Modelling Incomplete Information in Boolean Games using Possibilistic Logic☆

Sofie De Clercq[a], Steven Schockaert[b,*], Ann Nowé[c], Martine De Cock[d,a]

[a] *Dept. of Applied Mathematics, Computer Science & Statistics, Ghent University, Krijgslaan 281, 9000 Gent, Belgium*
[b] *School of Computer Science & Informatics, Cardiff University, Queen's Buildings, 5 The Parade, Roath, Cardiff CF24 3AA, UK*
[c] *Computational Modeling Lab, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium*
[d] *Institute of Technology, UW Tacoma, 1900 Commerce Street, Tacoma WA, 98402, US*

Abstract

Boolean games (BGs) are a framework for specifying strategic games in which the utility of an agent is determined based on the satisfaction of goals in propositional logic. The majority of existing work on BGs relies on the often unrealistic assumption that agents have perfect knowledge of each other's preferences. In this paper, we show how this issue can be addressed in a natural way, by replacing the use of classical logic for expressing agents' goals by possibilistic logic. We consider two such settings. In the first setting, possibilistic logic is used to encode knowledge about other agents' goals with different levels of certainty. In the second setting, which is based on generalized possibilistic logic, certainty levels are instead used to compactly encode priorities, while incompleteness is modelled in a binary way, similar as in epistemic modal logics. In both cases we introduce natural solution concepts, motivated by Schelling's theory on focal points: a certain pure Nash equilibrium (PNE) is preferred over another one due to the fact that all agents *know* it to be a PNE. Alternatively, an outcome might be preferred when all agents *consider it possible* of being a PNE. We prove that the associated computational complexity of these solution concepts does not increase compared to PNEs in Boolean games with complete information. Finally, to illustrate the practical relevance, we consider an application to negotiation, among others showing how knowledgeable agents can obtain a more desirable outcome than others.

Keywords: Boolean games, game theory, possibilistic logic, negotiation

☆This research is funded by the Research Foundation Flanders (FWO). Steven Schockaert was supported by a grant from the Leverhulme Trust (RPG-2014-164).

*Corresponding author

Email addresses: SofieR.DeClercq@ugent.be (Sofie De Clercq), SchockaertS1@cardiff.ac.uk (Steven Schockaert), ANowe@vub.ac.be (Ann Nowé), MDeCock@uw.edu (Martine De Cock)

1. Introduction

Game-theoretic frameworks often rely on the somewhat artificial assumption that agents are fully aware of each other's goals. In strategic settings, agents might deliberately conceal such information, or might not have had a chance to exchange it. Even if an agent knows its opponents well, it may not be fully certain about what exact goals the other agents are pursuing. For instance, suppose Alice and Bob are married and plan a night out. The options are going to the theatre and attending a sports game. Even if we assume that Alice and Bob have been married for an eternity and know each other *inside out*, they are not mind readers: it is still possible that Alice is not entirely sure whether Bob really prefers to join her to the theatre over attending the sports game alone. Games with incomplete information [1, 2] allow us to explicitly model the limitations of agents' knowledge about the preferences of the others. While this topic has been extensively researched for normal form games, Boolean games with incomplete information have hardly received any attention; see Section 2 for a discussion.

In this paper, we study the use of possibilistic logic [3] to model Boolean games with incomplete information. Possibilistic logic has the advantage of staying close to classical logic, while offering us more flexibility, and it can be naturally used to model (partial) ignorance [4]. Specifically, we study two different settings which differ in how the necessity degrees from the possibilistic logic theories are interpreted. In the first setting, we consider the usual interpretation of necessity degrees as certainty degrees. It uses possibilistic knowledge bases to encode necessary and sufficient conditions for the satisfaction of other agents subgoals. For instance, Alice can encode that she is absolutely certain that Bob reaches his highest utility when they both attend the sports game, while she is less certain that Bob reaches his highest utility when they both go to the theatre. We prove that this framework at the semantic level corresponds to a possibility distribution over all possible games.

In the second approach, necessity degrees are instead used to model preference. To keep the ability to model incomplete information, for this setting we switch from standard possibilistic logic to generalized possibilistic logic (GPL) [4]. In particular, each GPL theory semantically corresponds to a set of possibility distributions. In our context, these possibility distributions are interpreted as the utility functions that, according to a given agent, may correspond to the actual utility function of some other agent.

For both frameworks of Boolean games with incomplete information, we propose intuitive solution concepts, reflecting whether agents know or consider it possible that a certain outcome is a pure Nash equilibrium (PNE). We prove that the computational complexity of the associated decision problems does not increase compared to PNEs in Boolean games with complete information. To illustrate how these solution concepts could be useful, we briefly discuss an application to negotiation in Boolean games with incomplete information.

The remainder of this paper is structured as follows. We first discuss related work in Section 2 and give some background on Boolean games and possibilistic logic in Section 3. Next we introduce our two approaches in Section 4 and Section 5. Finally, we present the application to negotiation in Section 6. Finally note that this paper is an extension of our work in [5] (Section 4) and [6] (Sections 5 and 6). In addition to providing more detailed explanations and proofs, we have extended the framework from [5] with joint constraints and have added complexity results to the framework from [6].

2. Related Work

Although uncertainty has been studied extensively in the context of game theory (see e.g. [1]), the literature on Boolean games with incomplete information is currently limited. Note that we are only concerned with epistemic uncertainty in this paper (e.g. we do not consider stochastic actions, whose outcome cannot be predicted with certainty). To the best of our knowledge, stochastic uncertainty has not yet been studied in the context of Boolean games. Moreover, we are not aware of any existing approaches for modelling uncertainty w.r.t. the goals of other agents, in the context of Boolean games, although uncertain Boolean games have been studied for other purposes. For example, in [7] uncertain Boolean games are modelled by introducing a set of environment variables which are outside the control of any agent. Each agent has some (possibly incorrect) belief about the value of the environment variables. The focus of [7] is on manipulating Boolean games by making announcements about the true value of some environment variables, in order to create a stable solution if there were none without the announcements. In [8] uncertainty is modelled by extending the framework of Boolean games with a set of observable action variables for every agent, i.e. every agent can only observe the values assigned to a particular subset of action variables. As a result, agents are not able to distinguish between some outcomes, if these profiles only differ in action variables that are not observable to that agent. Three notions of verifiable equilibria are investigated, capturing respectively outcomes for which all agents know that they *might be* pure Nash equilibria (PNEs), outcomes for which all agents know that they *are* PNEs and outcomes for which it is common knowledge that they are PNEs, i.e. all agents know that they are PNEs and all agents know that all agents know that they are PNEs etc. The same authors have extended this framework to epistemic Boolean games [9], in which the logical language for describing goals is broadened to a multi-agent epistemic modal logic. Note, however, that agents are still fully aware of each other's goals in this framework, i.e. [9] considers agents whose goal is to obtain a particular epistemic state. For instance, *I not only want my husband to pick up our baby, I also want to know he is picking up our baby.*

In contrast, we study Boolean games with incomplete information, considering agents which have their own beliefs about the goals of other agents. Although probability theory is often used to model uncertainty in game theory [1], a possibilistic logic approach provides a simple and elegant mechanism for modelling partial ignorance, which is closely related to the notion of epistemic entrenchment [10]. Being based on ranking formulas (at the syntactic level) or ranking possible worlds (at the semantic level), possibilistic logic has the advantage of staying close to classical logic. As a result, we can introduce methods for solving possibilistic Boolean games that are closely similar to methods for solving standard Boolean games.

Within the broader context of game theory, several authors have looked at qualitative ways of modelling epistemic uncertainty. A common approach is to model the beliefs of an agent a about another agent b as a set of pairs (s, t), where s is a strategy and t is a so-called type (where types are used to model beliefs about beliefs in a hierarchical way). Such a belief structure is sometimes called *possibilistic* in the game theory literature (e.g. [11, 12]). However, it should be noted that these approaches are not related to possibility theory in the sense of [13], and are thus different in spirit from the "possibilistic" approach we present in this paper. In particular, while the aforementioned approaches rely on a Boolean uncertainty model (i.e. a given strategy s is either considered possible or not), we use possibility theory to model uncertainty (in Section 4) or preferences (in Section 5) based on a ranking semantics.

