t$ ”.

Our extension has the following semantics relative to a model:

**Definition 6 (Satisfaction relation, cont.)** *Let  $S = (B, V) \in \mathbf{S}$  and  $(S, Cxt) \in \mathbf{M}$ . Then,*

$$(S, Cxt) \models [+_m^t\alpha]\varphi \iff (S^{+_m^t\alpha}, Cxt) \models \varphi$$

with

$$\begin{aligned} S^{+_m^t\alpha} &= (B^{+_m^t\alpha}, V^{+_m^t\alpha}), \\ V^{+_m^t\alpha} &= V \cup \{now^{\geq t'} : t' \leq t\}, \\ B^{+_m^t\alpha} &= B \cup \{\alpha\} \cup \{now^{\geq t'} : t' \leq t\}. \end{aligned}$$

Intuitively speaking, agent  $m$ 's private learning that  $\alpha$  at time at least  $t$  simply consists in (i) adding the information  $\alpha$  to  $m$ 's belief base, and (ii) moving the objective time and  $m$ 's subjective view of time to index  $t$ .

As the following proposition indicates, the dynamic semantics given in Definition 6 is well-defined, as it guarantees that the structure resulting from a belief expansion operation belongs to the class  $\mathbf{M}$ , if the initial structure also belongs to  $\mathbf{M}$ .

**Proposition 1** *Let  $(S, Cxt) \in \mathbf{M}$ . Then,  $(S^{+_m^t\alpha}, Cxt) \in \mathbf{M}$ .*

Interestingly, adding belief expansion operators to the language  $\mathcal{L}$  does not increase the complexity of the corresponding satisfiability problem.

**Theorem 2** *Checking satisfiability of  $\mathcal{L}^+(ATM)$  formulas in the class  $\mathbf{M}$  is an NP-complete problem.*

SKETCH OF PROOF. The theorem is a consequence of Theorem 1 and the fact that we can find a polysize reduction of the satisfiability problem for  $\mathcal{L}^+(ATM)$  to the satisfiability problem for  $\mathcal{L}(ATM)$ . The reduction makes use of reduction axioms which allow us to eliminate dynamic operators from the input formula and to obtain a logically equivalent formula in  $\mathcal{L}(ATM)$ . ■

Let  $EVT = \{+_m^t\alpha : \alpha \in \mathcal{L}_0, t \in \mathbb{N}\}$  be the set of belief expansion events. It is reasonable to assume that such events have executability preconditions that are specified by the following function  $\mathcal{P} : EVT \rightarrow \mathcal{L}(ATM)$ . So, we can define the following operator of successful occurrence of an event in  $EVT$ :

$$\langle\langle+_m^t\alpha\rangle\rangle\varphi \stackrel{\text{def}}{=} \mathcal{P}(+_m^t\alpha) \wedge [+_m^t\alpha]\varphi.$$

The formula  $\langle\langle +_m^t \alpha \rangle\rangle \varphi$  has to be read “agent  $m$  can privately learn that  $\alpha$  at time  $t$  and  $\varphi$  holds after the occurrence of this learning event”.

In the next section, we will provide a formalization of the Yōkai board-game with the aid of the language  $\mathcal{L}^+(ATM)$ . We will represent agent  $m$ 's actions in the game as events in  $EVT$  affecting  $m$ 's beliefs. For every action of  $m$ , we will specify the corresponding executability precondition.

## 4 Game modeling

In this section, we first formalize all the static aspects of the game (the different rules, the initial state of the game, *etc.*) and then the aspects related to players actions. Finally, we briefly present some aspects that our language offers for planning.

### 4.1 Static aspects

Let be the following sets:

$$GRID = \{1, \dots, 32\} \times \{1, \dots, 32\},$$

$$IPOS = \{(l, c) \in GRID : l, c \in \{15, \dots, 18\}\},$$

$$COLORS = \{r, g, b, y\},$$

$$HINTS = 2^{COLORS} \setminus \{\{\}, \{r, g, b, y\}\},$$

$$AvHINTS \subseteq HINTS \quad \text{s.t.} \quad \begin{cases} |AvHINTS| = 7, \\ \exists h_1, h_2 \in AvHINTS : |h_1| = |h_2| = 1, \\ \exists h_1, h_2, h_3 \in AvHINTS : |h_1| = |h_2| = |h_3| = 2, \\ \exists h_1, h_2 \in AvHINTS : |h_1| = |h_2| = 3, \end{cases}$$

$$CARDS = \{1, \dots, 16\} \text{ and we note } CARDS^4 = \{X \in 2^{CARDS} : |X| = 4\}$$

There are seven hints at the start of the game, and each player must take turns activating a hint or using a hint to mark a card. Thus, each player plays 7 times. As each player must take 4 different actions in turn, the game lasts a maximum of  $(7 \times 4) \times 2 = 56$  time units (56 actions are executed during a game). Let  $\mathbb{N}$  be the set of time points, including initial state of the game at time 0 (and excluding it for  $\mathbb{N}^*$ ). End of the game is marked by the natural  $end \in \mathbb{N}$  such that  $end = 56$ .

However, each player only performs the action of moving a card once among the 4 actions she/it must perform during her/its turn. There are therefore a total of  $7 \times 2 = 14$  card moves. As in the initial state the 16 cards are placed in a square of  $4 \times 4$  cards, whatever the movements made during a game, the cards will evolve in a grid of  $32 \times 32$  positions represented by the set  $GRID$ , and  $IPOS$  represents the set of cards positions at the start of the game.

A hint is viewed as a subset of 1, 2 or 3 colors among 4 different colors (r for red, g for green, b for blue and y for yellow).  $HINTS$  is the set of all hints and

$|HINTS| = 14$ . When the game starts, only 7 hints are available for the game (represented by set  $AvHINTS$ ).

**Vocabulary.** We use the following notations: we denote a card by  $x \in CARDS$ , a color by  $c \in COLORS$ , a position by  $p \in GRID$ , a hint by  $h \in HINTS$  or  $h \in AvHINTS$ . Moreover, based on the sets defined above, we define the following set atomic propositions  $q^t$  (for  $q \in ATM$ ,  $t \in \mathbb{N}$ ) such that:

- $col_{x,c}^t$ : card  $x$  is of color  $c$  at time  $t$ ;
- $pos_{x,p}^t$ : card  $x$  is at position  $p$  at time  $t$ ;
- $avail_h^t$ : hint  $h$  is available at time  $t$ ;
- $active_h^t$ : hint  $h$  is activated at time  $t$ ;
- $mark_{x,h}^t$ : card  $x$  is marked with hint  $h$  at time  $t$ ;
- $neighbors_{p,p'}^t$ :  $p$  and  $p'$  are two adjacent positions (having one side in common) at time  $t$ .

