

# Boolean Games with Epistemic Goals

Thomas Ågotnes<sup>1</sup>, Paul Harrenstein<sup>2</sup>,  
Wiebe van der Hoek<sup>3</sup>, and Michael Wooldridge<sup>2</sup>

<sup>1</sup> University of Bergen, Norway, thomas.agotnes@infomedia.uib.no

<sup>2</sup> University of Oxford, UK, {mjlw,paulh}@cs.ox.ac.uk

<sup>3</sup> University of Liverpool, UK, wiebe@csc.liv.ac.uk

**Abstract.** We introduce and formally study games in which the goals of players relate to the epistemic states of players in the game. For example, one player might have a goal that another player knows a certain proposition, while another player might have as a goal that a certain player does not know some proposition. The formal model we use to study epistemic games is a variation of the increasingly popular *Boolean games* model in which each player controls a number of Boolean variables, but has limited ability to see the truth values of the overall set of formulae that hold in the game. Each player in an *epistemic Boolean game* has a goal, defined as a formula of modal epistemic logic. Using such a language for goals allows us to explicitly and compactly represent desirable epistemic states. After motivating and formally defining epistemic Boolean games as a concise representation of epistemic Kripke structures, we investigate their complexity and study their properties.

## 1 Introduction

In our everyday lives, we all quite naturally have goals and aspirations that relate to the epistemic states (knowledge and belief) of other agents. You want your children to know that you love them; you want your boss to know you work hard; the politician wants you to know he is honest; and so on. The formal analysis of such epistemic states is of course a well-established research topic in artificial intelligence, with modal logic and Kripke semantics being the pre-eminent tools of choice in such work [1]. Our aim in this paper is to begin to extend this research to the *game theoretic* aspects of systems in which the motivations of players relate to the epistemic states of others. Doing so, the following question seems very relevant: *If all players act rationally to bring about their goals, then what epistemic states will result in equilibrium?* We will refer to games in which the goals and preferences of players relate to the epistemic states of other players as epistemic games. In the present paper, we shall restrict our attention to knowledge, leaving the study of belief to future work.

As the basis for our study, we adapt the increasingly popular game theoretic model of *Boolean games* [2–5]. More precisely, we introduce *epistemic Boolean games* (EBGs). In such a game, (as in regular Boolean games), each player  $i$  is associated with a set of Boolean variables  $\Phi_i$ , which are under his control in the

sense that he can assign Boolean values to the variables  $\Phi_i$  in any way that he chooses. That is, the strategies available to a player  $i$  in an EBG are the possible Boolean assignments that can be made to the variables  $\Phi_i$ . The outcome of a Boolean game is a valuation for the overall set of Boolean variables  $\Phi$ , which will be composed of the individual assignments made by the players  $i$  in the game to their variables  $\Phi_i$ . A player  $i$  is not assumed to have perfect information of the game. This we formally capture by *visibility sets*  $\Theta_i$ , that is, sets of propositional formulae. If a formula  $\varphi$  is in  $\Theta_i$ —in which case we say that “ $i$  can see  $\varphi$ ”—it means that player  $i$  can distinguish states in which  $\varphi$  holds from ones in which this is not the case. A limit case is where every player can see every atomic formula, in which case the game is one of perfect information.

As in regular Boolean games, each player  $i$  is assumed to have a goal that he desires to be achieved. In conventional Boolean games, player  $i$ 's goal  $\gamma_i$  is represented as a formula of propositional logic. In our present work, however, the goal is assumed to be represented as a formula of modal multi-agent epistemic logic [1]. Thus, player  $i$  might have the goal that another player  $j$  comes to know something ( $\gamma_i = K_j\varphi$ ), while player  $j$  might have a goal that another player  $k$  does not know something ( $\gamma_j = \neg K_k\psi$ ). As in conventional Boolean games, the ability of a player  $i$  to influence whether his goal is achieved lies within the variables  $\Phi_i$  under his control; but our EBGs bring a new twist to this story, since the visibility sets of each player will have a part in determining whether a player's goal is achieved. The underlying EBG model we use (with sets of controlled variables and visibility sets) derives in part from the work of van der Hoek *et al.* on epistemic logics of propositional control [6], which in turn derives from logics of propositional control [7].

The remainder of this paper is structured as follows. First, in the following section we introduce EBGs. We then formalise Nash equilibria for our games, and investigate the complexity of decision problems relating to Nash equilibria. Sections 4 and 5 concern the conciseness of the representation of the strategic and epistemic situations that EBGs provide with respect to Kripke structures and regular Boolean games, respectively. We conclude in section 6 with a brief discussion of related work and some issues for future research.

## 2 Epistemic Boolean Games

We adapt the basic model of Boolean games (see, e.g., [2–5]) to model partial information.

*Epistemic Logic* Let  $\mathbb{B} = \{\top, \perp\}$  be the set of Boolean truth values, with “ $\top$ ” being truth and “ $\perp$ ” being falsity. Throughout the paper, we will use  $\Phi$  to denote a fixed, finite, non-empty set of Boolean variables, with typical members  $p, q, \dots$  etc. A *valuation* is a total function  $v : \Phi \rightarrow \mathbb{B}$ , assigning truth or falsity to every Boolean variable. Let  $\mathcal{V}$  denote the set of all valuations (over  $\Phi$ ). In the interests of brevity, we sometimes use a binary representation of valuations; for example, if  $\Phi = \{p, q\}$ , then the valuation 01 would be the one making  $p$  false and  $q$  true, assuming a natural order on the propositional variables.