3. Preliminaries

3.1. Background on Boolean Games

The logical language L_Φ associated with a finite set of atoms Φ contains the following formulas: (i) every atom of Φ, (ii) the logical constants $\bot$ and $\top$, and (iii) the formulas $\neg\varphi$ and $\varphi \wedge \psi$ for every φ, $\psi \in L_\Phi$. As usual, we use the abbreviations $\varphi \rightarrow \psi \equiv \neg(\varphi \wedge \neg\psi)$ and $\varphi \vee \psi \equiv \neg(\neg\varphi \wedge \neg\psi)$. We write $Lit(\Phi)$ to denote the set of literals of the language L_Φ, i.e. $Lit(\Phi) = \Phi \cup \{\neg p \,|\, p \in \Phi\}$. An interpretation of Φ is defined as a subset ν of $Lit(\Phi)$ such that for every atom $p \in \Phi$ either $p \in \nu$ or $\neg p \in \nu$. We denote the set of all interpretations of Φ as $Int(\Phi)$. An interpretation can be extended to a satisfaction relation on L_Φ in the usual way. If a formula $\varphi \in L_\Phi$ is satisfied by an interpretation ν, we denote this as $\nu \models \varphi$. An interpretation that satisfies a given formula is called a model of that formula. We denote the set of models of φ as $[\![\varphi]\!]$.

Originally, the utilities in BGs were binary, but several extensions have been introduced to allow more general preferences. Examples are the addition of costs [14], the use of a prioritized goal base instead of a single goal [15, 5] or the use of many-valued Łukasiewicz logic to formalize the idea of weighted goal satisfaction [16]. In our paper, we use the definition of a BG as stated in [5]. The latter is a particular case of generalized BGs [15] in which the preference relations are total. Additionally, we incorporate a constraint δ, restricting the possible joint actions of the agents. This is a generalization of the constraints in [17], which only restrict the individual actions of the agents. Such joint constraints might at first glance seem to conflict with the autonomous character of agents, since it creates a dependency between their actions: the strategy choice of one agent can restrict the available strategies of another agent. However, such dependencies are clearly all around us: agents cannot, for instance, buy the same house, complete the same one man's job, marry the same person etc. In particular, such dependencies naturally force agents to negotiate about their strategy choices, which is the subject of Section 6. Moreover, since the Boolean game framework with joint constraints is a generalization of the one from [17], we explain our theory for the former.

Definition 1 (Boolean Game). A Boolean game (BG) is a tuple $G = (N,\, (\Phi_i)_{i\in N},\, \delta,\, (\Gamma_i)_{i\in N})$. For every agent i in $N = \{1, \ldots, n\}$, Φ_i is a finite set of atoms such that $\Phi_i \cap \Phi_j =$

$\emptyset, \forall j \neq i$. We write $\Phi = \bigcup_{i \in N} \Phi_i$. For every $i \in N$, $\Gamma_i = \{\gamma_i^1; \ldots; \gamma_i^p\}$ is i's prioritized goal base. The formula $\gamma_i^m \in L_\Phi$ is agent i's goal of priority m. We assume that every agent has p priority levels and that $\delta \wedge \gamma_i^m \not\models \bot$ for every $i \in N$ and $m \in \{1, \ldots, p\}$. Finally, δ is a consistent formula in L_Φ, which encodes the integrity constraints of the game G.

The set Φ contains all action variables. Agent i controls Φ_i and can set these atoms to true or false. We also write $\Phi_{-i} = \Phi \setminus \Phi_i$ for the set of action variables outside i's control. By convention, goals are ordered from high (level 1) to low priority (level p).

Example 1. Alice and Bob, who share a car, are planning their afternoon. Alice controls $\Phi_1 = \{b_A, f_A, d_A\}$ and Bob controls $\Phi_2 = \{b_B, f_B\}$. Agent i can drive to the beach (set b_i to true) or to the forest (set f_i to true). If Alice sets d_A to true, she takes the dog with her. The game is constrained by $\delta = \neg(b_B \wedge f_B) \wedge \neg(b_A \wedge f_A) \wedge (b_B \to \neg f_A) \wedge (b_A \to \neg f_B)$. In words: neither Bob nor Alice can simultaneously go to the beach and the forest. Moreover, if Bob goes to the beach then Alice cannot go to the forest and similarly Bob cannot go to the forest if Alice goes to the beach. Alice and Bob's goal bases are:

$$\Gamma_1 = \{f_A \wedge f_B \wedge d_A; b_A \wedge b_B \wedge d_A; d_A\}$$
$$\Gamma_2 = \{b_B \wedge b_A \wedge \neg d_A; b_B \wedge b_A; f_B \wedge f_A\}$$

Thus Alice's first priority is to go to the forest with Bob and her dog. If this is not possible, she would like to go to the beach with Bob and her dog. Furthermore, we can infer that Alice prefers staying at home with her dog over leaving without it. Bob prefers to take Alice to the beach without the dog. However, he still prefers to go to the beach with Alice *and* the dog over going to the forest with Alice, and he prefers going to the forest with Alice over all remaining possibilities.

Definition 2 (Outcome). An interpretation of Φ that satisfies δ is called an outcome of G. We denote the set of all outcomes as $\mathcal{S}$, i.e. $\mathcal{S} = [\![\delta]\!]$.

Given an outcome ν, we write ν_{-i} for the projection of the outcome ν on Φ_{-i}, i.e. $\nu_{-i} = \nu \cap \Phi_{-i}$. Furthermore, we write $\mathcal{S}_i(\nu_{-i})$ for the set of partial outcomes $\nu_i \in \Phi_i$ that can extend the partial outcome $\nu_{-i} \in \Phi_{-i}$, i.e. $\mathcal{S}_i(\nu_{-i}) = \{\nu_i \in \mathit{Int}(\Phi_i) \mid (\nu_{-i} \cup \nu_i) \in \mathcal{S}\}$. Sometimes, we will write an outcome ν as (ν_{-i}, ν_i) to make explicit that the action variables from Φ_i are chosen in accordance with ν_i in the outcome ν. In particular, (ν_{-i}, ν_i') will denote a modification to the outcome ν in which the action variables for agent i are instead chosen in accordance with ν_i'. For the ease of presentation, we define a utility function that is scaled to the unit interval.

Definition 3 (Utility Function). For each $i \in N$ and $\nu \in \mathcal{S}$, the utility of i in ν is defined as

$$u_i(\nu) = \frac{p + 1 - \min\{k \mid 1 \leq k \leq p, \nu \models \gamma_i^k \wedge \delta\}}{p}$$

with $\min \emptyset = p + 1$.

Note that the utility can take $p+1$ possible values, namely those in $\Lambda_p = \{0, \frac{1}{p}, \frac{2}{p}, \ldots, 1\}$. We will denote the vector of utilities $(u_1(\nu), \ldots, u_n(\nu))$ corresponding to outcome ν as $\mathcal{U}(\nu)$. In Example 1, we have for instance:

$$\mathcal{U}(\{f_A, f_B, d_A, \neg b_A, \neg b_B\}) = (1, 0.33),$$
$$\mathcal{U}(\{b_B, b_A, \neg f_A, \neg f_B, \neg d_A\}) = (0, 1),$$
$$\mathcal{U}(\{b_A, b_B, d_A, \neg f_A, \neg f_B\}) = (0.67, 0.67),$$
$$\mathcal{U}(\{f_A, b_B, \neg f_B, \neg b_A, \neg d_A\}) = (0, 0).$$

Definition 4 (Pure Nash equilibrium). An outcome ν of a BG G is a pure Nash equilibrium (PNE) iff for every agent $i \in N$, ν_i is a best response (BR) to ν_{-i}, i.e. $u_i(\nu) \geq u_i(\nu_{-i}, \nu'_i)$ for all $\nu'_i \in \mathcal{S}_i$.