Moreover, we define a function  $\sigma : CARDS \rightarrow IPOS$  that assigns to each card  $x$  a position  $p$  in the initial positions set, and a function  $NEIG : GRID \rightarrow 2^{GRID}$  that assigns to each position  $(l, c)$  the set of its neighboring positions in the grid  $NEIG(l, c) = \{(l + 1, c), (l - 1, c), (l, c + 1), (l, c - 1)\} \cap GRID$ .

In the following, we consider that belief bases of agents are divided into two (sub)bases: the *core beliefs* base (containing all the beliefs that cannot be modified during the game), and the *volatile beliefs* base (containing all the belief that may change). As we want to model the game from  $m$ 's perspective, we only model its beliefs about the facts of the game or about the beliefs of agent  $\mathfrak{h}$ .

**Initial volatile beliefs of agent  $m$ .** Volatile belief base of agent  $m$  includes the fact that initial time is at least 0, each card position is known (we assume that function  $\sigma$  is known), the cards color is not known, only seven hints are available, none is activated yet, and none marks a card yet. Moreover, all adjacent positions of each cell are known:

$$\Sigma_v \supseteq \left\{ \begin{array}{l} now^{\geq 0}, \quad \bigwedge_{x \in CARDS} pos_{x, \sigma(x)}^0, \quad \bigwedge_{\substack{x \in CARDS \\ c \in COLORS}} \neg \Delta_m col_{x,c}^0 \\ \bigwedge_{h \in AvHINTS} avail_h^0, \quad \bigwedge_{h \in HINTS \setminus AvHINTS} \neg avail_h^0, \\ \bigwedge_{h \in HINTS} \neg active_h^0, \quad \bigwedge_{\substack{x \in CARDS \\ h \in HINTS}} \neg mark_{x,h}^0, \\ \bigwedge_{\substack{p \in GRID \\ p' \in NEIG(p)}} neighbors_{p,p'}^0, \quad \bigwedge_{\substack{p \in GRID \\ p' \in GRID \setminus NEIG(p)}} \neg neighbors_{p,p'}^0, \end{array} \right\}$$

In addition, we assume that  $\Sigma_v$  also includes  $\Delta_h^0 now^{\geq 0}$  and, for each of facts above of the form  $\bigwedge_X \alpha$ , a formula of the form  $\bigwedge_X \Delta_h^0 \alpha$ . (Agent  $\mathbf{m}$  believes that the initial state of the game is also known by agent  $\mathbf{h}$ .)

**Initial core beliefs of agent  $\mathbf{m}$ .** Core belief base includes all the agent  $\mathbf{m}$ 's beliefs that cannot change during the game. It concerns both integrity constraints (that describe rules of the game) and successor state axioms (SSAs) that describe the building of the belief base after the execution of an action. SSAs describe both what changes in the new state, and what do not (thanks to frame axioms). So, for every  $t \in \mathbb{N}$ ,  $t' \in \mathbb{N}^*$  we define the following integrity constraints (IC):

$$\Sigma_c \supseteq \left\{ \bigwedge_{x \in CARDS} \bigvee_{p \in GRID} pos_{x,p}^t, \right. \quad (\text{ICP1})$$

$$\bigwedge_{\substack{x \in CARDS \\ p, p' \in GRID: p \neq p'}} \neg(pos_{x,p}^t \wedge pos_{x,p'}^t), \quad (\text{ICP2})$$

$$\bigwedge_{\substack{p \in GRID \\ x, x' \in CARDS: x \neq x'}} \neg(pos_{x,p}^t \wedge pos_{x',p}^t), \quad (\text{ICP3})$$

$$\bigwedge_{x \in CARDS} \bigvee_{c \in COLORS} col_{x,c}^t, \quad (\text{ICC4})$$

$$\bigwedge_{\substack{x \in CARDS \\ c, c' \in COLORS: c \neq c'}} \neg(col_{x,c}^t \wedge col_{x,c'}^t), \quad (\text{ICC5})$$

$$\bigwedge_{c \in COLORS} \bigvee_{X \in CARDS^4} \bigwedge_{x \in X} col_{x,c}^t, \quad (\text{ICC6})$$

$$\bigwedge_{\substack{h \in HINTS \\ x, x' \in CARDS \\ x \neq x'}} \neg(mark_{x,h}^t \wedge mark_{x',h}^t), \quad (\text{ICH7})$$

$$\bigwedge_{h \in AvHINTS} (avail_h^t \vee active_h^t \vee \bigvee_{x \in CARDS} mark_{x,h}^t), \quad (\text{ICH8})$$

$$\bigwedge_{h \in HINTS} \neg(avail_h^t \wedge active_h^t), \quad (\text{ICH9})$$

$$\bigwedge_{\substack{x \in CARDS \\ h \in HINTS}} \left( \neg(active_h^t \wedge mark_{x,h}^t) \wedge \neg(avail_h^t \wedge mark_{x,h}^t) \right) \quad (\text{ICH10})$$

ICP1 and ICP2 mean that each card has at least one, and at most one, position respectively ; ICP3 means that each position can only accommodate one card. In the same way, ICC4 and ICC5 means that each card has at least one, and at most one, color respectively. ICC6 means (together with ICC5) that there are 4 cards of each color. ICH7 means a hint cannot marks two different cards. ICH8 means each hint must be either available, or activated, or marking a card. ICH9 and ICH10 mean that the state of a hint is unique: it is either available, activated, or marking.

