2015 20th International Conference on Control Systems and Science

Hiding and revealing information in Boolean Games

Matei Popovici
POLITEHNICA University of Bucharest
Bucharest, Romania
matei.popovici@cs.pub.ro

Lorina Negreanu
POLITEHNICA University of Bucharest
Bucharest, Romania
lorina.negreanu@cs.pub.ro

***Abstract*—Imperfect information is essential for modelling and reasoning about MAS which describe realistic systems or interactions. In this paper, we start from Boolean Games — an already established framework for capturing strategic behaviour in MAS, and introduce a new method for describing uncertainty. We illustrate situations in which the standard approach for imperfect information is problematic, and show how our proposal can deal with such situations. We also equip agents with the ability of formulating goals which express others' knowledge on the game outcome, and discuss how this is different from existing work.**

I. Introduction

Multi-Agent Systems (MAS) are a pervasive paradigm for modelling and reasoning about complex systems. The cornerstone of a MAS is the agent — commonly perceived as an abstraction for software components of a distributed system. However, MAS are far from being just another software architectural model. Essentially, a MAS can also describe an open system, whose participants may be software pertaining to different organizations, as well as humans. Hence, two major features regarding agents come into play: *heterogeneity* and *rationality*. With these in mind, MAS become models for competitive interactions involving both man and machine, in which partly-selfish and/or partly-cooperative agents attempt to maximize utility, or to achieve some desired outcome.

Game theory is a natural setting for describing and reasoning about such interactions. Game theory traditionally focuses on *equilibria* — situations in which (groups of) agents have no incentive to bring about change — and properties of such equilibria. It is less concerned with game representation and complexity. For instance, in describing a game, one assigns *utility values* to each agent and possible outcome of the game. These values measure the *welfare* of the corresponding agent, provided that the particular game outcome was realized. However, in a k-player game where each player has n actions available, we have n^k possible outcomes, a number which grows exponentially w.r.t. the number of players. Thus, specifying utilities for each individual outcome (as in the traditional matrix-based normal-form game representation, see e.g. [1]) is not feasible in the general case.

Boolean Games [2], [3] are an attractive class of games of compact representation. In a Boolean Game, each player controls a subset of *propositional symbols* and formulates a *goal* which is expressed as a boolean formula. Game outcomes are interpretations for the former, and utility is defined w.r.t. goal satisfaction.

More recently, imperfect (or incomplete) information has been considered in the context of Boolean Games [4], [5]. In this setting, players are no longer able to observe the actions of others, and thus may be unsure about the actual game outcome. This situation accurately reflects real-life scenarios from, e.g. communication, security and e-commerce.

The standard approach (e.g. [4]) for modelling imperfect information is to assign to each player, a *visibility set*, i.e. a set of propositional symbols whose valuation cannot be observed by the respective player.

In this paper, we illustrate scenarios where visibility sets are unable to faithfully describe the incomplete information that agents possess. We propose a solution for such scenarios based on *visibility relations* E_i, a generalization of visibility sets. A member $(q_\top, p) \in E_i$ models the fact that if variable q is assigned *true* ($\top$), then the truth value of p is revealed to player i.

We also extend Propositional Logic — the goal language of Boolean Games — in order to allow players to formulate goals related to the knowledge of others, as well as to the outcome of the game.

We illustrate that our version of Boolean Games is still a compact game representation. However, our small change in modelling imperfect information requires a reformulation of game equilibria. We illustrate using examples that the traditional imperfect information Nash Equilibrium [4], leads to unrealistic equilibria.

We discuss the modifications that are necessary in order to obtain a Nash Equilibrium with desirable properties.

Finally, we conjecture that our version of Boolean Games does not come with an increase in complexity: the verification problem ("*Is a given outcome a Nash Equilibrium of a particular game?*)" and stability problem ("*Do Nash Equilibria exist, for a given game?*") are likely to retain the same computational complexity as in the traditional case [4].