We make use of a *multi-agent epistemic modal logic*, which we will refer to as  $\mathcal{EL}$  [1]. The language of  $\mathcal{EL}$  is that of the well-known multi-modal logic  $S5_n$ . The language  $\mathcal{EL}$  extends classical propositional logic with a collection of indexed unary modal operators  $K_i$ , where the intended interpretation of a formula  $K_i\varphi$  is that “agent  $i$  knows  $\varphi$ .” Formally, given a set  $\Phi$  of propositional variables the syntax of  $\mathcal{EL}(\Phi)$  is given by the following grammar:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid K_i\varphi,$$

where  $p$  is a Boolean variable in  $\Phi$  and  $i$  is an agent. We assume that the remaining classical connectives—“ $\wedge$ ” (conjunction), “ $\rightarrow$ ” (material implication), “ $\leftrightarrow$ ” (material bi-implication)—are defined in the standard way. The propositional fragment of  $\mathcal{EL}$  over  $\Phi$ —i.e., the set of formulae without occurrences of the epistemic operators  $K_i$ —we denote by  $\mathcal{L}(\Phi)$ .

The semantics for the epistemic logic  $S5_n$ , as formulated in the language  $\mathcal{EL}(\Phi)$ , is defined with respect to tuples  $K = (W, R_1, \dots, R_n, \pi)$ , also referred to as  $S5_n$  Kripke structures. Here,  $W$  is a non-empty set of (*possible*) *worlds* and, for each agent  $i$ ,  $R_i \subseteq W \times W$  is an equivalence relation over  $W$ . Finally,  $\pi : W \times \Phi \rightarrow \mathbb{B}$  is a *valuation* function, indicating the truth value of every Boolean variable in every world. For the purposes of this paper we assume  $W$  to be finite. The formulae of  $\mathcal{EL}(\Phi)$  are interpreted with respect to *pointed structures*, i.e., pairs of the form  $(K, w)$  as follows [1, pp.18–19]: An atom  $p$  is true in  $(K, w)$  iff  $\pi(w, p) = \top$ ; the clauses for negation and disjunction are standard, and for knowledge, we have

$$(K, w) \models_{\mathcal{K}} K_i\varphi \quad \text{iff} \quad \text{for all worlds } w' \text{ with } R_i(w, w'): (K, w') \models_{\mathcal{K}} \varphi.$$

*Epistemic Boolean Games* The standard framework of Boolean games involves a set of agents or players, each of which has a goal formulated as a formula  $\gamma_i$  of classical propositional logic and a set of Boolean variables she controls, in the sense that he has the unique ability to set their value. Each player strives to satisfy her goal by appropriately setting the values of the variables she controls. Every such setting can be analysed as a strategic game in which the strategies of the players are given by the ways they can assign values to the variables they control. A profile of strategies then determines a unique truth value assignment or valuation to all propositional variables. These are taken as the outcomes of the game. Moreover, each player strictly prefers states that satisfy her goal to ones that do not and is indifferent otherwise. The difficulty is that a player’s goal  $\gamma_i$  may contain variables controlled by other players  $j \neq i$ , who will also be trying to choose values for their variables in  $\Phi_j$  so as to get their goals  $\gamma_j$  satisfied, the satisfaction of which may in turn depend on the variables  $\Phi_i$ .

The setting we consider in this paper is similar, except that each player’s goal is represented as a *formula  $\gamma_i$  of the epistemic language  $\mathcal{EL}(\Phi)$* . In order to evaluate the players’ epistemic goals, each player is endowed with a *visibility set*  $\Theta_i$  consisting of formulae in the propositional fragment  $\mathcal{L}(\Phi)$  of  $\mathcal{EL}(\Phi)$ . The idea is that player  $i$  can completely and correctly perceive truth values of the formulae in  $\Theta_i$ , i.e., player  $i$  can only distinguish outcomes that differ with

respect to the truth values they assign to some of the formulae contained in  $\Theta_i$ . A formula of the form  $K_i\varphi$  then holds in a particular outcome if  $\varphi$  holds in all outcomes  $i$  cannot distinguish from that outcome.

Formally, we define an *epistemic Boolean game* for  $\mathcal{EL}(\Phi)$  (hereafter simply “game”) as a tuple

$$G = (N, \Phi, \Phi_1, \dots, \Phi_n, \gamma_1, \dots, \gamma_n, \Theta_1, \dots, \Theta_n), \text{ where:}$$

- $N = \{1, \dots, n\}$  is a set of agents (also called the *players* of the game);
- $\Phi = \{p, q, \dots\}$  is a finite set of Boolean variables;
- $\Phi_i \subseteq \Phi$  is the set of Boolean variables under the unique control of  $i \in N$ ;
- $\gamma_i$  is an  $\mathcal{EL}(\Phi)$  formula representing the goal of player  $i \in N$ ; and
- $\Theta_i \subseteq \mathcal{L}(\Phi)$  is a finite *visibility set* for player  $i \in N$  consisting of propositional formulae, with the intended interpretation that player  $i$  is able to correctly observe the truth values of the propositional formulae in  $\Theta_i$ .

As usual in Boolean games, we will require that  $\Phi_i \cap \Phi_j = \emptyset$  for  $i \neq j$ , and that  $\Phi_1 \cup \dots \cup \Phi_n = \Phi$  (i.e., the sets  $\Phi_1, \dots, \Phi_n$  form a partition of  $\Phi$ ).

A *choice* for player  $i \in N$  is a function  $v_i : \Phi_i \rightarrow \mathbb{B}$ , i.e., an allocation of truth or falsity to all the variables under  $i$ ’s control. Let  $\mathcal{V}_i$  denote the set of choices for player  $i$ . The intuitive interpretation we give to  $\mathcal{V}_i$  is that it defines the *actions* or *strategies* available to player  $i$ . An *outcome* is a collection of choices, one for each player. Formally, an outcome for a game is a tuple  $\vec{v} = (v_1, \dots, v_n) \in \mathcal{V}_1 \times \dots \times \mathcal{V}_n$ . An outcome uniquely defines an overall valuation for the variables in  $\Phi$  and for this reason we often treat outcomes for games as if they were valuations, for example, writing  $\vec{v}(p)$  to denote the value of variable  $p \in \Phi$  under the assignment corresponding to outcome  $\vec{v}$ . We will also equivocate the set of valuations and the set of outcomes writing  $\mathcal{V}$  for  $\mathcal{V}_1 \times \dots \times \mathcal{V}_n$ .