Frame axioms (FA) describe the facts that does not change after the execution of an action. For convenience, we define the following FA abbreviations (for every  $C \subseteq \text{COLORS}$ ,  $X \subseteq \text{CARDS}$ ,  $H \subseteq \text{HINTS}$ ,  $P \subseteq \text{GRID}$ , and  $t \in \mathbb{N}^*$ ):

$$\begin{aligned} \text{posFA}_X^t &\stackrel{\text{def}}{=} \bigwedge_{\substack{x \in \text{CARDS} \setminus X \\ p \in \text{GRID}}} (\text{pos}_{x,p}^t \leftrightarrow \text{pos}_{x,p}^{t-1}) \\ \text{colFA}_X^t &\stackrel{\text{def}}{=} \bigwedge_{\substack{x \in \text{CARDS} \setminus X \\ c \in \text{COLORS}}} \left( (\text{col}_{x,c}^t \leftrightarrow \text{col}_{x,c}^{t-1}) \wedge (\neg \Delta_m \text{col}_{x,c}^t \leftrightarrow \neg \Delta_m \text{col}_{x,c}^{t-1}) \right) \\ \text{hintFA}_H^t &\stackrel{\text{def}}{=} \bigwedge_{h \in \text{HINTS} \setminus H} \left( (\text{avail}_h^t \leftrightarrow \text{avail}_h^{t-1}) \wedge (\text{active}_h^t \leftrightarrow \text{active}_h^{t-1}) \right) \wedge \\ &\quad \bigwedge_{\substack{x \in \text{CARDS} \\ h \in \text{HINTS} \setminus H}} (\text{mark}_{x,h}^t \leftrightarrow \text{mark}_{x,h}^{t-1}) \end{aligned}$$

$\text{posFA}_X^t P$  (resp.  $\text{colFA}_X^t C$ ) reads “the position (resp. color) of every cards except those in  $X$  is preserved from time  $t - 1$  to  $t$ ”.  $\text{hintFA}_X^t H$  reads “the status (available, active, or marking) of every hints except those in  $H$  is preserved from time  $t - 1$  to  $t$ ”.

Finally, we define the following successor state axioms (for every  $t \in \mathbb{N}$ ,  $t' \in \mathbb{N}^*$ ):

$$\begin{aligned} \Sigma_c \supseteq &\left\{ \begin{aligned} &\bigwedge_{\substack{p, p' \in \text{GRID} \\ p \neq p'}} (\text{neighbors}_{p,p'}^t \leftrightarrow \text{neighbors}_{p,p'}^{t-1}), & \text{(SSA11)} \\ &\bigwedge_{\substack{x \in \text{CARDS} \\ p \in \text{GRID}}} (\text{pos}_{x,p}^{t'} \wedge \neg \text{pos}_{x,p}^{t'-1} \rightarrow \text{posFA}_{\{x\}}^{t'} \wedge \text{colFA}_{\emptyset}^{t'} \wedge \text{hintFA}_{\emptyset}^{t'}), & \text{(SSA12)} \\ &\bigwedge_{\substack{x \in \text{CARDS} \\ c \in \text{COLORS}}} (\text{col}_{x,c}^{t'} \wedge \neg \Delta_m \text{col}_{x,c}^{t'-1} \rightarrow \text{posFA}_{\emptyset}^{t'} \wedge \text{colFA}_{\{x\}}^{t'} \wedge \text{hintFA}_{\emptyset}^{t'}), & \text{(SSA13)} \\ &\bigwedge_{h \in \text{HINTS}} (\text{active}_h^{t'} \wedge \neg \text{active}_h^{t'-1} \rightarrow \text{posFA}_{\emptyset}^{t'} \wedge \text{colFA}_{\emptyset}^{t'} \wedge \text{hintFA}_{\{h\}}^{t'}), & \text{(SSA14)} \\ &\bigwedge_{\substack{x \in \text{CARDS} \\ h \in \text{HINTS}}} (\text{mark}_{x,h}^{t'} \wedge \neg \text{mark}_{x,h}^{t'-1} \rightarrow \text{posFA}_{\emptyset}^{t'} \wedge \text{colFA}_{\emptyset}^{t'} \wedge \text{hintFA}_{\{h\}}^{t'}), & \text{(SSA15)} \end{aligned} \right\} \end{aligned}$$

SSA11 means that every neighboring positions of a position  $p$  never change. (Note it is a pure frame axiom because neighboring positions are never affected by any actions.) SSA12 means that if card  $x$  has a new position  $p$  at time  $t'$ , the position of others cards, the color of every cards, and the status of all hints, remain unchanged from  $t' - 1$  to  $t'$ . SSA13 means that if card  $x$  has a new color  $c$  at time  $t'$ , the position of every cards, the color of every other cards, and the status of all hints, remain unchanged from  $t' - 1$  to  $t'$ . SSA14 and SSA15 mean

that if a hint becomes active (resp. marks a card) between time  $t' - 1$  and  $t'$  then neither positions and colors of cards did change nor the status of other hints. Moreover, we assume that at least one action is performed at each time point. Formally, this is represented by the disjunction of antecedents of SSA12 to SSA15.

Finally, we suppose that agent  $\mathbf{m}$  believes that all the fact in its core belief base are also known by the other agent  $\mathbf{h}$ .

## 4.2 Dynamic aspects

**Vocabulary.** For convenience, we define the action set  $ACT \subseteq EVT$  of agent  $\mathbf{m}$  (for every  $t < end$ ,  $x \in CARDS$ ,  $p \in GRID$ ,  $h \in HINTS$ ):

$$\begin{aligned} +_{\mathbf{m}}^{col_{x,c}^{t+1}} &\stackrel{\text{def}}{=} +_{\mathbf{m}}^{t+1} \left( col_{x,c}^{t+1} \wedge (\Delta_{\mathbf{h}}^{t+1} \bigvee_{c' \in COLORS} \Delta_{\mathbf{m}} col_{x,c'}^{t+1}) \wedge \Delta_{\mathbf{h}}^{t+1} now^{\geq t+1} \right) \\ +_{\mathbf{m}}^{pos_{x,p}^{t+1}} &\stackrel{\text{def}}{=} +_{\mathbf{m}}^{t+1} (pos_{x,p}^{t+1} \wedge \Delta_{\mathbf{h}}^{t+1} pos_{x,p}^{t+1} \wedge \Delta_{\mathbf{h}}^{t+1} now^{\geq t+1}) \\ +_{\mathbf{m}}^{\text{actHint}_h^{t+1}} &\stackrel{\text{def}}{=} +_{\mathbf{m}}^{t+1} (active_h^{t+1} \wedge \Delta_{\mathbf{h}}^{t+1} active_h^{t+1} \wedge \Delta_{\mathbf{h}}^{t+1} now^{\geq t+1}) \\ +_{\mathbf{m}}^{\text{markHint}_{x,h}^{t+1}} &\stackrel{\text{def}}{=} +_{\mathbf{m}}^{t+1} (mark_{x,h}^{t+1} \wedge \Delta_{\mathbf{h}}^{t+1} mark_{x,h}^{t+1} \wedge \Delta_{\mathbf{h}}^{t+1} now^{\geq t+1}) \end{aligned}$$

As each action entails a new time point, each action execution entails the fact that the agent learns not only that time increases (which is taken into account directly in the semantics, see Definition 1), but also that time increases for agent  $\mathbf{h}$  (so,  $\Delta_{\mathbf{h}}^{t+1} now^{\geq t+1}$  must be added to  $\mathbf{m}$ 's belief base). Finally, each action about a fact also informs agent  $\mathbf{m}$  that  $\mathbf{h}$  believes this fact.