The rest of the paper is structured as follows: In Section II we introduce Boolean Games with Perfect Information; in Section III we discuss imperfect information under two settings: the traditional one (similar to [4]), and ours. We illustrate our modelling approach with examples. Finally, in Section IV we discuss related work and in Section V we conclude.

II. Boolean Games — Perfect Information

A. Preliminaries

Let $\mathcal{P}rops$ designate a finite set whose elements we call *propositions* or *variables* and $\mathbf{PL}(\mathcal{P}rops)$ denote the set of propositional formulae built with propositions from $\mathcal{P}rops$.

978-1-4799-1780-8/15 $31.00 © 2015 IEEE
DOI 10.1109/CSCS.2015.112

Whenever the latter is understood from the context, we simply write **PL** instead of $\mathbf{PL}(\mathcal{P}rops)$. An *interpretation* of $\mathcal{P}rops$ is a set $\xi \subseteq \mathcal{P}rops$ where propositions $p \in \xi$ and $p \in \mathcal{P}rops \setminus \xi$ are interpreted as *true* and *false*, respectively. The entailment relation $\models \subseteq 2^{\mathcal{P}rops} \times \mathbf{PL}(\mathcal{P}rops)$ is defined in the usual way. The set $clash(\xi, \xi') = \{p \in \mathcal{P}rops : p \in \xi \cup \xi' \wedge p \notin \xi \cap \xi'\}$ contains those variables which receive different truth values in ξ and ξ'.

B. Pregames

We reformulate the original definition of Boolean games [2], [6] by separating the utility definition from the rest of the setting. This makes discussing different types of games, easier. To this end, we introduce *Boolean Pregames*. A Boolean Pregame contains all ingredients which describe player interaction, but does not specify how players compute their utility.

Definition 1 (Pregame): A *Boolean Pregame* $\mathbb{B}(\mathcal{P}rops)$ over $\mathcal{P}rops$ is a tuple:

$$\mathbb{B}(\mathcal{P}rops) = \langle N, (\Theta_i)_{i \in N}, (\gamma_i)_{i \in N}, \varphi \rangle$$

where N is a finite set whose elements we interchangeably call *players* or *agents*, $(\Theta_i)_{i \in N}$ is a partition of $\mathcal{P}rops$: each Θ_i contains those variables which player i controls and for each $i \in N$, the formula $\gamma_i \in \mathbf{PL}(\mathcal{P}rops)$ is called the *goal* of player i. Finally, $\varphi \in \mathbf{PL}(\mathcal{P}rops)$ is a satisfiable formula which models a *game constraint*.

As before, we omit the specification of $\mathcal{P}rops$ from a Boolean Pregame, when it is understood from the context.

For each $i \in N$, the set of actions available to player i is 2^{Θ_i}. An *action profile* is a tuple $\overline{\xi} = (\xi_i)_{i \in N}$ of actions, one for each player. We use $\overline{\xi}$ to interchangeably refer to $(\xi_i)_{i \in N}$ and $\cup_{i \in N} \xi_i$. Hence, we write $\overline{\xi} \models \gamma$ instead of $\cup_{i \in N} \xi_i \models \gamma$.

A *feasible* action profile is an action profile $\overline{\xi}$ such that $\overline{\xi} \models \varphi$. We note that φ is useful for limiting the possible game outcomes, without artificially making the game (and players' goals) more complicated. Unless stated otherwise, we consider only feasible action profiles.

We also use "*action profiles*" and "*game outcomes*" interchangeably.

Example 1 (John & Mary - the pregame): We consider $N = \{j, m\}$ where j and m abbreviate the players *John* and *Mary*, respectively. Let $\mathcal{P}rops = \{jr, mr, jf, jf\}$, where the propositional symbol xy encodes the action "*x goes to y*" where $x \in \{j, m\}$ and $y \in \{r, f\}$. r and f abbreviate "*restaurant*" and "*fast-food*", respectively. The partition $\Theta_j = \{jr, jf\}$ and $\Theta_m = \{mr, mf\}$ intuitively reflects the powers of each player. The game constraint $\varphi = \neg(mf \wedge mr) \wedge \neg(jr \wedge jf)$, prevents those outcomes in which either John or Mary simultaneously go to the restaurant and the fast-food. Let:

$$\gamma_1 = (jr \wedge mr) \vee (jf \wedge mf) \quad \gamma_2 = (\neg jr \wedge mr) \vee (\neg jf \wedge mf)$$