To model partial information, for every visibility set  $\Theta_i$ , we define an equivalence relation  $\sim_{\Theta_i}$  over outcomes as follows such that for all  $\vec{v}, \vec{v}' \in \mathcal{V}$ :

$$\vec{v} \sim_{\Theta_i} \vec{v}' \quad \text{iff} \quad \text{for all } \varphi \in \Theta_i: \vec{v} \models \varphi \text{ iff } \vec{v}' \models \varphi.$$

Thus,  $\vec{v} \sim_{\Theta_i} \vec{v}'$  if and only if  $\vec{v}$  and  $\vec{v}'$  agree on the truth value of the formulae in  $\Theta_i$ . In the interests of readability, where there is no possibility of confusion, we will write  $\sim_i$  instead of  $\sim_{\Theta_i}$ .

One natural constraint to consider in our setting would be that  $\Theta_i \subseteq \Phi$ , i.e., that player  $i$  can only perceive the value of atomic propositions or propositional variables. Such games we will refer to as *atomic games*. Most of the examples in this paper pertain to this setting and we will see that they have quite a specific structure. It would also be natural to require that  $\Phi_i \subseteq \Theta_i$ , i.e., a player can see his own choice with respect to the variables he controls. However, we will generally not place this as a requirement on games. Furthermore, if  $G$  is such that for every player  $i \in N$  we have  $\Phi \subseteq \Theta_i$ , we say that  $G$  is a game of *perfect information*: in a game of perfect information, every player can see every propositional variable.

We interpret formulae of  $\mathcal{EL}(\Phi)$  with respect to *pointed games*, i.e., pairs  $(G, \vec{v})$  consisting of a game  $G = (N, \Phi, \Phi_1, \dots, \Phi_n, \gamma_1, \dots, \gamma_n, \Theta_1, \dots, \Theta_n)$  and an outcome  $\vec{v}$ , as follows (again, we only give the epistemic clause)

$$(G, \vec{v}) \models_{\mathcal{EL}} K_i \varphi \quad \text{iff} \quad \text{for all valuations } \vec{v}_2 \text{ with } \vec{v} \sim_{\Theta_i} \vec{v}_2: (G, \vec{v}_2) \models_{\mathcal{EL}} \varphi.$$

Observe that this semantics depend on neither  $\Phi_1, \dots, \Phi_n$  nor  $\gamma_1, \dots, \gamma_n$ .

We assume that a player  $i$  strictly prefers all those outcomes that satisfy its goal  $\gamma_i$  over all those that do not, but is indifferent between outcomes that satisfy its goal, and is indifferent between outcomes that do not satisfy its goal. We define for each player  $i$  a *utility function*  $u_i: \mathcal{V}_1 \times \dots \times \mathcal{V}_n \rightarrow \{0, 1\}$  over outcomes representing these preferences as follows.

$$u_i(\vec{v}) = \begin{cases} 1 & \text{if } (G, \vec{v}) \models \gamma_i, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

*Example 1.* Father (player 1) and Mother player (player 2) reason about picking up their child Baby (player 3) from nursery: either father does this ( $p$ ), or mother ( $q$ ). Assume  $\Theta_1 = \Phi_1 = \{p\}$ ,  $\Theta_2 = \Phi_2 = \{q\}$ , and  $\Theta_3 = \emptyset$ . Mother wants to know that Baby is being picked up, either by father or by herself. Father wants to know that Mother knows, but also wishes to avoid both parents showing up at nursery. Baby just wants to be picked up. Formally, the goals are:

$$\begin{aligned} \gamma_1 &= K_1(K_2(p \vee q) \wedge \neg(p \wedge q)) \\ \gamma_2 &= K_2(p \vee q) \text{ and } \gamma_3 = p \vee q \end{aligned}$$

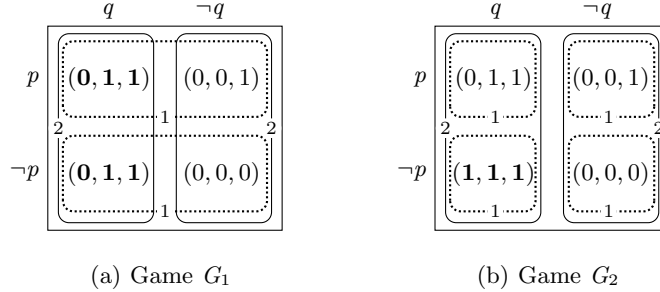
The situation (call the game  $G_1$ ) is depicted in Figure 1(a). Now suppose that Mother decides to pick up Baby, but Father does not, i.e.,  $\neg p \wedge q$ . This obviously fulfils the goal  $\gamma_2$ , and even  $K_2(p \vee q) \wedge \neg(p \wedge q)$ . Note, however, that Father is still unhappy with this outcome. Indeed, it is easy to see that  $\neg\gamma_1$  is true in this outcome. Note that each parent can bring about  $p \vee q$ , but not the fact that the spouse knows this.