**The separation constraint.** When moving a card from a position  $p$  to a position  $p'$ , it is forbidden to create two separate groups of cards (see Figure 1b). In other words, there must be a path between a given card  $x$  and all the other cards  $y$ , which automatically ensures that there is a path between any two different cards  $y'$  and  $y''$  of the game. So, for every  $p, p' \in GRID : p \neq p'$ ,  $t \in \mathbb{N}$ , and  $S \subseteq GRID : p, p' \notin S$ :

$$\begin{aligned} \text{linked}_{p,p',S}^t &\stackrel{\text{def}}{=} \left( \bigvee_{1 \leq i \leq |S|} path_{(p,p'),p',i}^t \right) \wedge \bigwedge_{p'' \in S \setminus NEIG(p)} \neg path_{(p,p'),p'',1}^t \\ &\quad \wedge \bigwedge_{\substack{2 \leq i \leq |S| \\ p'' \in S}} \left( path_{(p,p'),p'',i}^t \rightarrow \bigvee_{p''' \in NEIG(p'')} path_{(p,p'),p''',i-1}^t \right) \end{aligned}$$

that reads: “there is a path in  $S$  from position  $p$  to position  $p'$  iff: (1)  $p'$  is reachable in at most  $|S|$  steps, (2) no position which is not a neighbor of  $p$  in  $S$  is reachable from  $p$  in 1 step, (3) any position  $p''$  in the path is reachable from  $p$  within  $i$  steps only if there is at least a neighbor  $p'''$  of  $p''$  in  $S$  which is reachable in  $i - 1$  steps.”

Finally, for  $t \in \mathbb{N}$ ,  $p \in GRID$ ,  $x \in CARDS$ :

$$\begin{aligned} \text{emp}_p^t &\stackrel{\text{def}}{=} \bigwedge_{x \in CARDS} \neg \text{pos}_{x,p}^t \\ OPOS^t &= \{p \in GRID : \neg \text{emp}_p^t\} \\ \text{legMov}_{x,p}^t &\stackrel{\text{def}}{=} \bigvee_{\substack{p'' \in GRID \\ p'' \neq p}} \left( \text{pos}_{x,p''}^t \wedge \text{emp}_p^t \wedge \bigwedge_{p' \in NEIG(p'')} \text{linked}_{p,p',OPOS^t \setminus \{p''\}}^t \right) \end{aligned}$$

$\text{emp}_p^t$  reads “the position  $p$  is empty”;  $OPOS^t$  is the set of positions occupied by a card at time  $t$ ;  $\text{legMov}_{x,p}^t$  reads “the move of card  $x$  towards position  $p$  at time  $t$  is authorized by the rules of the game iff  $p$  is currently an empty position and there is a sequence of adjacent positions between  $p$  and any other occupied positions (except the initial position  $p''$  of card  $x$ ) through the set of currently occupied positions (excluding  $p''$ ).

**Action preconditions.** We assume now that the operators in *ACT* have the following executability preconditions, for  $t \in \mathbb{N}$  such that  $t < \text{end}$ ,  $x \in CARDS$ ,  $c \in COLORS$ ,  $h \in AvHINTS$  and  $\text{now}_m^t \stackrel{\text{def}}{=} \square_m \text{now}^{\geq t} \wedge \neg \square_m \text{now}^{\geq t+1}$ :

$$\begin{aligned} \mathcal{P}(+_m^{\text{col}_{x,c}^{t+1}}) &\stackrel{\text{def}}{=} \text{now}_m^t \wedge \square_m \bigwedge_{h \in HINTS} \neg \text{mark}_{x,h}^t \\ \mathcal{P}(+_m^{\text{pos}_{x,p}^{t+1}}) &\stackrel{\text{def}}{=} \text{now}_m^t \wedge \square_m \bigwedge_{h \in HINTS} \neg \text{mark}_{x,h}^t \wedge \square_m \text{legMov}_{x,p}^t \\ \mathcal{P}(+_m^{\text{actHint}_h^{t+1}}) &\stackrel{\text{def}}{=} \text{now}_m^t \wedge \square_m \text{avail}_h^t \\ \mathcal{P}(+_m^{\text{markHint}_{x,h}^{t+1}}) &\stackrel{\text{def}}{=} \text{now}_m^t \wedge \square_m \text{active}_h^t \wedge \square_m \bigvee_{c \in h} \text{col}_{x,c}^t \wedge \\ &\quad \square_m \bigwedge_{\substack{h' \in HINTS \\ h' \neq h}} \neg \text{mark}_{x,h'}^t \end{aligned}$$

$\mathcal{P}(+_m^{\text{col}_{x,c}^{t+1}})$  reads “agent  $m$  explicitly believes that time is currently equal to  $t$  and also explicitly believes that no hint marks card  $x$ ”;  $\mathcal{P}(+_m^{\text{pos}_{x,p}^{t+1}})$  reads “agent  $m$  explicitly believes that time is currently equal to  $t$  and also explicitly believes that no hint marks card  $x$  and that it is currently authorized, w.r.t. rules of the game, to move cards  $x$  to position  $p$ ”;  $\mathcal{P}(+_m^{\text{actHint}_h^{t+1}})$  reads “agent  $m$  explicitly believes that time is currently equal to  $t$  and also explicitly believes that hint  $h$  is currently available”;  $\mathcal{P}(+_m^{\text{markHint}_{x,h}^{t+1}})$  reads “agent  $m$  explicitly believes that time is currently equal to  $t$ , that hint  $h$  is currently active, that  $h$  includes the colors of card  $x$ , and that no other hint  $h'$  already marks card  $x$ ”.