Also, let:

$$\mathbb{B}_\alpha = \langle N, (\Theta_i)_{i \in N}, (\gamma_j^\alpha, \gamma_m^\alpha), \varphi \rangle \text{ where } \alpha \in \{\heartsuit, \dagger\}$$

and $\gamma_j^\heartsuit = \gamma_m^\heartsuit = \gamma_1$, $\gamma_j^\dagger = \gamma_1, \gamma_m^\dagger = \gamma_2$. Thus, in the pregame $\mathbb{B}_\heartsuit$, John and Mary prefer going out together, no matter the venue, while in $\mathbb{B}_\dagger$, John would like to go out to whatever location Mary is going, while Mary would like to go out but avoid John.

C. Games

A Boolean Game is a pair:

$$\mathbb{G} = \langle \mathbb{B}, (u_i)_{i \in N} \rangle$$

consisting of a Boolean pregame $\mathbb{B}$ and a tuple of utility functions $u_i : 2^{\mathcal{P}rops} \to \mathbb{R}$, one for each player $i \in N$. The value $u_i(\overline{\xi})$, reflects the payoff which player i receives if the interpretation $\overline{\xi}$ is the outcome of the game.

An action profile $\overline{\xi}$ is attacked by i via $\overline{\xi'}$ (written $\overline{\xi'} \hookrightarrow_i \overline{\xi}$) iff there exists $i \in N$ such that: (i) $\overline{\xi}_{-i} = \overline{\xi'}_{-i}$[1] and (ii) : $u_i(\overline{\xi'}) > u_i(\overline{\xi})$. The intuition is that player i can implement an action profile $\overline{\xi'}$ by himself/herself (condition (i)), in which he/she obtains a strictly higher utility, thus rendering him/her strictly better of (condition (ii)). Whenever such a i and $\overline{\xi'}$ exist for a given $\overline{\xi}$, we say $\overline{\xi}$ *is attacked*.

Definition 2 (Nash Equilibrium): A *Nash Equilibrium* (short. NE) is an action profile $\overline{\xi}$ which is not attacked. We designate by $\mathbf{NE}(\mathbb{G})$, the set of Nash Equilibria of a Boolean Game $\mathbb{G}$.

In what follows, we distinguish between Boolean Games of perfect information, and — in the following section — between two types of games of imperfect information.

Definition 3 (Perfect information): A *Boolean Game of perfect information* is a Boolean Game where the utility functions are defined as:

$$u_i(\overline{\xi}) = \begin{cases} 1 & \overline{\xi} \models \gamma_i \\ 0 & \overline{\xi} \not\models \gamma_i \text{ and } \Theta_i \cap \overline{\xi} = \emptyset \quad \forall i \in N \\ -1 & \text{otherwise} \end{cases}$$

The intuition is that an agent will receive 0 utility if his goal is not satisfied, but none of the variables he/she controls were set: thus, the agent performed a "*do nothing*" action. However, if an agent "*does something*" and his goal is not satisfied, his utility is negative.

Remark 1 (Game definition): Our game definition does not include costs nor player resources, which may make sense in practice. The above definition can be easily extended to include costs (as in, e.g. [3]), and this will not affect our subsequent results.

We note that each Boolean Pregame $\mathbb{B}$ corresponds to one Boolean Game of perfect information, $\mathbb{G}(\mathbb{B})$.

Example 2 (John & Mary - utilities): In the table below, each entry (x, y) shows the utilities x, y of John and Mary, respectively, in different outcomes of $\mathbb{G}(\mathbb{B}_\heartsuit)$. The actions of John and Mary are given on the columns and rows, respectively.We recall that $\emptyset$ is the action denoting *stay at home*,

[1] The notation a_{-i} has it's standard game-theoretic meaning

for both John and Mary. Outcomes which do not satisfy the constraint φ are not considered.