We are now in a position to apply the well-known notion of (pure strategy) Nash equilibrium [8] to EBGs. Formally, an outcome  $(v_1, \dots, v_i, \dots, v_n)$  is a *Nash equilibrium* if there is no player  $i \in N$  and choice  $v'_i \in \mathcal{V}_i$  for  $i$  such that

$$u_i(v_1, \dots, v'_i, \dots, v_n) > u_i(v_1, \dots, v_i, \dots, v_n).$$

Thus, an outcome is a Nash equilibrium if no player can unilaterally deviate to obtain a better outcome for themselves, under the assumption that every other player stays with their choice. The Nash equilibria of game  $G$  will be denoted by  $\mathcal{N}(G)$ . Observe that, in general, we have that  $\vec{v}$  is a Nash equilibrium in  $G$  if and only if for all players  $i$ ,

$$(G, \vec{v}) \not\models \gamma_i \quad \text{implies} \quad (G, (v_1, \dots, v_{i-1}, v'_i, v_{i+1}, \dots, v_n)) \not\models \gamma_i \text{ for all } v'_i \in \mathcal{V}_i.$$



**Fig. 1.** The games  $G_1$  and  $G_2$  as in Example 1. The utilities to the players of each outcome are indicated by vectors  $(x, y, z)$ , where  $x$  represents father's utility,  $y$  mother's, and  $z$  baby's. The outcomes father and mother cannot distinguish are indicated by solid and dotted boxes, respectively and the Nash equilibria are in bold face.

*Example 1 (continued).* Recall the nursery Example 1. We noted that the outcome where  $\neg p \wedge q$  satisfies  $K_2(p \vee q) \wedge \neg(p \wedge q)$ . This is indeed a Nash equilibrium, even though Father's goal  $\gamma_1$  is not satisfied (Father does not *know* that Mother will go). It is a Nash equilibrium, because, given this outcome, father on his own cannot fulfil his goal: the epistemic goal  $K_1 K_2(p \vee q)$  cannot be satisfied in our model. Note that, since Baby cannot influence any proposition, every outcome has the property that it cannot unilaterally deviate and improve its outcome, i.e., it can never prevent an outcome from being a Nash equilibrium.

Note now that knowledge is power: if we change the game to  $G_2$ , in which  $\Theta_1 = \{p, q\}$  (but keep the rest as in  $G_1$ ), then Father will be informed as to whether mother satisfies her own goal by making  $q$  true, and hence  $\neg p \wedge q$  is a Nash equilibrium for this game in which both parents and the child are happy.

### 3 Computational Complexity

In this section we analyse the computational complexity of a number of natural problems relating to EBGs. First, the MODEL CHECKING problem for  $\mathcal{EL}$  is as follows (cf. [10]):

MODEL CHECKING:  
*Given:* Pointed game  $(G, \vec{v})$  and  $\mathcal{EL}$  formula  $\varphi$ .  
*Question:* Is it the case that  $(G, \vec{v}) \models_{\mathcal{EL}} \varphi$ ?

We leave out the proof of the following due to lack of space.

**Proposition 2.** *The MODEL CHECKING problem for  $\mathcal{EL}$  is PSPACE-complete.*

As an aside, we remark that this result may at first seem surprising, given that the problem of model checking formulae of  $S5_n$  over Kripke structures may be solved in polynomial time [1, pp. 63–64]. There is no contradiction,

<p><b>Input:</b> Game <math>G</math> and outcome <math>\vec{v}</math> for <math>G</math></p> <p><b>Output:</b> “yes” if <math>\vec{v} \in \mathcal{N}(G)</math>, “no” otherwise.</p> <pre> 1  for <math>i := 1</math> to <math>n</math> do 2    if <math>(G, (v_1, \dots, v_i, \dots, v_n)) \not\models_{\mathcal{EL}} \gamma_i</math> then 3      for each <math>v'_i \in \mathcal{V}_i</math> do 4        if <math>(G, (v_1, \dots, v'_i, \dots, v_n)) \models_{\mathcal{EL}} \gamma_i</math> then 5          return “no” 6        end if 7      end for 8    end if 9  end for 10 return “yes” </pre>
--

**Algorithm 1:** Algorithm for MEMBERSHIP

however: we are interpreting our  $\mathcal{EL}$  formulae with respect to games, which can be understood as *compact* representations of Kripke structures. The polynomial time computational results for model checking over Kripke structures assumes that each state is explicitly listed in the input to the problem, which we do not assume. Thus, model checking  $S5_n$  over Kripke structures assumes an input that in the worst case is exponentially larger than our game representation. As a general rule of thumb in complexity analysis, the more compact a representation is, the higher will be the complexity of the decision problems associated with these structures (see, e.g., [11, pp. 492–495]).

Two natural decision problems suggest themselves relating to the Nash equilibria of EBGs (cf. [12, pp. 8–9]). The first asks whether a given outcome  $\vec{v}$  is a Nash equilibrium of a game  $G$ . We call this problem MEMBERSHIP:

MEMBERSHIP:

*Instance:* Game  $G$  and outcome  $\vec{v}$  for  $G$ .

*Question:* Is it the case that  $\vec{v} \in \mathcal{N}(G)$ ?

We first prove the upper bound for the MEMBERSHIP problem; we postpone the lower bound for the moment.

**Proposition 3.** MEMBERSHIP for EBGs is in PSPACE.

*Proof.* First observe that from Proposition 2, the model-checking problem for  $\mathcal{EL}$  is PSPACE-complete. Let  $G$  be the game and let  $\vec{v} = (v_1, \dots, v_i, \dots, v_n)$  be the outcome given in the problem instance. Then, Algorithm 1 decides the problem in PSPACE. The loop on lines (1)–(9) checks whether any player has a beneficial deviation: if it finds such a beneficial deviation, the algorithm returns “no” (line 5), indicating  $\vec{v} \notin \mathcal{N}(G)$ . We claim that the overall algorithm operates in PSPACE. To see this, observe that for each player  $i \in N$  we carry out a single PSPACE check (line 2, Proposition 2) followed by a loop (lines 3–7) that iterates through all elements of  $\mathcal{V}_i$ , in each case carrying out a PSPACE check (Proposition 2 again). The loop on lines 3–7 is in PSPACE as a consequence of

the fact that

$$\text{PSPACE} = \text{PSPACE}^{\text{PSPACE}}.$$

Since PSPACE is closed under sequential PSPACE operations, the outer loop operates in PSPACE. We conclude that the overall algorithm operates in PSPACE.