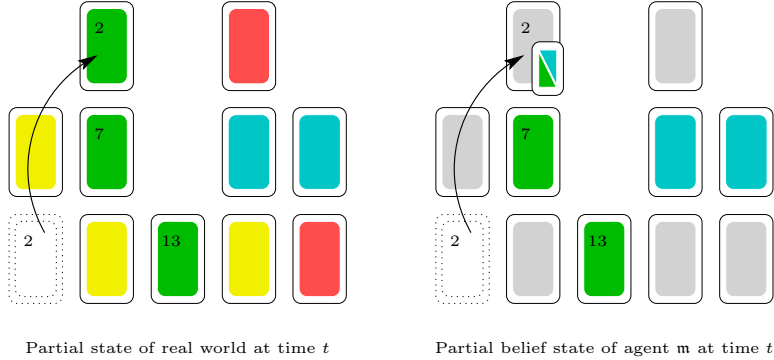


Figure 2: Example of game states

### 4.3 Planning a game round for the artificial agent

At each round of  $\mathbf{m}$  at time  $t$  we consider a planning task which consists in finding a sequence of four relevant actions from  $ACT$  to achieve a goal  $\alpha_G$ . The initial belief state of agent  $\mathbf{m}$  is given by its core beliefs  $\Sigma_c$  and an expansion of its volatile beliefs  $\Sigma_v$  after the execution of the actions of the last round of agent  $\mathbf{h}$ . The construction of the goal  $\alpha_G$  as a conjunction of subgoals is driven by a game strategy. For example, a strategy could provide a subgoal  $\alpha_{G_1}$  which is intended to maximize the number of cards with a neighbor of the same color.  $\alpha_{G_1}$  states that, from the point of view of  $\mathbf{m}$ , there is a card  $x$  whose color is known and having no neighbors of the same color, and there exists another card  $x''$  of the same color as  $x$  such that card  $x$  was moved to a previously unoccupied neighboring position at the end of the round:

$$\alpha_{G_1} = \bigvee_{\substack{x \in CARDS \\ c \in COLORS \\ p \in OPOS^t}} \left( pos_{x,p}^t \wedge col_{x,c}^t \wedge \bigwedge_{\substack{p' \in NEIG(p) \\ x' \in CARDS \setminus \{x\}}} (pos_{x',p'}^t \rightarrow \neg col_{x',c}^t \vee \neg \Delta_{\mathbf{m}} col_{x',c}^t) \wedge \bigvee_{\substack{x'' \in CARDS \\ p'' \in OPOS^t \\ p''' \in NEIG(p'') \setminus OPOS^t}} (pos_{x'',p''}^t \wedge col_{x'',c}^t \wedge pos_{x,p'''}^{t+4}) \right)$$

Figure 2 shows an example of game states at the beginning of a round of  $\mathbf{m}$ . On this example, we consider that agent  $\mathbf{m}$  knows that card 7 is green, and that  $\mathbf{h}$  knows that  $\mathbf{m}$  knows the color of this card:

$$\Sigma_v \supseteq \{ col_{7,g}^t, \Delta_{\mathbf{h}}^t \bigvee_{c \in COLORS} \Delta_{\mathbf{m}} col_{7,c}^t \}.$$

At previous round,  $\mathbf{h}$  just looked at cards 2 and 7, and then moved card 2 near card 7 and marked card 2 with the hint  $\{b, g\}$ . This leads to an expansion of

$\Sigma_v$ :

$$\Sigma_v \supseteq \left\{ \bigvee_{c \in \text{COLORS}} \Delta_{\mathfrak{h}}^t \text{col}_{2,c}^t, \bigvee_{c \in \text{COLORS}} \Delta_{\mathfrak{h}}^t \text{col}_{7,c}^t, \right. \\ \left. \text{neighbors}_{2,7}^t \wedge \neg \text{neighbors}_{2,7}^{t-1}, \text{mark}_{2,\{\text{b,g}\}}^t \wedge \neg \text{mark}_{2,\{\text{b,g}\}}^{t-1} \right\}$$

Hence, thanks to a rule in core beliefs  $\Sigma_c \supseteq \{ \text{col}_{7,g}^t \wedge \Delta_{\mathfrak{h}}^t (\bigvee_{c \in \text{COLORS}} \Delta_{\mathfrak{m}} \text{col}_{7,c}^t) \wedge (\bigvee_{c \in \text{COLORS}} \Delta_{\mathfrak{h}}^t \text{col}_{2,c}^t) \wedge (\bigvee_{c \in \text{COLORS}} \Delta_{\mathfrak{h}}^t \text{col}_{7,c}^t) \rightarrow \text{col}_{2,g}^t \}$  given by the game strategy,  $\mathfrak{m}$  implicitly deduces that card 2 is green:  $\Sigma_v \models \Box_{\mathfrak{m}} \text{col}_{2,g}^t$ . Then, subgoal  $\alpha_{G_1}$  leads agent  $\mathfrak{m}$  to move card 13 near card 7 or card 2 ( $+_{\mathfrak{m}}^{\text{pos}_{13,p}^{t+1}}$  where  $p \in \text{NEIG}(7)$  or  $p \in \text{NEIG}(2)$ ) which are better choices than moving card 7 near card 13 avoided by the latter implicit belief.

## 5 Conclusion

We have introduced a simple epistemic language for representing an artificial player’s knowledge and actions in the context of the cooperative board-game Yōkai. We have shown that this game requires a combination of Theory of Mind (ToM), temporal and spatial reasoning to be played effectively by the artificial agent. Our approach relies on SAT given the existence of a polysize satisfiability preserving translation of the epistemic language into propositional logic.

Future work will be organized in two steps. First, we intend to specify a belief revision module for the artificial player which spells out how it should change its beliefs after the human player’s moves. Secondly, we plan to implement an artificial player based on our formalization and to use existing SAT solvers for automating the reasoning and planning for the artificial player during the game play. We plan to formalize and implement a variety of game strategies for the artificial player, in line with the methodology sketched in Section 4.3, and to test their performances experimentally.

## References

- [1] G. Aucher and F. Schwarzentruber. On the complexity of dynamic epistemic logic. In *Proceedings of the 14th Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2013)*, 2013.
- [2] Guillaume Aucher. Private announcement and belief expansion: an internal perspective. *Journal of Logic and Computation*, 22(3):451–479, 2012.
- [3] A. Baltag, L. Moss, and S. Solecki. The logic of public announcements, common knowledge and private suspicions. In Itzhak Gilboa, editor, *Proceedings of the Seventh Conference on Theoretical Aspects of Rationality and Knowledge (TARK’98)*, pages 43–56, San Francisco, CA, 1998. Morgan Kaufmann.