	$\{jr\}$	$\{jf\}$	$\emptyset$
$\{mr\}$	$(1,1)$	$(-1,-1)$	$(0,-1)$
$\{mf\}$	$(-1,-1)$	$(1,1)$	$(0,-1)$
$\emptyset$	$(-1,0)$	$(-1,0)$	$(0,0)$

The utilities of $\mathbb{G}(\mathbb{B}_\dagger)$ are as follows:

	$\{jr\}$	$\{jf\}$	$\emptyset$
$\{mr\}$	$(1,-1)$	$(-1,1)$	$(0,1)$
$\{mf\}$	$(-1,1)$	$(1,-1)$	$(0,1)$
$\emptyset$	$(-1,0)$	$(-1,0)$	$(0,0)$

Example 3 (John and Mary - perfect information): Let $\mathbb{G}_\heartsuit = \mathbb{G}(\mathbb{B}_\heartsuit)$ and $\mathbb{G}_\dagger = \mathbb{G}(\mathbb{B}_\dagger)$ be the Boolean Games corresponding to the Pregames from Example 1. We note that $\mathbf{NE}(\mathbb{G}_\dagger) = \emptyset$. We only give an informal argument: Mary is better of going out when John stays in, and John is better of going at x whenever Mary is going at x. Given such outcome, Mary prefers the other venue. On the other hand, $\mathbf{NE}(\mathbb{G}_\heartsuit) = \{\{jr, mr\}, \{jf, mf\}, \emptyset\}$ reflecting both players' preference to be together. Also, note that if one player stays in, the other's best response is to also stay in.

III. Boolean Games with imperfect information

Imperfect information is often given different interpretations in the context of game theory. In this paper, we assume players are aware of other players' goals and actions, as well as their own. However, they may not distinguish between several game outcomes.

Formally, for each player $i \in N$, we denote by $\sim_i \subseteq 2^{\mathcal{P}rops} \times 2^{\mathcal{P}rops}$ the *indistinguishability relation* of player i. All $\sim_i$ are reflexive and symmetric. The intuition is that, for two outcomes $\overline{\xi}$ and $\overline{\xi'}$ such that $\overline{\xi} \sim_i \overline{\xi'}$, player i cannot distinguish one outcome from the other. Let $[\overline{\xi}]_i = \{\overline{\xi'} : \overline{\xi} \sim_i \overline{\xi'}\}$ be the set of outcomes which player i considers possible, if $\overline{\xi}$ is the outcome of the game.

A. Allowing goals to express agent knowledge

We extend **PL** by adding a new *knowledge* unary operator: K_i. Informally, the formula $K_i\varphi$ expresses that "*agent i knows φ is true*". We denote by **KPL** the resulting language. The semantics of **KPL** is defined w.r.t. interpretations, players, their indistinguishability relations $\sim_i$ and **KPL**-formulae, as follows:

$$\begin{array}{lcl} \overline{\xi}, i \models K_x\gamma & \iff & \forall \overline{\xi'} \in [\overline{\xi}]_x : \overline{\xi'}, x \models \gamma \\ \overline{\xi}, i \models \gamma & \iff & \overline{\xi} \models \gamma \text{ where } \gamma \in \{p, \neg\gamma', \gamma' \wedge \gamma''\} \end{array}$$

$K_x\gamma$, where $x \in N$, is satisfied by an outcome $\overline{\xi}$ and player i, iff on all outcomes $\overline{\xi'}$ which x considers possible, $\overline{\xi'}, x \models \gamma$. In other words, γ must be true for all outcomes which x considers possible.

We note that $\overline{\xi}, i \models K_i\gamma$ expresses *subjective knowledge* on the satisfaction of γ since every outcome which i considers possible satisfies γ. However, the epistemic alternatives to $\overline{\xi}$ are given not by the player i w.r.t. which γ is evaluated, but with respect to the player x designated in K_x. Hence, $\overline{\xi}, i \models K_j\gamma$, $i \neq j$ expresses that $K_j\gamma$ is objectively true, hence i considers possible that j knows γ.