The second problem simply asks whether there exist *any* Nash equilibria for a given EBG  $G$ . We call this problem NON-EMPTINESS:

NON-EMPTINESS:

*Instance:* Game  $G$ .

*Question:* Is it the case that  $\mathcal{N}(G) \neq \emptyset$ ?

The proofs of the following propositions are left out due to lack of space.

**Proposition 4.** NON-EMPTINESS for EBGs is PSPACE-complete.

Given this, we can prove the lower bound for the MEMBERSHIP problem:

**Proposition 5.** MEMBERSHIP for EBGs is PSPACE-hard.

## 4 Epistemic Boolean Games *versus* Kripke Structures

The semantics we have given to  $\mathcal{EL}$  in terms of epistemic Boolean games is very close to Kripke semantics, with outcomes for games essentially playing the role of possible worlds. Let us make this idea both explicit and precise. We argue that Boolean games of incomplete information can represent the same situations modelled by  $S5_n$  Kripke structures and, moreover, that they sometimes can do so significantly more concisely.

We adapt the standard definition of *bisimulation* for modal logic [13, pp.64–67] to games and Kripke structures. Given a game  $G$  for  $\mathcal{EL}(\Phi)$  and a Kripke structure  $K$  for  $\mathcal{EL}(\Phi')$ , for any  $\Psi \subseteq \Phi \cap \Phi'$ , we say that  $G$  and  $K$  are  $\Psi$ -bisimilar if there exists a relation

$$\mathcal{Z} \subseteq \mathcal{V} \times \mathcal{W}$$

such that for all outcomes  $\vec{v}$  and all possible worlds  $w$ :

- if  $\mathcal{Z}(\vec{v}, w)$ , then for all  $p \in \Psi$ , we have  $\vec{v}(p) = \pi(w, p)$ ;
- if  $\mathcal{Z}(\vec{v}, w)$  and for some  $i \in N$  and  $\vec{v}'$  we have  $\vec{v} \sim_i \vec{v}'$ , then there is some  $w'$  such that  $\mathcal{Z}(\vec{v}', w')$  and  $R_i(w, w')$ ; and
- if  $\mathcal{Z}(\vec{v}, w)$  and for some  $i \in N$  and  $w'$  we have  $R_i(w, w')$ , then there is some  $\vec{v}'$  such that  $\mathcal{Z}(\vec{v}', w')$  and  $\vec{v} \sim_i \vec{v}'$ .

We write  $(G, \vec{v}) \cong_{\Psi} (K, w)$  to mean that there is a  $\Psi$ -bisimulation relation  $\mathcal{Z}$  between  $G$  and  $K$  such that  $\mathcal{Z}(\vec{v}, w)$ . If  $\Phi = \Phi' = \Psi$  we omit the subscript in  $\cong_{\Psi}$ . The key point about bisimulations is the following, readily established result.

**Proposition 6.** Let  $\Psi \subseteq \Phi$ . If  $(G, \vec{v}) \cong_{\Psi} (K, w)$ , then for all formulae  $\varphi \in \mathcal{EL}(\Psi)$  we have:

$$(G, \vec{v}) \models_{\mathcal{EL}} \varphi \quad \text{iff} \quad (K, w) \models_{\mathcal{K}} \varphi.$$



We easily show that for every game  $G$  there exists a bisimilar Kripke structure  $K_G$ . Given an EBG  $G = (N, \Phi, \Phi_1, \dots, \Phi_n, \gamma_1, \dots, \gamma_n, \Theta_1, \dots, \Theta_n)$ , we define the Kripke structure  $K^G = (W^G, R_1^G, \dots, R_n^G, \pi^G)$  induced by  $G$  as follows.

- $W^G = \mathcal{V}$ , i.e.,  $W^G$  is the set of outcomes for  $G$ ;
- $R_i^G(\vec{v}, \vec{v}')$  iff  $\vec{v} \sim_i \vec{v}'$ ; and
- for all  $\vec{v}$  and  $p \in \Phi$ , we have  $\pi^G(\vec{v}, p) = \vec{v}(p)$ .

Then it is immediate by construction that we have:

**Proposition 7.** *Let  $G$  be a game,  $\vec{v}$  an outcome, and  $K^G$  be the Kripke structure induced by  $G$ . Then,  $(G, \vec{v}) \cong_{\Phi} (K^G, \vec{v})$ .*

Notice that the construction of the set of worlds in  $K^G$  makes the exponential blow up of size in moving from games  $G$  to Kripke structures explicit. The previous result essentially tells us that for every game  $G$  there is an “equivalent” Kripke structure  $K$ , where equivalence is measured in terms of bisimulation.

Likewise, for every finite Kripke structure  $K = (W, R_1, \dots, R_n, \pi)$  for  $\mathcal{EL}(\Phi)$ , a  $\Phi$ -bisimilar EBG  $G^K$  exists. This EBG, however, is defined on an extended language, which also involves propositional variables for every possible world  $w \in W$ . Assuming without loss of generality that  $W \cap \Phi = \emptyset$ , we let  $\Psi = \Phi \cup W$ . Moreover, define for every  $w \in W$  a formula  $\chi(w)$  as follows.

$$\chi(w) = w \wedge \bigwedge_{w' \neq w} \neg w' \wedge \bigwedge_{\pi(w,p)=\top} p \wedge \bigwedge_{\pi(w,p)=\perp} \neg p$$

Let for each possible world  $w$  the outcome  $\vec{v}_w : \Psi \rightarrow \mathbb{B}$  such that for all  $p \in \Psi$ ,

$$\vec{v}_w(p) = \begin{cases} \top & \text{if either } \pi(w, p) = \top \text{ or } p = w, \\ \perp & \text{otherwise.} \end{cases}$$