- [4] N. Bard, J. N. Foerster, S. Chandar, N. Burch, M. Lanctot, H. F. Song, E. Parisotto, V. Dumoulin, S. Moitra, E. Hughes, I. Dunning, S. Mourad, H. Larochelle, M. G. Bellemare, and M. Bowling. The hanabi challenge: A new frontier for AI research. *Artificial Intelligence*, 280, 2020.
- [5] T. Bolander. Seeing is believing: Formalising false-belief tasks in dynamic epistemic logic. In A. Herzig and E. Lorini, editors, *Proceedings of the European conference on Social Intelligence (ECSI-2014)*, pages 87–107, 2014.
- [6] T. Bolander and M. B. Andersen. Epistemic planning for single- and multi-agent systems. *Journal of Applied Non-Classical Logics*, 21(1):656–680, 2011.
- [7] T. Bolander, H. van Ditmarsch, A. Herzig, E. Lorini, P. Pardo, and F. Schwarzentruher. Announcements to attentive agents. *Journal of Logic, Language and Information*, 25(1):1–35, 2015.
- [8] T. Caridroit, J.-M. Lagniez, D. Le Berre, T. de Lima, and V. Montmirail. A sat-based approach for solving the modal logic s5-satisfiability problem. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*, pages 3864–3870. AAAI Press, 2017.
- [9] L. Dissing and T. Bolander. Implementing theory of mind on a robot using dynamic epistemic logic. In C. Bessiere, editor, *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, (IJCAI 2020)*, pages 1615–1621, 2020.
- [10] M. Eger and C. Martens. Practical specification of belief manipulation in games. In B. Magerko and J. P. Rowe, editors, *Proceedings of the Thirteenth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE-17)*, pages 30–36. AAAI Press, 2017.
- [11] M. Eger, C. Martens, and M. A. Cordoba. An intentional AI for hanabi. In *IEEE Conference on Computational Intelligence and Games, CIG 2017*, pages 68–75. IEEE, 2017.
- [12] R. Fagin, J. Halpern, Y. Moses, and M. Vardi. *Reasoning about Knowledge*. MIT Press, Cambridge, 1995.
- [13] J. Gerbrandy and W. Groeneveld. Reasoning about information change. *Journal of Logic, Language, and Information*, 6:147–196, 1997.
- [14] J. Y. Halpern and Y. Moses. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence*, 54(2):319–379, 1992.
- [15] E. Lorini. In praise of belief bases: Doing epistemic logic without possible worlds. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)*, pages 1915–1922. AAAI Press, 2018.

- [16] E. Lorini. Rethinking epistemic logic with belief bases. *Artificial Intelligence*, 282, 2020.
- [17] E. Lorini and F. Romero. Decision procedures for epistemic logic exploiting belief bases. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019)*, pages 944–952. IFAAMAS, 2019.
- [18] C. Lutz. Complexity and succinctness of public announcement logic. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2006)*, pages 137–143. ACM, 2006.
- [19] J. A. Plaza. Logics of public communications. In M. Emrich, M. Pfeifer, M. Hadzikadic, and Z. Ras, editors, *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*, 201-216, 1989.
- [20] Y. Shoham. Logical theories of intention and the database perspective. *Journal of Philosophical Logic*, 38(6):633–648, 2009.
- [21] H. P. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic Epistemic Logic*. Kluwer Academic Publishers, 2007.

## A Detailed proof of Theorem 1

In this section, we are going to provide a polysize reduction of the satisfiability problem of  $\mathcal{L}$  to SAT. The reduction consists of three steps.

As a first step, we put  $\mathcal{L}$  formulas in negation normal form (NNF) via the following function  $nnf$ :

$$\begin{aligned}
nnf(p^t) &= p^t, \\
nnf(now^{\geq t}) &= now^{\geq t}, \\
nnf(\Delta_h^t \alpha) &= \Delta_h^t \alpha, \\
nnf(\Delta_m \alpha) &= \Delta_m \alpha, \\
nnf(\Box_m \alpha) &= \Box_m nnf(\alpha), \\
nnf(\Diamond_m \alpha) &= \Diamond_m nnf(\alpha), \\
nnf(\varphi \wedge \psi) &= nnf(\varphi) \wedge nnf(\psi), \\
nnf(\varphi \vee \psi) &= nnf(\varphi) \vee nnf(\psi), \\
nnf(\neg p) &= \neg p, \\
nnf(\neg now^{\geq t}) &= \neg now^{\geq t}, \\
nnf(\neg \Delta_h^t \alpha) &= \neg \Delta_h^t \alpha, \\
nnf(\neg \Delta_m \alpha) &= \neg \Delta_m \alpha, \\
nnf(\neg \neg \varphi) &= nnf(\varphi), \\
nnf(\neg(\varphi \wedge \psi)) &= nnf(\neg \varphi \vee \neg \psi), \\
nnf(\neg(\varphi \vee \psi)) &= nnf(\neg \varphi \wedge \neg \psi), \\
nnf(\neg \Box_m \alpha) &= \Diamond_m nnf(\neg \alpha), \\
nnf(\neg \Diamond_m \alpha) &= \Box_m nnf(\neg \alpha).
\end{aligned}$$

Let us define the NNF variant  $\mathcal{L}_0^{NNF}$  of the language  $\mathcal{L}_0$  by the following grammar:

$$\beta ::= p^t \mid \neg p^t \mid now^{\geq t} \mid \neg now^{\geq t} \mid \Delta_h^t \alpha \mid \neg \Delta_h^t \alpha \mid \Delta_m \alpha \mid \neg \Delta_m \alpha \mid \beta_1 \wedge \beta_2 \mid \beta_1 \vee \beta_2,$$

where  $p^t$  ranges over  $ATM$ ,  $t$  ranges over  $\mathbb{N}$  and  $\alpha$  ranges over  $\mathcal{L}_0$ .

Furthermore, let us define the language  $\mathcal{L}^{NNF}$  by the following grammar. For  $\beta$  ranging over  $\mathcal{L}_0^{NNF}$ :

$$\varphi ::= \beta \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \Box_m \beta \mid \Diamond_m \beta.$$

**Proposition 2** *Let  $\varphi \in \mathcal{L}$ . Then,  $\varphi \leftrightarrow nnf(\varphi)$  is valid in the class  $\mathbf{M}$ , and  $nnf(\varphi) \in \mathcal{L}^{NNF}$ .*

Note that the size of  $nnf(\varphi)$  is polynomial in the size of  $\varphi$ .