Finally, in the second definition, $\overline{\xi}, i \models \gamma$ expresses that γ is *objectively true* in $\overline{\xi}$. Actually, in the absence of knowledge operators the new semantics "*falls back*" to the standard **PL**semantics. Players are simply ignored.

Example 4 (Knowledge goals in $\mathbb{G}_\heartsuit$ and $\mathbb{G}_\dagger$): We continue the previous examples. We slightly modify the player goals from Example 1:

$$\begin{array}{lcl} \gamma_j^\heartsuit & = & (jr \wedge K_j mr) \vee (jf \wedge K_j mf) \\ \gamma_m^\heartsuit & = & (K_m jr \wedge mr) \vee (K_m jf \wedge mf) \\ \gamma_j^\dagger & = & (jr \wedge K_j mr) \vee (jf \wedge K_j mf) \\ \gamma_m^\dagger & = & (K_m(\neg jr) \wedge mr) \vee (K_m(\neg jf) \wedge mf) \end{array}$$

in order to reflect that each player desires subjective knowledge about the other's whereabouts. For instance $\gamma_m^\dagger$ expresses that Mary is at the restaurant and she knows John is not at the restaurant or Mary is at the fastfood and she knows John is not at the fastfood.

B. Imperfect information - general setting

A *Boolean Game of imperfect information* is a Boolean Game where the utility functions are defined w.r.t. $(\sim_i)_{i\in N}$, as:

$$u_i(\overline{\xi}) = \begin{cases} 1 & \overline{\xi}, i \models \gamma_i \\ 0 & \overline{\xi}, i \not\models \gamma_i \text{ and } \Theta_i \cap \overline{\xi} = \emptyset \quad \forall i \in N \\ -1 & \text{otherwise} \end{cases}$$

Note that the functions u_i extend those from Section 3, by considering the new "*knowledge*" semantics $\models \subseteq 2^{\mathcal{P}rops} \times N \times \mathbf{PL}(\mathcal{P}rops)$.

Remark 2 (Specifying $\sim_i$): We note that the specification of $\sim_i$ requires exponential space w.r.t. to $\mathcal{P}rops$, since there is an exponential number of valuations of the variables in $\mathcal{P}rops$. To prevent having an exponential description of a game with imperfect information, we need a compact representation for $\sim_i$.

In what follows, we introduce two such representations.

1) Visibility sets [6]: For each player $i \in N$, we define $\Gamma_i \subseteq \mathcal{P}rops$ as the *visibility set* of player i. Intuitively, i can observe only those variables in Γ_i. We further require that $\Gamma_i \cap \Theta_i = \emptyset$: each agent can observe the variables he/she controls. Each Γ_i induces an indistinguishability relation, defined as follows:

$$\overline{\xi} \sim_i \overline{\xi'} \text{ iff } p \in clash(\overline{\xi}, \overline{\xi'}) \implies p \in \Gamma_i$$

In words, two outcomes $\overline{\xi}$ and $\overline{\xi'}$ are indistinguishable by player i iff they assign different truth values only for those variables in Γ_i.

We denote by $\mathbb{G}^i_{\overline{\Gamma}}(\mathbb{B})$ the Boolean Game with imperfect information which corresponds to the Boolean Pregame $\mathbb{B}$ and to indistinguishability relations $(\sim_i)_{i\in N}$ induced by $\overline{\Gamma} = (\Gamma_i)_{i\in N}$.

Example 5 (John and Mary, visibility sets): Assume $\Gamma_j = \{mf, mr\}$ and $\Gamma_m = \{jf, jr\}$, hence neither John nor Mary can observe the actions of the other.