The concept of utility function highlights the close connection there is between Boolean games and normal-form games (NFGs). The main difference between the two frameworks is that in NFGs a utility value for each outcome is encoded explicitly, while in BGs utilities are defined implicitly based on logical formulas. This means that in practice, BGs can be exponentially more compact than the NFG representation of the same game. Clearly, for every BG we can always construct an NFG that is equivalent to it (in the sense that it induces the same utility function). Conversely, given an NFG in which all utilities belong to Λ_p, we can always construct an equivalent BG as well (e.g. by introducing one action variable for each action from the NFG, and adding the constraint that each agent has to set exactly one of these action variables to true). Note that in this paper, we will only be concerned with solution concepts that rely on the relative ordering of the utility values, such as PNEs (but unlike e.g. mixed equilibria). Clearly the requirement that all utilities belong to Λ_p does then not restrict the kind of games that can be encoded. Furthermore note that in finite settings, we can always model games with real-valued utilities, by considering an injective mapping from utility scores in Λ_p to $\mathbb{R}$. On the other hand, the results we present in this paper would not be directly applicable to settings where agents can choose between an infinite number of actions (as in Łukasiewicz games [16]), nor in settings where utilities are partially ordered.

In the context of bargaining (see Section 6), it is natural that agents try to achieve an outcome that is, among others, efficient. A well-known efficiency concept is Pareto optimality.

Definition 5 (Pareto Optimality). For every $\nu, \nu' \in \mathcal{S}$ it holds that ν Pareto dominates ν', denoted as $\nu >_{par} \nu'$, iff

$$(\forall i \in N : u_i(\nu) \geq u_i(\nu')) \wedge (\exists i \in N : u_i(\nu) > u_i(\nu'))$$

We denote the set of Pareto optimal outcomes in $\mathcal{S}$ as

$$PO = \{\nu \in \mathcal{S} \mid \neg(\exists \nu' \in \mathcal{S} : \nu' >_{par} \nu)\}$$

Intuitively, an outcome is Pareto optimal if no agent can be better off without another agent being worse off. It is easy to see that every BG has at least one Pareto optimal outcome.

A well-known refinement of the Pareto ordering incorporating a notion of fairness is the discrimin ordering [18]. To define it, we denote the set of agents whose utility is the same in ν and ν' as $eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))$, i.e. $eq(\mathcal{U}(\nu),\mathcal{U}(\nu')) = \{i \in N \mid u_i(\nu) = u_i(\nu')\}$.

Definition 6 (Discrimin Ordering). For every $\nu, \nu' \in \mathcal{S}$ it holds that $\nu >_{discr} \nu'$ iff

$$\min_{j \notin eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))} u_j(\nu) > \min_{j \notin eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))} u_j(\nu')$$

We define the set of discrimin optimal outcomes as

$$DO = \{\nu \in \mathcal{S} \mid \neg(\exists \nu' \in \mathcal{S} : \nu' >_{discr} \nu)\}$$

It is easy to see that $>_{discr}$ is a strict order relation on $\mathcal{S}$. It holds that $DO \subseteq PO$ and $DO \neq \emptyset$. In Example 1, $\{b_A, b_B, d_A, \neg f_A, \neg f_B\}$ is the unique discrimin optimal outcome, although it is not the only Pareto optimal outcome.

3.2. Background on Possibility Theory

3.2.1. Possibility theory

Possibilistic logic (see e.g. [19] for a more comprehensive overview) is a popular tool to encode and reason about uncertain information in an intuitive and compact way. The semantics is defined based on possibility distributions [13]:

Definition 7 (Possibility Distribution). A possibility distribution on the universe $\mathcal{W}$ is a mapping $\pi : \mathcal{W} \rightarrow [0, 1]$.

The elements of the universe $\mathcal{W}$ are referred to as possible worlds[1]. Let us denote an arbitrary element of the universe $\mathcal{W}$ as ν. If $\pi(\nu) = 1$, ν is considered to be completely possible, whereas $\pi(\nu) = 0$ corresponds to ν being completely impossible. A possibility distribution such that $\pi(\nu) = 1$ for every $\nu \in \mathcal{W}$ corresponds to a state of complete ignorance, since all options are completely possible. For instance, consider two possible worlds, denoted as w and $\neg w$, corresponding respectively to a specific nation being weaponized or not. If one has no information regarding the weapons of this nation, both worlds are completely possible, i.e. $\pi(w) = \pi(\neg w) = 1$. A possibility distribution π is called normalized if at least one world is considered completely possible, i.e. $\exists \nu \in \mathcal{W} : \pi(\nu) = 1$. Note that in contrast to probability distributions, there is no requirement that possibility distributions satisfy $\sum_{\nu \in \mathcal{W}} \pi(\nu) = 1$. A possibility distribution encodes which worlds cannot be excluded based on available knowledge. Therefore, smaller possibility degrees are more specific, as they reflect a higher degree of certainty that some worlds can be ruled out. For instance, suppose you suspect that your neighbouring nation is weaponized. This could be encoded through the possibility distribution $\pi(w) = 1$ and $\pi(\neg w) = 0.5$. If you have evidence that your neighbouring nation is

[1] Outside possibilistic logic, possibility distributions can also be defined over universa that do not represent possible worlds, e.g. the real numbers.

weaponized (you are certain of it), this could be encoded through the possibility distribution $\pi(w) = 1$ and $\pi(\neg w) = 0$. The latter is more specific than the former. An ordering $\leq$ on all possibility distributions on $\mathcal{W}$ can be defined as $\pi_1 \leq \pi_2$ iff it holds that $\pi_1(\nu) \leq \pi_2(\nu)$, $\forall \nu \in \mathcal{W}$, assuming the natural ordering on $[0,1]$. We say that π_1 is at least as specific as π_2 when $\pi_1 \leq \pi_2$. Given a set of possibility distributions, the maximal elements w.r.t. $\leq$ are called the least specific possibility distributions, as these correspond to the smallest amount of information. A possibility and necessity measure are induced by a possibility distribution in the following way.

Definition 8 (Possibility and Necessity Measure). Given a possibility distribution π on a universe $\mathcal{W}$, the possibility $\Pi(A)$ and necessity $N(A)$ that an event $A \subseteq \mathcal{W}$ occurs is defined as:

$$\Pi(A) = \sup_{\nu \in A} \pi(\nu); \qquad N(A) = \inf_{\nu \notin A} (1 - \pi(\nu))$$

Note that $\Pi(A)$ measures the degree to which the event A is compatible with available evidence, whereas $N(A)$ measures the degree to which the event A is implied by the available evidence.

Example 2. Suppose a specific nation has two neighbouring nations. If the first (respectively the second) is weaponized, we denote this as w_1 (respectively w_2). Consider the universe $\mathcal{W} = \{\{w_1, w_2\}, \{w_1, \neg w_2\}, \{\neg w_1, w_2\}, \{\neg w_1, \neg w_2\}\}$ and the possibility distribution π on $\mathcal{W}$:

$$\begin{aligned} \pi(\{w_1, w_2\}) &= 1 & \qquad \pi(\{\neg w_1, w_2\}) &= 0.4 \\ \pi(\{w_1, \neg w_2\}) &= 0.4 & \qquad \pi(\{\neg w_1, \neg w_2\}) &= 0 \end{aligned}$$

Thus the nation considers it impossible that none of its neighbours is armed and has limited certainty that both of the neighbours are armed. Now consider the event A that the first neighbour is weaponized, i.e. $A = \{\{w_1, w_2\}, \{w_1, \neg w_2\}\}$. Then the possibility $\Pi(A)$ is 1, i.e. it is completely possible that the first neighbour is armed. The necessity $N(A)$ is 0.6, i.e. it is necessary to degree 0.6 that the first neighbour is armed.

Another uncertainty measure which is sometimes used in possibility theory is called guaranteed possibility. It is defined for an event $A \subseteq \mathcal{W}$ as follows:

$$\Delta(A) = \inf_{\nu \in A} \pi(\nu)$$

In other words, $\Delta(A)$ reflects the degree to which all elements of A are considered possible. This measure will be useful to express limitations on the knowledge of an agent.

3.2.2. Possibilistic logic

Definition 9 (Possibilistic Knowledge Base). A possibilistic knowledge base is a finite set $\{(\varphi_1, \alpha_1), \ldots, (\varphi_m, \alpha_m)\}$ of pairs of the form (φ_i, α_i), with $\varphi_i \in L_\Phi$ and $\alpha_i \in]0,1]$. It encodes a possibility distribution, namely the least specific possibility distribution satisfying the constraints $N(\varphi_i) \geq \alpha_i$.