As a second step, we define the following modal language  $\mathcal{L}^{Mod}$  into which the language  $\mathcal{L}^{NNF}$  will be translated:

$$\begin{aligned}\omega & ::= q \mid \neg\omega \mid \omega_1 \wedge \omega_2 \mid \omega_1 \vee \omega_2, \\ \varphi & ::= q \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \blacksquare\omega \mid \blacklozenge\omega,\end{aligned}$$

where  $q$  ranges over the following set of atomic formulas:

$$\begin{aligned}ATM^+ = & ATM \cup \{p_{now \geq t} : t \in \mathbb{N}\} \cup \{p_{\Delta_m \alpha} : \alpha \in \mathcal{L}_0\} \cup \\ & \{p_{\Delta_b^t \alpha} : t \in \mathbb{N} \text{ and } \alpha \in \mathcal{L}_0\}.\end{aligned}$$

So  $p_{now \geq t}$ ,  $p_{\Delta_m \alpha}$  and  $p_{\Delta_b^t \alpha}$  are nothing but special propositional variables.

We interpret the language  $\mathcal{L}^{Mod}$  w.r.t. a pair  $(M, w)$ , called pointed Kripke model, where  $M = (W, \Rightarrow, \pi)$ ,  $W$  is a set of worlds,  $\Rightarrow \subseteq W \times W$  and  $\pi : ATM^+ \rightarrow 2^W$ . (Boolean cases are again omitted as they are defined in the usual way.)

**Definition 7** *The semantic interpretation for formulas in  $\mathcal{L}^{Mod}$  w.r.t. a pointed Kripke model  $(M, w)$  is as follows:*

$$\begin{aligned}(M, w) \models q & \iff w \in \pi(q); \\ (M, w) \models \blacksquare\omega & \iff \forall v \in W, \text{ if } w \Rightarrow v \text{ then } (M, v) \models \omega; \\ (M, w) \models \blacklozenge\omega & \iff \exists v \in W \text{ s.t. } w \Rightarrow v \text{ and } (M, v) \models \omega.\end{aligned}$$

The class of pointed Kripke models is denoted by  $\mathbf{K}$ . Satisfiability and validity of formulas in  $\mathcal{L}^{Mod}$  relative to the class  $\mathbf{K}$  is defined in the usual way.

Let  $tr_1 : \mathcal{L}^{NNF} \rightarrow \mathcal{L}^{Mod}$  be a translation such that:

$$\begin{aligned}tr_1(p^t) & = p^t, \\ tr_1(\neg p^t) & = \neg p^t, \\ tr_1(now^{\geq t}) & = p_{now \geq t}, \\ tr_1(\neg now^{\geq t}) & = \neg p_{now \geq t}, \\ tr_1(\Delta_b^t \alpha) & = p_{\Delta_b^t \alpha}, \\ tr_1(\neg \Delta_b^t \alpha) & = \neg p_{\Delta_b^t \alpha}, \\ tr_1(\varphi_1 \wedge \varphi_2) & = tr_1(\varphi_1) \wedge tr_1(\varphi_2), \\ tr_1(\varphi_1 \vee \varphi_2) & = tr_1(\varphi_1) \vee tr_1(\varphi_2), \\ tr_1(\Delta_m \alpha) & = p_{\Delta_m \alpha} \wedge \blacksquare tr_1(\alpha), \\ tr_1(\neg \Delta_m \alpha) & = \neg p_{\Delta_m \alpha}, \\ tr_1(\square_m \beta) & = \blacksquare tr_1(\beta), \\ tr_1(\lozenge_m \beta) & = \blacklozenge tr_1(\beta).\end{aligned}$$

As the following theorem indicates, the polynomial translation  $tr_1$  guarantees the transfer of satisfiability from model class  $\mathbf{M}$  to model class  $\mathbf{K}$ .

**Theorem 3** Let  $\varphi \in \mathcal{L}^{NNF}$ . Then,  $\varphi$  is satisfiable in the class **M** if and only if  $\blacksquare(\bigwedge_{\alpha \in \Gamma_\varphi} tr_1(\alpha)) \wedge tr_1(\varphi)$  is satisfiable in the class **K**, where  $\Gamma_\varphi$  is defined as follows:

$$\begin{aligned} \Gamma_\varphi = & \{now^{\geq 0}\} \cup \{now^{\geq t} \rightarrow now^{\geq t'} : t' \leq t \text{ and } now^{\geq t}, now^{\geq t'} \in SF(\varphi)\} \cup \\ & \{now^{\geq t} \leftrightarrow \Delta_m now^{\geq t} : now^{\geq t} \in SF(\varphi)\} \cup \\ & \{\neg now^{\geq t} \leftrightarrow \neg \Delta_m now^{\geq t} : now^{\geq t} \in SF(\varphi)\}, \end{aligned}$$

and  $SF(\varphi)$  is the set of subformulas of  $\varphi$  which is inductively defined as follows:

$$\begin{aligned} SF(p^t) &= \{p^t\}, \\ SF(\neg p^t) &= \{p^t\} \cup SF(p^t), \\ SF(now^{\geq t}) &= \{now^{\geq t}\}, \\ SF(\neg now^{\geq t}) &= \{\neg now^{\geq t}\} \cup SF(now^{\geq t}), \\ SF(\Delta_b^t \alpha) &= \{\Delta_b^t \alpha\} \cup SF(\alpha), \\ SF(\neg \Delta_b^t \alpha) &= \{\neg \Delta_b^t \alpha\} \cup SF(\Delta_b^t \alpha), \\ SF(\Delta_m \alpha) &= \{\Delta_m \alpha\} \cup SF(\alpha), \\ SF(\neg \Delta_m \alpha) &= \{\neg \Delta_m \alpha\} \cup SF(\Delta_m \alpha), \\ SF(\varphi_1 \wedge \varphi_2) &= \{\varphi_1 \wedge \varphi_2\} \cup SF(\varphi_1) \cup SF(\varphi_2), \\ SF(\varphi_1 \vee \varphi_2) &= \{\varphi_1 \vee \varphi_2\} \cup SF(\varphi_1) \cup SF(\varphi_2), \\ SF(\Box_m \beta) &= \{\Box_m \beta\} \cup SF(\beta), \\ SF(\Diamond_m \beta) &= \{\Diamond_m \beta\} \cup SF(\beta). \end{aligned}$$

As a last step, we provide a polysize reduction of  $\mathcal{L}^{Mod}$ -satisfiability to SAT, where the underlying propositional logic language  $\mathcal{L}^{PL}$  is built from the following set of atomic propositions:

$$ATM^{++} = \{q_x : q \in ATM^+ \text{ and } x \in \mathbb{N}\} \cup \{r_{x,y} : x, y \in \mathbb{N}\}.$$

Let  $tr_2 : \mathcal{L}^{Mod} \times \mathbb{N} \times \mathbb{N} \rightarrow \mathcal{L}^{PL}$  be the following translation function:

$$\begin{aligned} tr_2(q, x, y) &= q_x, \\ tr_2(\neg \varphi, x, y) &= \neg tr_2(\varphi, x, y), \\ tr_2(\varphi_1 \wedge \varphi_2, x, y) &= tr_2(\varphi_1, x, y) \wedge tr_2(\varphi_2, x, y), \\ tr_2(\varphi_1 \vee \varphi_2, x, y) &= tr_2(\varphi_1, x, y) \vee tr_2(\varphi_2, x, y), \\ tr_2(\blacksquare \omega, x, y) &= \bigwedge_{0 \leq z \leq y} (r_{x,z} \rightarrow tr_2(\omega, z, y)), \\ tr_2(\blacklozenge \omega, x, y) &= \bigvee_{0 \leq z \leq y} (r_{x,z} \wedge tr_2(\omega, z, y)). \end{aligned}$$

Translation  $tr_2$  is similar to the translation of modal logic S5 into propositional logic given in [8] and, more generally, to the standard translation of modal logic into FOL in which accessibility relations are encoded by special predicates.