We note that $\mathbf{NE}(\mathbb{G}^i_{\overline{\Gamma}}(\mathbb{B}_\heartsuit)) = \{\emptyset\}$. Consider for instance $\overline{\xi} = \{jr, mr\}$. Thus $\overline{\xi}, j \not\models K_j mr$ since there exists $\overline{\xi'} = \{jr\} \in [\overline{\xi}]_i$, such that $\overline{\xi'}, j \not\models mr$. Thus, the goal of John is not satisfied, since, subjectively, he does not know Mary joined him at the restaurant. Informally, no player can observe the actions of the other: John and Mary are not able to coordinate.

Remark 3 (Visibility sets): While visibility sets are a compact way for specifying incomplete information, they face limitations. Consider our previous example, which exhibits two problematic issues. The first is that the indistinguishability relation of each player is counter-intuitive. It is natural for John not to know the whereabouts of Mary, if, for instance, John is at the restaurant, while Mary is not. Mary may be at the fast-food, or at home. However, if both John and Mary are at the restaurant, we expect them to both be aware of this.

To generalize, it is not uncommon that certain (otherwise unobservable) variables do become observable, in specific outcomes.

The second issue is related to equilibria: due to imperfect information, the players cannot coordinate, although we would intuitively expect them to.

2) Visibility relations: We have argued that the approach based on visibility sets may be over-restrictive for specifying imperfect information, in certain settings. We propose an alternative, which relies on the following:

For each variable $p \in \mathcal{P}rops$ we introduce elements $p_\top$ and $p_\perp$. They intuitively correspond to a truth-value being assigned to p. Let $\mathcal{P} = \{(p, x) : p \in \mathcal{P}rops, x \in \{\top, \perp\}\}$ designate the set of truth value assignments for variables in $\mathcal{P}rops$. For conveniency, we write p_x instead of (p, x).

For each player i, we define the relation $E_i \subseteq \mathcal{P} \times \mathcal{P}rops$, where $(q_x, p) \in E_i$ means that if q has the truth-value of x, then player i can observe the truth value of p (p is *revealed* by q_x), where $x \in \{\top, \perp\}$. Each E_i induces an indistinguishability relation as follows:

$$\overline{\xi} \sim_i \overline{\xi'} \text{ iff } \forall p \in clash(\overline{\xi}, \overline{\xi'}) : \quad \begin{array}{l} (q_\top, p) \in E_i \iff q \notin \overline{\xi} \cup \overline{\xi'} \wedge \\ (q_\perp, p) \in E_i \iff q \in \overline{\xi} \cup \overline{\xi'} \end{array}$$

Thus, two outcomes $\overline{\xi}, \overline{\xi'}$ are indistinguishable iff each variable p on which the outcomes disagree is not *revealed* by setting some q to *true* (resp. *false*) in either outcome.

We denote by $\mathbb{G}^i_{\overline{E}}(\mathbb{B})$ the Boolean Game with imperfect information which corresponds to the Boolean Pregame $\mathbb{B}$ and to indistinguishability relations $(\sim_i)_{i \in N}$ induced by $\overline{E} = (E_i)_{i \in N}$.

Example 6 (John & Mary - visibility relations): Consider:

$$\begin{array}{rcl} E_j &=& \{(jr_\top, mr), (jf_\top, mf)\} \cup \\ && \{(jx_y, jx) : x \in \{r, f\}, y \in \{\top, \perp\}\} \\ E_m &=& \{(mr_\top, jr), (mf_\top, jf)\} \cup \\ && \{(mx_y, mx) : x \in \{r, f\}, y \in \{\top, \perp\}\} \end{array}$$

Informally, the visibility relations express that both John and Mary are able to observe if the other player joined them at the respective venue. Also, the variables they control are observable.

We first note that $\{jf, mf\} \not\sim_j \{jf\}$, since the two interpretations disagree on mf, but the latter is revealed in the first interpretation by $jf_\top$. If John went to the restaurant, he knows whether or not Mary joined him. However, $\{jr\} \sim_j \{jr, mf\}$ — if John goes to the restaurant, he cannot distinguish between those outcomes where Mary goes to the fast-food or stays at home. Also note that $\{jr\} \not\sim_j \{mr\}$, the valuation of jr is different in the two outcomes, but $jr_\perp$ reveals its truth-value. Informally, John can distinguish between an outcome where he is at the restaurant and one where he is at home.