Note that $N(\varphi)$ is an abbreviation for $N(\llbracket\varphi\rrbracket)$ or thus $N(\{\nu \in Int(\Phi) \mid \nu \models \varphi\})$. When it is clear from the context, we abbreviate *possibilistic knowledge base* to *knowledge base*.

Example 3. Recall the context of Example 2. To encode the situation in which a nation found evidence that at least one of its neighbouring nations is weaponized, it adds $(w_1 \vee w_2, 1)$ to its knowledge base. The corresponding constraint $N(w_1 \vee w_2) \geq 1$ implies that the world in which nation w_1 and nation w_2 are not weaponized is considered impossible. If the nation is rather certain that both of its neighbouring nations are weaponized, it adds $(w_1 \wedge w_2, 0.6)$ to its knowledge base. This has the effect that every world in which the neighbours are not both armed has a possibility degree of at most 0.4. It is easy to see that the possibility distribution encoded by these formulas is π as defined in Example 2.

The possibility distribution $\pi_{\mathcal{K}}$ encoded by a knowledge base $\mathcal{K}$ is well-defined because there is a unique least specific possibility distribution which satisfies the constraints of $\mathcal{K}$ [3]. It is given by

$$\pi_{\mathcal{K}}(\nu) = \min\{1 - \alpha \mid (\phi, \alpha) \in \mathcal{K} \text{ and } \nu \not\models \phi\}$$

The following inference rules are associated with possibilistic logic:

- $(\neg p \vee q, \alpha); (p \vee r, \beta) \vdash (q \vee r, \min(\alpha, \beta))$ (resolution rule),
- if p entails q classically, then $(p, \alpha) \vdash (q, \alpha)$ (formula weakening),
- for $\beta \leq \alpha$, $(p, \alpha) \vdash (p, \beta)$ (weight weakening),
- $(p, \alpha); (p, \beta) \vdash (p, \max(\alpha, \beta))$ (weight fusion).

The axioms consist of all propositional axioms with weight 1. These inference rules and axioms are sound and complete in the following sense [3]: it holds that $\mathcal{K} \vdash (\varphi, \alpha)$ iff $N(\varphi) \geq \alpha$ for the necessity measure N induced by $\pi_{\mathcal{K}}$. Another useful property is $\mathcal{K} \vdash (\varphi, \alpha)$ iff $\mathcal{K}_\alpha \vdash \varphi$ (in the classical sense) [19], with $\mathcal{K}_\alpha = \{\varphi \mid (\varphi, \beta) \in \mathcal{K}, \beta \geq \alpha\}$ the α-cut of $\mathcal{K}$.

3.2.3. Generalized Possibilistic Logic

Generalized possibilistic logic (GPL) is a recent extension of possibilistic logic which has been introduced to model incomplete knowledge about the beliefs of another agent [4]. Whereas a possibilistic knowledge base corresponds to a single possibility distribution, a knowledge base in GPL corresponds to a set of possibility distributions. Each of these possibility distributions corresponds to a possible model of the beliefs of the other agent. Syntactically, the language of GPL uses propositional combinations of modal formulas of the form $\mathbf{N}_\lambda(\alpha)$, where $\mathbf{N}_\lambda(\alpha)$ intuitively means that the necessity of α is at least λ. A possibilistic knowledge base intuitively corresponds to conjunctions of such modal formulas. In GPL, on the other hand, we can express formulas such as $\mathbf{N}_1(\alpha_1) \vee \mathbf{N}_1(\alpha_2)$, which encodes the fact that we know that the other agent is either certain of α_1 or certain of α_2, but we are ignorant about which of these two possibilities is the case.

Example 4. Suppose a specific nation is reacting to the fact that one of its two neighboring nations is weaponized. While the considered nation knows which of its two neighbours is weaponized, we do not possess this knowledge, and hence we cannot precisely model the beliefs of that nation. However, we can encode the knowledge $\mathbf{N}_1(w_1) \vee \mathbf{N}_1(w_2)$, which states that we know that the nation either knows that w_1 is the case or that they know that w_2 is the case.

Formally the language of GPL is defined as follows:

- For each formula α from L_Φ (with L_Φ defined as before) and each certainty degree from $\{\frac{1}{k}, \frac{2}{k}, ..., 1\}$ we have that $\mathbf{N}_\Lambda(\alpha)$ is a GPL formula. GPL formulas of this form are also called meta-atoms.
- If γ_1 and γ_2 are GPL formulas, then $\neg\gamma_1$ and $\gamma_1 \wedge \gamma_2$ are also GPL formulas.

Interpretations in GPL are normalized possibility distributions π, whose possibility degrees are taken from $\{0, \frac{1}{k}, \frac{2}{k}, ..., 1\}$. Such a GPL interpretation π satisfies the formula $\mathbf{N}_\lambda(\alpha)$ iff $N(\alpha) \geq \lambda$, with N the necessity measure corresponding to π. The notion of satisfaction is then extended to (sets of) propositional combinations of meta-atoms in the usual way. We define a GPL knowledge base as a finite set of GPL formulas. An interpretation that a satisfies all formulas of a GPL knowledge base $\mathcal{K}$ is called a model of $\mathcal{K}$. The set of all models of $\mathcal{K}$ will be denoted as $Mod(\mathcal{K})$.

Note that the semantics of GPL are defined in terms of lower bounds on necessity measures. However, because only finitely many certainty degrees are used, we can easily express lower bounds on possibility measures as well. In particular, for $\alpha \in L_\Phi$, we have that $\Pi(\alpha) \geq \lambda$ iff $\Pi(\alpha) > \lambda - \frac{1}{k}$, which is equivalent to $N(\neg\alpha) < 1 - \lambda + \frac{1}{k}$. This means that syntactically we can express the constraint $\Pi(\alpha) \geq \lambda$ using the GPL formula $\neg\mathbf{N}_{1-\lambda+\frac{1}{k}}(\neg\alpha)$. For the ease of presentation, we will use the abbreviation $inv(\lambda) = 1 - \lambda + \frac{1}{k}$ and:

$$\mathbf{\Pi}_\lambda(\alpha) \equiv \neg\mathbf{N}_{inv(\lambda)}(\neg\alpha)$$

If the set of atoms in Φ is finite, constraints of the form $\Delta(\alpha) \geq \lambda$ can also be expressed in GPL. In particular, we introduce the following notation:

$$\mathbf{\Delta}_\lambda(\alpha) \equiv \bigwedge_{\nu \in [\![\alpha]\!]} \mathbf{\Pi}_\lambda(\nu)$$

The guaranteed possibility measure is closely related to the notion of 'only knowing' from non-monotonic reasoning [20]. For example, a formula such as $\mathbf{N}_1(\alpha) \wedge \Delta_1(\alpha)$ states that the agent knows that α is true but nothing more than that.

Example 5. In the running example, $\mathbf{N}_1(w_1 \vee w_2) \wedge \Delta_1(w_1 \vee w_2)$ means that the considered nation knows that one of its neighbors is weaponized but that it does not know which one it is. This can be contrasted with $\mathbf{N}_1(w_1) \vee \mathbf{N}1(w_2)$ which states that the nation knows which of its neighbors is weaponized, and with $\mathbf{N}_1(w_1 \vee w_2)$, which leaves in the middle whether the considered nation knows which one of its neighbors is weaponized.

The use of GPL formulas for expressing limitations on the knowledge of others is explored in more detail in [4].

Finally note that the following links between the syntax and semantics of GPL can be easily verified:

$$(\mathcal{K} \models \mathbf{N}_\lambda(\alpha)) \equiv \big(\forall \pi \in \mathit{Mod}(\mathcal{K}), \forall \nu \in \mathcal{W} : (\nu \not\models \alpha) \Rightarrow \pi(\nu) \leq 1 - \lambda\big)$$
$$(\mathcal{K} \models \mathbf{\Pi}_\lambda(\alpha)) \equiv \big(\forall \pi \in \mathit{Mod}(\mathcal{K}), \exists \nu \in \mathcal{W} : (\nu \models \alpha) \wedge \pi(\nu) \geq \lambda\big)$$
$$(\mathcal{K} \models \mathbf{\Delta}_\lambda(\alpha)) \equiv \big(\forall \pi \in \mathit{Mod}(\mathcal{K}), \forall \nu \in \mathcal{W} : (\nu \models \alpha) \Rightarrow \pi(\nu) \geq \lambda\big)$$

4. Using Possibilistic Logic for Encoding Uncertain Boolean Games

In this section, we show how Boolean games can be naturally extended to model situations in which agents are uncertain about other agents' goals. We first use uncertainty measures from possibility theory to semantically define (solution concepts to) Boolean games with incomplete information in Section 4.1. Then in Section 4.2 we present a syntactic characterization of these semantics. We prove that the semantic and the syntactic approach are equivalent in Section 4.3. Finally, we characterize the computational complexity of the related solution concepts in Section 4.4.