Note that the formula  $\chi(w)$  characterises outcome  $\vec{v}_w$ , in the sense that

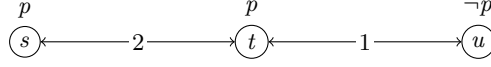
$$(G, \vec{v}_w) \models \chi(w') \quad \text{if and only if} \quad w = w'.$$

Given  $K = (W, R_1, \dots, R_n, \pi)$ , we define for each agent  $i$  and each world  $w$ ,

$$\theta_i^K(w) = \bigvee_{R_i(w, w')} \chi(w').$$

Intuitively,  $\theta_i^K(w)$  characterises (in  $\mathcal{L}(\Psi)$ ) the set of worlds that player  $i$  cannot distinguish from  $w$  in  $K$  along with their interpretation of the propositional variables. That is, if  $\theta_i^K(w)$  is in player  $i$ 's visibility set,  $i$  can distinguish worlds that are in  $\{w' \in W : R_i(w, w')\}$  from those that are not. Moreover, as  $K$  is an  $S5_n$  model, each  $R_i$  is an equivalence relation. Hence,  $R_i(w, w')$  implies  $\theta_i^K(w) = \theta_i^K(w')$ . The game  $G^K = (N, \Phi^K, \Phi_1^K, \dots, \Phi_n^K, \gamma_1^K, \dots, \gamma_n^K, \Theta_1^K, \dots, \Theta_n^K)$  induced by  $K$  is now defined such that  $\Phi^K = \Phi \cup W$  and for each agent  $i$ ,

$$\Theta_i^K = \{\theta_i^K : w \in W\}.$$



**Fig. 2.** S5 Kripke structure with no bisimilar atomic game models.

The choices for  $\Phi_1^K, \dots, \Phi_n^K$  and  $\gamma_1^K, \dots, \gamma_n^K$  are arbitrary. Observe, however, that the game  $G^K$  does *not* in general represent the situation much more concisely than the Kripke structure  $K$ .

Now define the relation  $\mathcal{Z}$  such that for all  $\vec{v} \in \mathcal{V}$  and all  $w \in W$ ,

$$\mathcal{Z}(\vec{v}, w) \quad \text{if and only if} \quad \vec{v} = \vec{v}_w.$$

We find that  $\mathcal{Z}$  is witness to the fact that  $G^K$  and  $K$  are  $\Phi$ -bisimilar, giving us the following result (proof left out due to lack of space).

**Proposition 8.** *Let  $K$  be a Kripke structure,  $w$  a world, and let  $G^K$  be the game induced by  $K$ . Then,  $(G^K, \vec{v}_w) \cong_{\Phi} (K, w)$ .*

However, if we restrict ourselves to atomic games, i.e., games with visibility sets  $\Theta_i$  that consist of propositional variables, Proposition 8 (or the converse of Proposition 6) does not hold.

**Proposition 9.** *There are Kripke structures  $K$  over  $N, \Phi$  for which there is no atomic game  $G$  over  $N, \Phi$  such that  $G \cong K$ .*

*Proof.* Consider the S5<sub>n</sub> Kripke structure  $K$  with  $N = \{1, 2\}$  and  $\Phi = \{p\}$  as depicted in Figure 2. (Both  $R_1$  and  $R_2$  are equivalence relations, but for clarity we have not drawn the reflexive arrows). Note that  $K_1 p \vee K_1 \neg p$  means that 1 knows the value of  $p$ . In  $K$ , this is true in  $s$ , but not in  $t$  or  $u$ . Hence we have  $K, s \models_{\mathcal{K}} K_1 p \wedge \neg K_2 (K_1 p \vee K_1 \neg p)$ . However, we claim that in any atomic game  $G$  and any valuation  $\vec{v}$  we have

$$(G, \vec{v}) \models_{\varepsilon \mathcal{L}} K_1 p \rightarrow K_2 (K_1 p \vee K_1 \neg p)$$

To see the latter, suppose that  $(G, \vec{v}) \models_{\varepsilon \mathcal{L}} K_1 p$ . This means not only that  $\vec{v}(p) = \top$ , but also that  $p \in \Theta_1$ . To appreciate the latter, observe that  $p \notin \Theta_1$  would imply that there is some  $\vec{u}$ , which only differs from  $\vec{v}$  in that it assigns  $\perp$  to  $p$  rather than  $\top$ , and, since  $\vec{v}$  and  $\vec{u}$  agree on all other variables (in particular those in  $\Theta_1$ ), for which  $\vec{v} \sim_i \vec{u}$ . But this in turn implies that  $(G, \vec{v}) \models_{\varepsilon \mathcal{L}} \neg K_i p$ , a contradiction. But since  $p \in \Theta_1$ , it is obvious that  $G \models_{\varepsilon \mathcal{L}} K_1 p \vee K_1 \neg p$ , and hence we have  $(G, \vec{v}) \models_{\varepsilon \mathcal{L}} K_2 (K_1 p \vee K_1 \neg p)$ . So there is no  $(G, \vec{v})$  that satisfies  $K_1 p \wedge \neg K_2 (K_1 p \vee K_1 \neg p)$ , and hence, by Proposition 2, there is no  $(G, \vec{v})$  for which  $(G, \vec{v}) \cong (K, s)$ .

The upshot of this discussion is that atomic games induce Kripke structures with a very particular structure. Specifically, they induce Kripke structures in which: (i) the set of worlds corresponds exactly to the set of possible valuations;

(ii) the accessibility relations  $R_i$  are derived solely from the visibility sets  $\Theta_i$  that consist of propositional variables; and (iii) which variables agents can see, which variables they can manipulate, and what goals they have, is all common knowledge. Issues like this (in the closely related model of *interpreted systems* for knowledge [1, pp.103–114]) were studied in depth by Lomuscio [14].