We continue Example 5. Suppose $\overline{\xi} = \{jf, mf\}$. We note that $[\overline{\xi}]_j = \{\overline{\xi}\}$, and thus: $\overline{\xi}, j \models jf \wedge K_j mf$. We also observe that $\mathbf{NE}(\mathbb{G}^i_{\overline{E}}(\mathbb{B}_\heartsuit)) = \{\{jf, mf\}, \{jr, mr\}, \emptyset\}$, which confirms our intuition that the two players may coordinate in this particular case.

Next, we analyse $\mathbb{G}^i_{\overline{E}}(\mathbb{B}_\dagger)$. Suppose $\overline{\xi} = \{jr, mf\}$, and recall that $[\overline{\xi}]_j = \{\{jr, mf\}, \{jr\}\}$. We also note that $\overline{\xi}, j \models K_j \neg mr$: since John is at the restaurant and Mary is not, John does not consider possible that Mary is at the restaurant. Moreover, John knows Mary is not at the restaurant.

We also note that $\overline{\xi'} = \{jf, mf\} \hookrightarrow_j \overline{\xi}$, since if John were to go to the restaurant, he would be in the company of Mary, and moreover, have subjective knowledge of this, since $jf_\top$ reveals mf.

Such a deviation of John is *objective* ("theoretically" possible) but not *subjective*: given the current outcome, John could not know if Mary stayed at home or went to the restaurant.[2]

The main issue regarding the current Nash Equilibrium definition is that attacks are considered objectively, and not in the terms of the knowledge possessed by the player performing the attack. A possible fix of Definition 2, is to re-define the notion of an attack:

$$\overline{\xi}' \hookrightarrow_i \overline{\xi} \text{ iff } \forall \overline{\xi}'' \in [\overline{\xi'}] : \overline{\xi}_{-i} = \overline{\xi}''_{-i} \wedge u_i(\overline{\xi}'') > u_i(\overline{\xi})$$

In words, $\overline{\xi'}$ is an attack, iff in all outcomes which player i considers possible from $\overline{\xi}$, the original attack conditions hold. We defer a more careful study of the former definition, and on the relation to the original one, for future work.

IV. Related Work

Imperfect information is extensively studied in games [1] as well as in logics for Multiagent systems [7]. In this section, we restrict our attention to imperfect information and how it is modelled in Boolean Games. We consider the work from [4] to be the closest to our approach. The authors of [4] model imperfect information using visibility sets, and introduce *equilibria verifiability*, as a means for selection of stable outcomes. A (Nash) equilibrium is *verifiable* iff it satisfies some *observational property*, for each player. The authors distinguish between three types of such properties (and implicitly, of types

[2] In this particular case, John may deduce that Mary is at the restaurant, since she is better of not staying home, but if we consider more than two venues, this reasoning no longer holds.

of verifiability). We recall only two: *weak verifiability* which requires that players consider a Nash Equilibrium to be a possible game outcome and *strong verifiability* which requires that players know a Nash Equilibrium was played (a non-Nash Equilibrium as a game outcome is impossible).

We conjecture that *strong verifiability* is equivalent to our definition of Nash Equilibrium, in games of imperfect information induced by visibility sets, where goals γ_i are of the form $K_i\gamma$, where $\gamma \in \mathbf{PL}$ (hence no nested knowledge operators are present). Thus, our work extends that from [4], by introducing a more general model of imperfect information, and by allowing agents to express goals related to the knowledge of their peers in the game. This latter feature is similar to Public Announcement Games (PAG) [8]. In a PAG, agents are also able to express knowledge-related goals in an incomplete information setting. However, their available actions consist in making public (truthful) announcements. Thus, the dynamics of PAG relates to how information is shared among agents. While information-sharing is possible in our setting as well (e.g. some agent may set some variable p to *true/false*, and thus reveal another variables' valuation to other agents), the extent to which this is possible depends on the model itself (the relations E_i).