Throughout this paper, we consider Boolean games with prioritized goal bases (see Definition 1). Recall that in these games, an agent is most eager to achieve the goal with the highest priority. If this goal cannot be achieved, the agent will settle for the goal with the second-highest priority, etc.

Example 6. Bob and Alice are going out: they can attend a sports game or go to the theatre. Alice – agent 1 – controls action variable a, and Bob – agent 2 – controls b. Setting their action variable to true corresponds to attending a sports game; setting it to false corresponds to going to the theatre. Bob and Alice's first priority is to go out together. If they do not go out together, Bob prefers a sports game, whereas Alice prefers the theatre. This can be represented as a Boolean game with priorities, where Alice's and Bob's goal bases are:

$$\Gamma_1 = \{a \leftrightarrow b; \neg a\}, \quad \Gamma_2 = \{a \leftrightarrow b; b\}$$

Both agents have utility 2 in the PNEs $\{a, b\}$ and $\emptyset$, which respectively correspond to attending a sports game together and going to the theatre together.

Our aim in this section is to propose an extension to the Boolean game framework in which agents can be uncertain about other agents' goals. An important concern is that the resulting framework should still enable a compact and intuitive representation of games, as this is the main strength of Boolean games. In Section 4.2 we will provide a compact characterization of the semantics proposed in Section 4.1. Using this extended Boolean game framework, we aim to determine rational behaviour for agents which are uncertain about the other agents' goals. We illustrate this with the following example.

Example 7. Consider again the scenario of Example 6, but now assume that Bob and Alice are not fully aware of each other's goals. For instance, if Bob knows Alice's goal, but Alice thinks that Bob does not want to join her to the theatre, then, based on their beliefs, attending a sports game together is a 'better' solution than going to the theatre together. Indeed, Alice believes that Bob will not agree to go to the theatre together (or might have an incentive to leave if he would go), but they both believe that the other will agree to attend a sports game together.

The results presented in this section can easily be generalized to accommodate for partially ordered preference relations. However, as modelling preferences is not the focus here, we prefer the simpler setting of Definition 1, for clarity.

For the ease of presentation, we will impose an additional restriction on the kind of goal bases that are considered in our framework for Boolean games with incomplete information. In particular, we will consider Boolean games $G = (N, (\Phi_i)_{i\in N}, \delta, (\Gamma_i)_{i\in N})$ satisfying the conditions of Definition 1, in which each goal base Γ_i belongs to the following set:.

$$\mathcal{G} = \{\{\gamma^1; \ldots; \gamma^p\} \mid (\gamma^1 \wedge \delta \not\models \bot) \wedge \forall k \in \{1, \ldots, p\} : \gamma^k \in L_\Phi^{CNF} \wedge (k \neq p \Rightarrow \gamma^k \models \gamma^{k+1})\} \tag{1}$$

with L_Φ^{CNF} all formulas of L_Φ in conjunctive normal form. Note that we can make this restriction without loss of generality as any goal base $\{\gamma^1; \ldots; \gamma^p\}$ violating the condition $\gamma^k \models \gamma^{k+1}$ can be transformed into a semantically equivalent[2] goal base which does satisfy the restriction (1), namely $\{\gamma^1; \gamma^1 \vee \gamma^2; \ldots; \bigvee_{m=1}^{p} \gamma^m\}$. Moreover, the set of possible goal bases is independent of the agent, i.e. every agent has the same set $\mathcal{G}$ of possible goal bases.

4.1. Semantics of Possibilistic Games

In the considered setting, the set of agents N, the corresponding partitions of action variables Φ_i, and the global constraint δ are known to all the agents. However, we assume that agents are uncertain about the goals of the other agents. We can formalise this by considering for each agent i a possibility distribution π_i on the set of possible Boolean games, where this latter set is defined as follows:

$$\mathcal{BG}(N, \Phi_1, \ldots, \Phi_n, \delta) = \{(N, (\Phi_i)_{i\in N}, \delta, (\Gamma_i)_{i\in N}) \mid \forall i \in N : \Gamma_i \in \mathcal{G}\}$$

When the set N of agents, the action variables $\Phi_1, \ldots, \Phi_n$ and the constraint δ are clear from the context, we abbreviate $\mathcal{BG}(N, \Phi_1, \ldots, \Phi_n, \delta)$ to $\mathcal{BG}$. The knowledge of an agent i about the goals of the other agents can then be captured by a possibility distribution π_i over $\mathcal{BG}$, encoding agent i's beliefs about what is the actual game being played. Note that this possibility distribution π_i is different for each agent.

Example 8. Recall the scenario of Example 6. Suppose Bob has perfect knowledge of Alice's preferences, then $\pi_2 : \mathcal{BG} \to \{0, 1\}$ maps every Boolean game to 0, except the Boolean

[2] Here equivalent means that they induce the same utility function.

games with the preference orderings of Example 6, i.e. the actual game being played is the only one considered possible by Bob up to logical equivalence. Suppose Alice is certain that Bob wants to attend a sports game together, or attend the game on his own if attending it together is not possible. Then $\pi_1 : \mathcal{BG} \to \{0,1\}$ maps all Boolean games to 0, except those with the following preference orderings and corresponding payoff matrix:

$$\{a,b\} =_1 \emptyset >_1 \{b\} >_1 \{a\}$$
$$\{a,b\} >_2 \{b\} >_2 \emptyset =_2 \{a\}$$

Bob \ Alice	a	$\neg a$
b	$(1,1)$	$(0.5,0.5)$
$\neg b$	$(0,0)$	$(0,1)$

Our first aim is to determine to which degree a specific outcome ν is necessarily or possibly a PNE according to agent i. Intuitively, it is possible to degree λ that an outcome ν is a PNE according to agent i iff there exists a Boolean game $G \in \mathcal{BG}$ such that ν is a PNE in G and such that agent i considers it possible to degree λ that G is the real game being played, i.e.

$$\Pi_i(\{G \in \mathcal{BG} \mid \nu \text{ is a PNE in } G\}) = \lambda$$

Similarly, it is certain to degree λ that an outcome ν is a PNE according to agent i iff for every $G \in \mathcal{BG}$ such that ν is not a PNE, it holds that i considers it possible to degree at most $1-\lambda$ that G is the real game being played, i.e.

$$N_i(\{G \in \mathcal{BG} \mid \nu \text{ is a PNE in } G\}) = \lambda$$

Using the aforementioned degrees, we can define measures which offer a way to distinguish between multiple equilibria, motivated by Schellings' notion of focal points [21]. An equilibrium is a focal point if, for some reason other than its utility, it stands out from the other equilibria. In our case, the reason can be that agents have a higher certainty that the outcome is actually a PNE. This motivation is similar as the one for verifiability in [8]: only certain PNEs carry sufficient information such that the agents can tell that an equilibrium has been played. Such a property provides an argument to play such an equilibrium, instead of the alternatives. Note that there might not exist an outcome which every agent believes is necessarily a PNE, even when the (unknown) game being played has one or more PNEs. In such cases, the degree to which various outcomes are possibly a PNE could be used to guide decisions.