## 5 Epistemic Boolean Games *versus* Boolean Games

What distinguishes epistemic Boolean games from regular Boolean games is the language in which the players' goals are phrased, viz., the language of  $\mathcal{EL}$  and raw propositional logic, respectively. Otherwise, the strategic structure of epistemic Boolean games is very similar to that of regular Boolean games. Rather, in this section, we argue that for every epistemic Boolean game a corresponding regular Boolean game can be found that is in an important sense strategically equivalent to it. However, we also show that, in the case of atomic games, the regular Boolean game may be exponentially larger than the epistemic Boolean game it corresponds to.

Given a set  $\Phi$  of propositional variables the set of  $2^{|\Phi|}$  outcomes are determined by the set of valuations. As the only role of the visibility sets in EBGs is to enable the evaluation of the epistemic formulae in  $\mathcal{EL}(\Phi)$ , it can now easily be seen that for every epistemic Boolean game there is a regular Boolean game with the same *strategic* properties. (The claim in the opposite direction is trivial.)

Let  $G = (N, \Phi, \Phi_1, \dots, \Phi_n, \gamma_1, \dots, \gamma_n, \Theta_1, \dots, \Theta_n)$  be an epistemic Boolean game. Associate with each outcome  $\vec{v}$  a *propositional formula*  $\chi(\vec{v})$  in much the same way as in the previous section.

$$\chi(\vec{v}) = \bigwedge_{\vec{v}(p)=\top} p \wedge \bigwedge_{\vec{v}(p)=\perp} \neg p.$$

Now consider the regular Boolean game  $G' = (N, \Phi, \Phi_1, \dots, \Phi_n, \gamma'_1, \dots, \gamma'_n)$ , where for each player  $i$ ,

$$\gamma'_i = \bigvee_{(G, \vec{v}) \models_{\mathcal{EL}} \gamma_i} \chi(\vec{v}).$$

We then have that for players  $i$  and all outcomes  $\vec{v}$ ,

$$(G, \vec{v}) \models_{\mathcal{EL}} \gamma_i \quad \text{if and only if} \quad (G, \vec{v}) \models_{\mathcal{EL}} \gamma'_i.$$

Hence,  $G$  and  $G'$  agree on the players, strategies, and outcomes. Moreover, each goal  $\gamma_i$  induces the same preferences over the outcomes in  $G$  as  $\gamma'_i$  does in  $G'$ . Accordingly,  $G$  and  $G'$  could justifiably be said to be strategically equivalent.

It should be observed though, that the size of  $\gamma'_i$ , as defined above, is exponential in the size of  $\gamma_i$ . This raises the obvious question of whether we are in fact gaining anything by using the epistemic language: can we find an equivalence preserving translation  $\tau$  from epistemic formulae to propositional formulae such that  $\tau(\varphi)$  is guaranteed to be of size polynomial in the size of  $\varphi$ ? If the answer

was “yes”, then this would indicate that our epistemic language was somewhat redundant, in terms of raw expressive power and succinctness. In fact, there is a compelling complexity theoretic argument that the epistemic language is *exponentially more succinct* than the propositional language on our game structures:

**Proposition 10.** *If  $P \neq PSPACE$ , then the epistemic language is exponentially more succinct than the propositional language over game structures.*

*Proof.* Recall that model checking for the epistemic language over our game structures is PSPACE-complete, while the model checking problem for the propositional language is in P. Now, suppose there exists a polynomial time computable equivalence preserving translation  $\tau$  from the epistemic language to the propositional language. The existence of such a translation would imply PSPACE=P: to solve a PSPACE hard model checking problem for the epistemic language, we could apply the polynomial time translation  $\tau$  and apply the polynomial time model checking algorithm for the resulting formula, yielding a polynomial time decision procedure for the PSPACE-hard model checking problem. We conclude that no such translation can exist unless P=PSPACE.

Note that it is considered *highly* unlikely that PSPACE=P.

Thus, the epistemic language provides concrete benefits with respect to succinctness. We might also note that the use of epistemic modalities of course provides benefits with respect to naturalness of expression and the readability of formulae when compared to the use of raw propositional logic.

## 6 Related Work and Conclusions

In this paper, we formally defined and investigated games in which the goals of players relate to the epistemic states of other players. In such a game, a player will be strategically motivated to act in such a way as to bring about states of knowledge—or indeed ignorance—in other players. We formally defined epistemic Boolean games, an extension to the now well-established Boolean games model in which players have goals represented not as propositional logic formulae, but as formulae of modal epistemic logic. We then investigated the computational complexity of questions relating to Nash equilibria in such games.

Our research is closely related to several other papers that have appeared in the literature. In [15] we also considered Boolean games with visibility sets, but the setting in that paper is very different: goals are propositional rather than epistemic, and the focus is rather on identifying *verifiable* Nash equilibria; equilibria that players know to be equilibria.

Van Otterloo *et al.* introduced *knowledge condition games* [16]. Knowledge condition games are extensive form games, in which state sets are explicitly listed. The basic question considered in the work of van Otterloo *et al.* is when sets of players in extensive form games are able to act in such a way as to bring about a state of knowledge; the basic result relating to this problem is that the problem is  $\Sigma_2^P$ -complete. Because the problem relates to strategic ability, van

Otterloo also considered the use of the strategic ability logic ATL [17] and its epistemic variant ATEL [18]. The work differs from the present paper in that we consider a Boolean games framework, with visibility sets implicitly defining epistemic accessibility relations; we also allow for players to have goals explicitly represented as formulae of epistemic logic.