Incomplete information is also tackled in [5]: the authors separate between *environment variables* and *agent-controlled variables*. Players have incomplete information only on valuations of the former. Imperfect information is also modelled using visibility sets. The technical work of [5] is focused on making announcements regarding the valuation of *environment variables*, and — in this respect — is closer to PAG than to ours.

Finally, in [9], imperfect information is considered only with respect to the beliefs regarding other players' goals.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we have introduced an alternative modelling approach for incomplete information, together with the ability to express knowledge-related goals in Boolean Games. The proposal has been illustrated using examples: all interesting results have been deferred for future work. For instance, it is interesting to see what is the relation between (Nash) equilibria of Boolean Games with imperfect information modelled using visibility sets (Section III-B1), i.e. *strongly verifiable equilibria* [4], and those with imperfect information using visibility relations (Section III-B2).

Also, we have claimed that complexity is expected to remain unchanged for the verification and stability problems mentioned in the introductory section. From a computational view, our approach simply replaces a tractable utility function with another. However, this remains to be shown formally.

Another interesting direction to explore is related to mechanism design. Unlike game theory, which studies equilibria under given settings, mechanism design seeks those settings under which given equilibria are realised. In our setting, mechanism design is useful for determining those incentives ("*type/amount of information*") which agents require in order to implement desirable outcomes.

REFERENCES

[1] Y. Shoham and K. Leyton-Brown, *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, 2008.

[2] P. Harrenstein, W. van der Hoek, J.-J. Meyer, and C. Witteveen, "Boolean games," in *Proceedings of the 8th Conference on Theoretical Aspects of Rationality and Knowledge*, ser. TARK '01. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2001, pp. 287–298. [Online]. Available: http://dl.acm.org/citation.cfm?id=1028128.1028159

[3] P. E. Dunne, W. van der Hoek, S. Kraus, and M. Wooldridge, "Cooperative boolean games," in *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 2*, ser. AAMAS '08. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems, 2008, pp. 1015–1022. [Online]. Available: http://dl.acm.org/citation.cfm?id=1402298.1402363

[4] T. Ågotnes, P. Harrenstein, W. Van Der Hoek, and M. Wooldridge, "Verifiable equilibria in boolean games," in *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, ser. IJCAI '13. AAAI Press, 2013, pp. 689–695. [Online]. Available: http://dl.acm.org/citation.cfm?id=2540128.2540229

[5] J. Grant, S. Kraus, M. Wooldridge, and I. Zuckerman, "Manipulating games by sharing information," *Studia Logica*, vol. 102, no. 2, pp. 267–295, 2014. [Online]. Available: http://dx.doi.org/10.1007/s11225-014-9544-5

[6] T. Ågotnes, P. Harrenstein, W. Van Der Hoek, and M. Wooldridge, "Verifiable equilibria in boolean games," in *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, ser. IJCAI '13. AAAI Press, 2013, pp. 689–695. [Online]. Available: http://dl.acm.org/citation.cfm?id=2540128.2540229

[7] P.-Y. Schobbens, "Alternating-time logic with imperfect recall," *Electronic Notes in Theoretical Computer Science*, vol. 85, no. 2, pp. 82 – 93, 2004, {LCMAS} 2003, Logic and Communication in Multi-Agent Systems. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1571066105826040

[8] T. gotnes and H. van Ditmarsch, "What will they say?public announcement games," *Synthese*, vol. 179, no. 1, pp. 57–85, 2011. [Online]. Available: http://dx.doi.org/10.1007/s11229-010-9838-8

[9] S. De Clercq, S. Schockaert, M. De Cock, and A. Now, "Possibilistic boolean games: Strategic reasoning under incomplete information," in *Logics in Artificial Intelligence*, ser. Lecture Notes in Computer Science, E. Ferm and J. Leite, Eds. Springer International Publishing, 2014, vol. 8761, pp. 196–209.