When we consider PNEs from an epistemic game theory point of view, a PNE is a sensible solution concept in games with complete information, as it corresponds to the fact that it is common knowledge that every agent is acting individually rationally and thus playing a best response. In other words: every agent knows that every agent plays a best response, and thus has no incentive to individually deviate. Moreover, every agent knows that every agent knows that every agent plays a best response etc. When knowledge is incomplete, this property is no longer valid. It makes sense to consider weakened properties, such as: every agent believes that every agent is playing rationally. Note that to obtain common belief, one should also model the beliefs of the other agents, and the beliefs of the beliefs of the other agents etc. Such hierarchical beliefs have been studied for both probabilistic and Boolean models of agents' beliefs. A standard technique for modelling such beliefs, proposed in [22, 23, 24], is to consider a notion of types. The main idea is to assign to each agent a type, which determines its beliefs about the utilities and types of other players. This allows us to express hierarchical beliefs without explicitly construction a hierarchical model. Among others, this has the advantage that infinite belief hierarchies can be encoded in a finite way. While a similar approach could be pursued in our possibilistic setting, for simplicity we will focus on agents' beliefs about the goal bases of other agents.

Definition 10. Given the possibility measures Π_i for every agent i, the degree to which all agents find it possible that the outcome ν is a PNE is

$$poss(\nu) = \min_{i \in N} \Pi_i(\{G \in \mathcal{BG} \,|\, \nu \text{ is a PNE in } G\})$$

Similarly, given the necessity measures N_i for every agent i, the degree to which all agents believe it necessary that ν is a PNE is defined as

$$nec(\nu) = \min_{i \in N} N_i(\{G \in \mathcal{BG} \,|\, \nu \text{ is a PNE in } G\})$$

4.2. Syntactic Characterization

While the two possibilistic solution concepts from Section 4.1 are useful, the formulation in Definition 10 cannot directly be used in applications, since the number of Boolean games in $\mathcal{BG}$ is double-exponential. In this section, we present a syntactic counterpart which will allow for a more compact representation of the agents' knowledge about the game being played, as well as a more efficient implementation.

Definition 11 (Goal-Knowledge Base). A goal knowledge base (goal-KB) $\mathcal{K}_i^j$ of agent i w.r.t. agent j contains formulas of the form $(\varphi \rightarrow g_j^k, \lambda)$ or $(\varphi \leftarrow g_j^k, \lambda)$, where $1 \leq k \leq p$, $\varphi \in L_\Phi$, $\lambda \in]0,1]$ and g_j^k is a new atom, encoding agent j's goal of priority k. We further assume that $\mathcal{K}_i^j$ contains $\{(g_j^k \rightarrow g_j^{k+1}, 1) \,|\, 1 \leq k \leq p-1\}$. Finally, we require that a goal-KB $\mathcal{K}_i^j$ satisfies the following criteria, which we will refer to as *goal-consistency*:

- For every $\varphi, \psi \in L_\Phi$ such that $(\varphi \rightarrow g_j^k, \lambda) \in \mathcal{K}_i^j$ and $(\psi \leftarrow g_j^k, \mu) \in \mathcal{K}_i^j$, it holds that $\varphi \models \psi$.
- For all $(\psi_1 \leftarrow g_j^k, \lambda_1), ..., (\psi_r \leftarrow g_j^k, \lambda_r) \in \mathcal{K}_i^j$ it holds that $\psi_1 \wedge ... \wedge \psi_r \wedge \delta \not\models \bot$. Furthermore if $\psi_1 \wedge ... \wedge \psi_r \wedge \delta$ has a single model ν, we assume that $\mathcal{K}_i^j$ contains a formula $(\bigwedge\{l \,|\, l \in \nu\} \rightarrow g_j^k, \lambda)$ with $\lambda \geq \min(\lambda_1, ..., \lambda_r)$.

We will sometimes also encode $\mathcal{K}_i^j$ using formulas of the form $(\varphi \leftrightarrow g_j^k, \lambda)$ as an abbreviation for the two formulas $(\varphi \rightarrow g_j^k, \lambda)$ and $(\varphi \leftarrow g_j^k, \lambda)$. Furthermore, for the ease of presentation we do not normally mention the formulas the formulas $\{(g_j^k \rightarrow g_j^{k+1}, 1) \,|\, 1 \leq k \leq p-1\}$ in examples, as these formulas belong to $\mathcal{K}_i^j$ by definition.

A goal-KB $\mathcal{K}_i^j$ captures the knowledge of agent i about the goal base of agent j. These formulas express that, if agent j's utility is at least $\frac{p+1-k}{p}$, it is at least $\frac{p-k}{p}$. Furthermore, the information that we like to express in $\mathcal{K}_i^j$ consists of necessary and/or sufficient conditions for the utility of agent j. For instance, agent i might believe with certainty λ that φ is a sufficient condition for satisfying the goal with priority k, i.e. for achieving a utility of at least $\frac{p+1-k}{p}$. This is encoded as $(\varphi \rightarrow g_j^k, \lambda) \in \mathcal{K}_i^j$. Similarly, agent i might believe with certainty λ that φ is a necessary condition for achieving the goal with priority k, i.e. $(\varphi \leftarrow g_j^k, \lambda) \in \mathcal{K}_i^j$. These formulas can be combined as $(\varphi \leftrightarrow g_j^k, \lambda) \in \mathcal{K}_i^j$. Note how adding the atoms g_j^k to the language allows us to explicitly encode what an agent knows about the goals of another agent. This is inspired by the approach from [25] for merging conflicting sources, where similarly additional atoms are introduced to encode knowledge about the unknown meaning of vague properties, in the form of necessary and sufficient conditions.

Example 10. Recall the scenario of Example 6. Suppose Bob has a good idea of what Alice's goal base looks like: $\mathcal{K}_2^1 = \{((a \leftrightarrow b) \leftrightarrow g_1^1, 0.9), (((a \leftrightarrow b) \vee \neg a) \leftrightarrow g_1^2, 0.6)\}$. He is very certain that Alice's first priority is to go out together and rather certain that she prefers the theatre in case they do not go out together. Although Alice is very certain that Bob will be pleased if they attend a sports game together, she is only a little certain that Bob would be just as pleased if they go to the theatre together. She knows Bob prefers to go to a sports game as a second priority. Her knowledge of Bob's goal base can be captured by $\mathcal{K}_1^2 = \{((a \wedge b) \rightarrow g_2^1, 0.8), ((\neg a \wedge \neg b) \rightarrow g_2^1, 0.3), (b \rightarrow g_2^2, 1)\}$.

It is natural to assume that $\mathcal{K}_i^i$ contains the formulas $(g_i^k \leftrightarrow \bigvee_{m=1}^k \gamma_i^m, 1)$ for all $k \in \{1, \ldots, p\}$, i.e. every agent knows its own goal base and the corresponding utility. However, this assumption is not necessary for the results in this section. By requiring goal-consistency in Definition 11, we ensure that the knowledge base $\mathcal{K}_i^j$ only encodes beliefs about the goal of agent j. Without this assumption, it could be possible to derive from $\mathcal{K}_i^j$ formulas of the form $\varphi \rightarrow \psi$ that are not implied yet by the constraints of the game, i.e. $\delta \not\models (\varphi \rightarrow \psi)$. Such formulas encode dependencies between the action variables of agents, which might be useful for modelling suspected collusion, but this will not be considered in this work. However, we do not demand that the beliefs of an agent are correct, i.e. we do not assume that each agent considers the actual game possible.

Definition 12 (Boolean Game with Incomplete Information). A Boolean game with incomplete information is a tuple $G = (N, (\Phi_i)_{i \in N}, \delta, (\Gamma_i)_{i \in N}, (\mathcal{K}_i)_{i \in N})$ with $N, (\Phi_i)_{i \in N}, \delta$ and $(\Gamma_i)_{i \in N}$ as in Definition 1 and $\mathcal{K}_i$ the set $\{\mathcal{K}_i^1, \ldots, \mathcal{K}_i^n\}$, where $\mathcal{K}_i^j$ is a goal-KB of i w.r.t. j (see Definition 11).

Let us now consider how to compute the necessity and possibility that agent j plays a best response in the outcome ν according to agent i. First recall that whenever we write an interpretation ν where a formula is expected, this should be interpreted as the conjunction of ν's literals, i.e. $\bigwedge\{l \mid l \in \nu\}$.