The size of an  $\mathcal{L}^{Mod}$  formula,  $size(\varphi)$ , is defined by:

$$\begin{aligned} size(p^t) &= 1, \\ size(\varphi_1 \wedge \varphi_2) &= size(\varphi_1) + size(\varphi_2) + 1, \\ size(\varphi_1 \vee \varphi_2) &= size(\varphi_1) + size(\varphi_2) + 1, \\ size(\neg\varphi) &= size(\varphi) + 1, \\ size(\blacksquare\omega) &= size(\blacklozenge\omega) = size(\omega) + 1. \end{aligned}$$

Note that the size of  $tr_2(\varphi, 0, size(\varphi))$  is polynomial in the size of  $\varphi$ .

**Theorem 4** *Let  $\varphi \in \mathcal{L}^{Mod}$ . Then,  $\varphi$  is satisfiable in the class  $\mathbf{K}$  if and only if  $tr_2(\varphi, 0, size(\varphi))$  is satisfiable in propositional logic.*

Note that the size of  $tr_2(\varphi, 0, size(\varphi))$  is polynomial in the size of  $\varphi$ . Therefore, Theorem 1 follows from Proposition 2, Theorem 3 and Theorem 4.

## B Detailed proof of Theorem 2

The following equivalences are valid in the class  $\mathbf{M}$ :

$$\begin{aligned} [+_{\mathbf{m}}^t\alpha]\alpha' &\leftrightarrow \begin{cases} \top, & \text{if } \alpha' = \Delta_{\mathbf{m}}\alpha \text{ or } (\alpha' = \Delta_{\mathbf{m}}now^{\geq t'} \text{ and } t' \leq t), \\ \alpha', & \text{otherwise;} \end{cases} \\ [+_{\mathbf{m}}^t\alpha]\neg\varphi &\leftrightarrow \neg[+_{\mathbf{m}}^t\alpha]\varphi; \\ [+_{\mathbf{m}}^t\alpha](\varphi_1 \wedge \varphi_2) &\leftrightarrow [+_{\mathbf{m}}^t\alpha]\varphi_1 \wedge [+_{\mathbf{m}}^t\alpha]\varphi_2; \\ [+_{\mathbf{m}}^t\alpha](\varphi_1 \vee \varphi_2) &\leftrightarrow [+_{\mathbf{m}}^t\alpha]\varphi_1 \vee [+_{\mathbf{m}}^t\alpha]\varphi_2; \\ [+_{\mathbf{m}}^t\alpha]\square_{\mathbf{m}}\alpha' &\leftrightarrow \square_{\mathbf{m}}((\alpha \wedge \bigwedge_{t' \leq t} now^{\geq t'}) \rightarrow \alpha'); \\ [+_{\mathbf{m}}^t\alpha]\blacklozenge_{\mathbf{m}}\alpha' &\leftrightarrow \blacklozenge_{\mathbf{m}}((\alpha \wedge \bigwedge_{t' \leq t} now^{\geq t'}) \wedge \alpha'). \end{aligned}$$

Thanks to these equivalences we can define the following reduction *red* trans-



forming every  $\mathcal{L}^+$  formula  $\varphi$  into an equivalent  $\mathcal{L}$  formula  $red(\varphi)$ :

$$\begin{aligned}
red(p^t) &= p^t, \\
red(\Delta_b^t \alpha) &= \Delta_b^t \alpha, \\
red(\Delta_m \alpha) &= \Delta_m \alpha, \\
red(\neg \varphi) &= \neg red(\varphi), \\
red(\varphi_1 \wedge \varphi_2) &= red(\varphi_1) \wedge red(\varphi_2), \\
red(\varphi_1 \vee \varphi_2) &= red(\varphi_1) \vee red(\varphi_2), \\
red(\Box_m \varphi) &= \Box_m red(\varphi), \\
red(\Diamond_m \varphi) &= \Diamond_m red(\varphi), \\
red([+^t_m \alpha] \alpha') &= \begin{cases} \top, & \text{if } \alpha' = \Delta_m \alpha \text{ or } (\alpha' = \Delta_m now^{\geq t'} \text{ and } t' \leq t), \\ red(\alpha'), & \text{otherwise;} \end{cases} \\
red([+^t_m \alpha] \neg \varphi) &= red(\neg [+^t_m \alpha] \varphi), \\
red([+^t_m \alpha] (\varphi_1 \wedge \varphi_2)) &= red([+^t_m \alpha] \varphi_1 \wedge [+^t_m \alpha] \varphi_2), \\
red([+^t_m \alpha] (\varphi_1 \vee \varphi_2)) &= red([+^t_m \alpha] \varphi_1 \vee [+^t_m \alpha] \varphi_2), \\
red([+^t_m \alpha] \Box_m \alpha') &= red\left(\Box_m \left((\alpha \wedge \bigwedge_{t' \leq t} now^{\geq t'}) \rightarrow \alpha'\right)\right); \\
red([+^t_m \alpha] \Diamond_m \alpha') &= red\left(\Diamond_m \left((\alpha \wedge \bigwedge_{t' \leq t} now^{\geq t'}) \wedge \alpha'\right)\right).
\end{aligned}$$

**Proposition 3** *Let  $\varphi \in \mathcal{L}^+$ . Then,  $\varphi \leftrightarrow red(\varphi)$  is valid in the class  $\mathbf{M}$ , and  $red(\varphi) \in \mathcal{L}$ .*

Theorem 2 is a consequence of Theorem 1, Proposition 3 and the fact that the size of  $red(\varphi)$  is polynomial in the size of  $\varphi$ .