Ågotnes and van Ditmarsch also considered closely related issues [19]. They introduced *public announcement games*. The foundation for their study was the growing body of work on *dynamic epistemic logic* (DEL) [20]. In dynamic epistemic logic, agents are allowed to make announcements, which may be simple objective statements about the state of the world (e.g., “ $p$  is true”) or may be more complex announcements involving statements about the knowledge or ignorance of agents (e.g., “I don’t know  $p$ ”). Ågotnes and van Ditmarsch investigated how players in a game could use such announcements to bring about states of knowledge. The main difference is that the announcements considered by Ågotnes and van Ditmarsch are much richer than the mechanisms available to players in our games for bringing about epistemic states; and the semantics are correspondingly much more technically involved than in our setting. Our setting offers a much more *compact* representation for epistemic games than that of Ågotnes and van Ditmarsch, and one that is much closer to computational models. Additionally, we are able to prove results relating to (e.g.) the complexity of Nash equilibria that were not considered in [19].

Finally, and perhaps the closest to the present paper is the work of Grant *et al.* on Boolean games in which players have (possibly incorrect) beliefs about the value of certain variables, and where an external principal is able to make announcements about these variables in order to influence the beliefs of the players within the game, and hence the rational choices that they subsequently make [21]. The model of belief studied by Grant *et al* is much more restricted than the model of knowledge we use in the present paper (every player simply believes every proposition is either true or false).

Several avenues suggest themselves for future work. First, we might look at richer computational models than simply setting variables to be true or false. For example, we might model games using a practical system specification language such as REACTIVE MODULES [22], and, correspondingly, allow for richer strategies as in the Alternating-time Temporal Logic [17]. This would bring the setting of epistemic games much closer to practical systems, and would make it possible to model practical protocols and systems. Second, we might consider the use of mixed (probabilistic) strategies. If players are permitted to use randomised strategies, then the set  $\mathcal{N}(G)$  can be replaced by a probability distribution over outcomes. In this case, instead of asking whether, for example, player  $i$  knows  $\varphi$  in all outcomes, we can ask what the probability is that player  $i$  comes to know  $\varphi$ , assuming that players play Nash equilibrium strategies.

Third, the epistemic concept we have worked with in this paper is *knowledge*, in the sense of S5 modal logic [1]. It would also be interesting to consider the use of *belief*, e.g., in the sense of the modal logic KD45. Finally, it looks worthwhile to see whether the exponential blow-up when moving from games to Kripke

models, can be avoided by adopting a slightly different perspective on Kripke models, as proposed in [9], viewing them as compositions of smaller agent models. **Acknowledgements:** Harrenstein and Wooldridge were supported by the ERC under Advanced Investigator Grant 291528 (“RACE”).

## References

1. Fagin, R., Halpern, J.Y., Moses, Y., Vardi, M.Y.: Reasoning About Knowledge. The MIT Press (1995).
2. Harrenstein, P., van der Hoek, W., Meyer, J.J., Witteveen, C.: Boolean games. In van Benthem, J., ed.: Proc. TARK VIII, (2001) pp. 287–298.
3. Bonzon, E., Lagasque, M., Lang, J., Zanuttini, B.: Boolean games revisited. In: Proc. ECAI-2006.
4. Dunne, P.E., Kraus, S., van der Hoek, W., Wooldridge, M.: Cooperative boolean games. In: Proc. AAMAS-2008.
5. Endriss, U., Kraus, S., Lang, J., Wooldridge, M.: Designing incentives for boolean games. In: Proc. AAMAS-2011.
6. van der Hoek, W., Troquard, N., Wooldridge, M.: Knowledge and control. In: Proc. AAMAS-2011.
7. van der Hoek, W., Wooldridge, M.: On the logic of cooperation and propositional control. *Artificial Intelligence* **164** (2005) 81–119.
8. Osborne, M.J., Rubinstein, A.: A Course in Game Theory. MIT Press (1994).
9. van Eijck, J., Sietsma, F., Wang, Y.: Composing Models. *Journal of Applied Non-Classical Logic* **21**:3–4 (2011) 397–425.
10. Clarke, E.M., Grumberg, O., Peled, D.A.: Model Checking. MIT Press (2000).
11. Papadimitriou, C.H.: Computational Complexity. Addison-Wesley (1994).
12. Chalkiadakis, G., Elkind, E., Wooldridge, M.: Computational Aspects of Cooperative Game Theory. Morgan-Claypool (2011).
13. Blackburn, P., de Rijke, M., Venema, Y.: Modal Logic. Cambridge University Press (2001).
14. Lomuscio, A.: Knowledge Sharing among Ideal Agents. PhD thesis, School of Computer Science, University of Birmingham, Birmingham, UK (1999)
15. Ågotnes, T., Harrenstein, P., van der Hoek, W., Wooldridge, M.: Verifiable equilibria in boolean games. In: Proc. IJCAI-2013.
16. van Otterloo, S., van der Hoek, W., Wooldridge, M.: Knowledge condition games. *Journal of Logic, Language, and Information* **15** (2006) 425–452.
17. Alur, R., Henzinger, T.A., Kupferman, O.: Alternating-time temporal logic. *Journal of the ACM* **49** (2002) 672–713.
18. van der Hoek, W., Wooldridge, M.: Time, knowledge, and cooperation: Alternating-time temporal epistemic logic and its applications. *Studia Logica* **75** (2003) 125–157.
19. Ågotnes, T., van Ditmarsch, H.: What will they say? – public announcement games. *Synthese* **179** (2011) 57–85.
20. van Ditmarsch, H., van der Hoek, W., Kooi, B.: Dynamic Epistemic Logic. Springer-Verlag (2007).
21. Grant, J., Kraus, S., Wooldridge, M., Zuckerman, I.: Manipulating boolean games through communication. In: Proc. IJCAI-2011.
22. Alur, R., Henzinger, T.A.: Reactive modules. *Formal Methods in System Design* **15** (1999) 7–48.