Agent j plays a best response in the outcome ν iff for every alternative strategy $\nu_j' \in \mathcal{S}_j(\nu_{-j})$ it holds that $u_j(\nu) \geq u_j(\nu_{-j}, \nu_j')$. Essentially this boils down to the fact that, for some $k \in \{0, \ldots, p\}$, $u_j(\nu) \geq \frac{k}{p}$ and $\forall \nu_j' \in \mathcal{S}_j(\nu_{-j}) : u_j(\nu_{-j}, \nu_j') \leq \frac{k}{p}$. Note that for $k = 0$, the first condition is always fulfilled. Similarly, for $k = p$, the second condition becomes trivial. Similarly, agent j does not play a best response in ν iff there exists a $\nu_j' \in \mathcal{S}_j(\nu_{-j})$

such that $u_j(\nu) < u_j(\nu_{-j}, \nu'_j)$. This means that, for all $k \in \{0, \dots, p\}$, $u_j(\nu) < \frac{k}{p}$ or $\exists \nu'_j \in \mathcal{S}_j(\nu_{-j}) : u_j(\nu_{-j}, \nu'_j) > \frac{k}{p}$. The possibility of agent j playing a best response is dual to the necessity of agent j playing no best response. These insights motivate the following definition.

Definition 13. Let $i, j \in N$ be two agents in a Boolean game with incomplete information G and let ν be an outcome of G. We denote $g_j^{p+1} = \top$ and $g_j^0 = \bot$ for every j. We say that j plays a best response in ν with necessity λ according to i, written $\mathrm{BR}_i^{nec}(j, \nu) = \lambda$, iff λ is the greatest value in $[0, 1]$ for which there exists some $k \in \{0, \dots, p\}$ such that the following two conditions are satisfied:

1. $\mathcal{K}_i^j \vdash \nu \to (g_j^{k+1}, \lambda)$
2. $\mathcal{K}_i^j \vdash ((\nu_{-j} \wedge \neg \nu_j) \to (\neg g_j^k \vee \neg \delta), \lambda)$

Let λ^* be the smallest value greater than $1 - \lambda$ which occurs in $\mathcal{K}_i^j$. Agent i believes it is possible to degree λ that agent j plays a best response in ν, written $\mathrm{BR}_i^{pos}(j, \nu) = \lambda$, iff λ is the greatest value in $]0, 1]$ for which there exists some $k \in \{0, ..., p\}$ such that the following two conditions are satisfied:

1. $\mathcal{K}_i^j \nvdash (\nu \to \neg g_j^{k+1}, \lambda^*)$
2. $\forall \nu'_j \in \mathcal{S}_j(\nu_{-j}) : \mathcal{K}_i^j \nvdash (\nu_{-j} \wedge \nu'_j \to g_j^k, \lambda^*)$

If no such λ exists, then $\mathrm{BR}_i^{pos}(j, \nu) = 0$.

Importantly, the syntax in Definition 13 allows us to express the certainty or possibililty that an agent plays a best response, from the point of view of another agent. This forms an important base from which to define interesting solution concepts or measures in Boolean games with incomplete information. In this paper, we introduce the following measures that respectively reflect to what degree all agents believe it is necessary and possible that ν is a PNE. In other words, the degree to which all agents believe it is necessary or possible that all other agents will not have the incentive to deviate from the outcome.

Definition 14. Let G be a Boolean game with incomplete information. For every outcome $\nu \in \mathcal{S}$, we define the degree $PNE^{nec}(\nu)$ to which ν is necessarily a PNE and the degree $PNE^{pos}(\nu)$ to which ν is possibly a PNE as follows:

$$PNE^{nec}(\nu) = \min_{i \in N} \min_{j \in N} \mathrm{BR}_i^{nec}(j, \nu), \quad PNE^{pos}(\nu) = \min_{i \in N} \min_{j \in N} \mathrm{BR}_i^{pos}(j, \nu)$$

In Section 4.3 below, we will show that these notions indeed correspond to the notions of necessary and possible PNE that were defined semantically in Section 4.1.

If we assume that all agents know their own goal, then $\mathrm{BR}_i^{nec}(i, \nu) = \mathrm{BR}_i^{pos}(i, \nu) = 0$ if ν is not a PNE. Consequently, if ν is not a PNE, then we have $PNE^{nec}(\nu) = PNE^{pos}(\nu) = 0$. Note that the measures from Definition 14 induce a total ordering on $\mathcal{S}$, so there always exists a $\nu \in \mathcal{S}$ such that PNE^{nec} or PNE^{pos} is maximal.

Example 10 (continued). Let G be the Boolean game with incomplete information, defined by the aforementioned goal-KBs, and assume that Bob and Alice know their own goals. It can be computed that:

	$\{\neg a, \neg b\}$	$\{a, \neg b\}$	$\{\neg a, b\}$	$\{a, b\}$
$\min_{j \in N} \mathrm{BR}_1^{nec}(j, .)$	0.3	0	0	0.8
$\min_{j \in N} \mathrm{BR}_2^{nec}(j, .)$	0.9	0	0	0.9
$PNE^{nec}(.)$	0.3	0	0	0.8

The outcome $\{a, b\}$ has the highest value for PNE^{nec}. Note that if Bob had the 'dual' beliefs of Alice, i.e. $\mathcal{K}_2^1 = \{((\neg a \wedge \neg b) \rightarrow g_1^1, 0.8), ((a \wedge b) \rightarrow g_1^1, 0.3), (\neg a \rightarrow g_1^2, 1)\}$, then $\{\neg a, \neg b\}$ and $\{a, b\}$ both would have had value 0.3 for PNE^{nec}.

In [26] we showed that many solution concepts for Boolean games can be found by using a reduction to answer set programming. The concepts in this section, such as PNE^{nec}, can be computed using a straightforward generalization of these ideas.

4.3. Soundness and Completeness

In this section, we show that the solution concepts for Boolean games with incomplete information that were introduced in Section 4.2 indeed correspond to their semantic counterparts from Section 4.1.

With a given goal base $\Gamma_j = \{\gamma_j^1; \ldots; \gamma_j^p\} \in \mathcal{G}$ we can associate a classical knowledge base $T_j = \{\gamma_j^k \leftrightarrow g_j^k \mid k \in \{1, \ldots, p\}\}$, which is simply encoding that each goal g_j^k is defined as in the goal base Γ_j. Using this formulation of a goal base as a logical theory, we can now associate $\mathcal{K}_i^j$ with a possibility distribution π_i^j on $\mathcal{G}$ in the following natural way, inspired by [27], with $\max \emptyset = 0$:

$$\pi_i^j(\Gamma_j) = 1 - \max\{\alpha_l \mid (\varphi_l, \alpha_l) \in \mathcal{K}_i^j, T_j \not\models \varphi_l\} \tag{2}$$

Note that $T_j \not\models \varphi_l$ means that the goal base Γ_j does not satisfy the formula α. In other words, it means that if Γ_j is the actual goal base of agent j, then the belief α of agent i about this goal base is incorrect. The higher the certainty of the formulas violated by the theory associated with Γ_j, the lower the possibility that Γ_j is the real goal base of agent j according to agent i. Note that if we make the reasonable assumption that an agent knows its own goals, then π_i^i maps all elements of $\mathcal{G}$ to 0 except the real goal base[3] of i, which is mapped to 1. Given the Boolean game with incomplete information G, and using the possibility distributions on $\mathcal{G}$ for every j, we can define a possibility distribution π_i^G on the set of possible Boolean games $\mathcal{BG}$:

$$\pi_i^G(G') = \min_{j \in N} \pi_i^j(\Gamma_j^{G'})$$

[3]and any semantically equivalent goal base

with $\Gamma_j^{G'}$ the goal base of agent j in the Boolean game G'. This possibility distribution is the natural semantic counterpart of the Boolean game with incomplete information G. We now show that these possibility distributions π_i^G allow us to interpret the solution concepts that have been defined syntactically in Section 4.2 as instances of the semantically defined solution concepts from Section 4.1. This is formalized in the following proposition and corollary, which we prove further on. We use the notation $\mathrm{br}_j(\nu, \Gamma_j)$ for the propositional variable corresponding to the statement "agent j with goal base Γ_j plays a best response in ν".

Proposition 1. *For every $\nu \in \mathcal{S}$, $i, j \in N$ and $\lambda \in]0, 1]$, it holds that*

$$BR_i^{nec}(j, \nu) \geq \lambda \Leftrightarrow \big(\forall \Gamma_j \in \mathcal{G} : \neg \mathrm{br}_j(\nu, \Gamma_j) \Rightarrow \pi_i^j(\Gamma_j) \leq 1 - \lambda\big) \tag{3}$$

$$BR_i^{pos}(j, \nu) \geq \lambda \Leftrightarrow \big(\exists \Gamma_j \in \mathcal{G} : \mathrm{br}_j(\nu, \Gamma_j) \wedge \pi_i^j(\Gamma_j) \geq \lambda\big) \tag{4}$$

Corollary 1. *Let us denote the possibility and necessity measure associated with π_i^G as Π_i^G and N_i^G. For every $\nu \in \mathcal{S}$ it holds that*

$$N_i^G(\{G' \in \mathcal{BG} \,|\, \nu \text{ is a PNE in } G'\}) = \min_{j \in N} BR_i^{nec}(j, \nu) \tag{5}$$

$$\Pi_i^G(\{G' \in \mathcal{BG} \,|\, \nu \text{ is a PNE in } G'\}) = \min_{j \in N} BR_i^{pos}(j, \nu) \tag{6}$$

Consequently, it holds that:

$$nec_G(\{G' \in \mathcal{BG} \,|\, \nu \text{ is a PNE in } G'\}) = PNE^{nec}(\nu)$$
$$poss_G(\{G' \in \mathcal{BG} \,|\, \nu \text{ is a PNE in } G'\}) = PNE^{pos}(\nu)$$

Before we prove Proposition 1 and Corollary 1, a lemma is stated which deals with the construction of specific goal bases in $\mathcal{G}$, given specific knowledge bases encoding information about these goal bases.

Lemma 1. *Given a goal-KB $\mathcal{K}_i^j$, there exists a goal base $\Gamma_j \in \mathcal{G}$ such that $\pi_i^j(\Gamma_j) = 1$.*

Proof. We show that the condition $\pi_i^j(\Gamma_j) = 1$ is satisfied for the goal base $\Gamma_j = (\gamma_j^1; \ldots; \gamma_j^p)$, with γ_j^k a formula in CNF which is equivalent to $\bigwedge\{\varphi \,|\, \varphi \in L_\Phi, \exists \lambda > 0 : \mathcal{K}_i^j \vdash (g_j^k \rightarrow \varphi, \lambda)\}$. Indeed, first note that due to the fact that $\mathcal{K}_i^j$ is goal consistent, we have $\gamma_j^k \wedge \delta \not\models \bot$. We also clearly have $\gamma_j^k \models \gamma_j^{k+1}$ for $k < p$, hence it holds that $\Gamma_j \in \mathcal{G}$. To prove $\pi_i(\Gamma_j) = 1$ we have to show $T_j \models \phi$ for each $(\phi, \alpha) \in \mathcal{K}_i^j$. By construction we already have that $T_j \models \phi$ for each ϕ of the form $g_j^k \rightarrow \varphi$ or $g_j^k \rightarrow g_j^{k+1}$. Since $\mathcal{K}_i^j$, by definition of goal KB, is goal consistent, it also follows that $T_j \models \phi$ for all ϕ of the form $\varphi \rightarrow g_j^k$. □

Note that the construction of Γ_j relies on the (constrained) syntax of the formulas in $\mathcal{K}_i^j$. We now prove Proposition 1.

Proof of Proposition 1. $\boxed{\Rightarrow \text{ of } (3)}$ Suppose $\mathrm{BR}_i^{nec}(j, \nu) \geq \lambda$ and let $\Gamma_j \in \mathcal{G}$. We show by contraposition that $\pi_i^j(\Gamma_j) > 1 - \lambda$ implies $\mathrm{br}_j(\nu, \Gamma_j)$. From $\pi_i^j(\Gamma_j) > 1 - \lambda$ and (2), we find that $\forall(\varphi_l, \alpha_l) \in \mathcal{K}_i^j : T_j \not\models \varphi_l \Rightarrow \alpha_l < \lambda$. By Definition 13, $\mathrm{BR}_i^{nec}(j, \nu) \geq \lambda$ implies that there exists a $k' \in \{0, \ldots, p\}$ such that $\mathcal{K}_i^j \vdash (\nu \to g_j^{k'+1}, \lambda)$ and $\mathcal{K}_i^j \vdash ((\nu_{-j} \wedge \neg\nu_j) \to (\neg g_j^{k'} \vee \neg\delta), \lambda)$. It follows that $T_j \models \nu \to g_j^{k'+1}$ and $T_j \models (\nu_{-j} \wedge \neg\nu_j) \to (\neg g_j^{k'} \vee \neg\delta)$. Consequently, by definition of T_j, if $k' \in \{1, \ldots, p-1\}$, it holds that $T_j \models \nu \to \gamma_j^{k'+1}$ and $T_j \models (\nu_{-j} \wedge \neg\nu_j) \to (\neg\gamma_j^{k'} \vee \neg\delta)$. This means that j indeed plays a best response in ν since $\gamma_j^{k'+1} \wedge \delta$ is satisfied in ν (noting that $\nu \models \delta$ by the definition of outcome) and for every alternative strategy of j, $\gamma_j^{k'} \wedge \delta$ is not satisfied.

$\boxed{\Leftarrow \text{ of } (3)}$ Suppose that $\mathrm{BR}_i^{nec}(j, \nu) < \lambda$, i.e. for every $k \in \{0, \ldots, p\}$ either $\mathcal{K}_i^j \nvdash (\nu \to g_j^{k+1}, \lambda)$ or $\mathcal{K}_i^j \nvdash ((\nu_{-j} \wedge \neg\nu_j) \to \neg g_j^k \vee \neg\delta, \lambda)$. Let k' be the greatest index for which $\mathcal{K}_i^j \nvdash (\nu \to g_j^{k'}, \lambda)$. Note that $k' \geq 1$ because of (i). Furthermore note that we then have $\mathcal{K}_i^j \nvdash (\nu \to g_j^k, \lambda)$ for all $k \leq k'$ and thus in particular for $k = 1$. We now construct a goal base $\Gamma_j = (\gamma_j^1; \ldots; \gamma_j^p)$ with γ_j^k defined as follows. For $k \leq k'$, γ_j^k is a formula in CNF equivalent to $\Phi_k \vee (\Psi_k \wedge \neg\nu)$ where:

$$\Phi_k = \bigvee\{\varphi \mid \varphi \in L_\Phi, \mathcal{K}_i^j \vdash (\varphi \to g_j^k, \lambda)\}$$
$$\Psi_k = \bigwedge\{\psi \mid \psi \in L_\Phi, \mathcal{K}_i^j \vdash (\psi \leftarrow g_j^k, \lambda)\}$$

For $k > k'$, γ_j^k is a formula in CNF equivalent to $\Phi_k \vee \Psi_k$.

We first show that $\Gamma_j \in \mathcal{G}$. Clearly, we have $\gamma_j^k \models \gamma_j^{k+1}$ for $k < p$. To see why $\gamma_j^1 \wedge \delta \not\models \bot$, note that $\gamma_j^1 = \Phi_1 \vee (\Psi_1 \wedge \neg\nu)$ and that $\Psi_1 \wedge \neg\nu \wedge \delta \models \bot$ would mean that ν is the only model of $\Psi_1 \wedge \delta$, which by definition of goal consistency would mean that $\mathcal{K}_i^j \vdash (\nu \to g_j^1, \lambda)$, which is a contradiction.

Next we show that $\pi_i^j(\Gamma_j) > 1 - \lambda$ by checking that for every formula $(\varphi, \alpha) \in \mathcal{K}_i^j$ with $\alpha \geq \lambda$, it holds that $T_j \models \varphi$. This is clear for formulas φ of the form $\phi \to g_j^k$ and for formulas of the form $g_j^k \to g_j^{k+1}$. The fact that formulas of the form $\phi \leftarrow g_j^k$ are satisfied can be seen by the fact that then $\phi \models \Psi_k$ by definition of goal consistency.

Finally we show that j does not play a best response in ν with the constructed Γ_j. First note that because $\mathcal{K}_i^j \nvdash (\nu \to g_j^k, \lambda)$ for $k \leq k'$ we know that $\nu \not\models \Phi$, and thus we have $\nu \not\models \gamma_j^k$ for any $k \leq k'$. Moreover, since $\mathcal{K}_i^j \vdash (\nu \to g_j^{k'+1}, \lambda)$ we must have $\mathcal{K}_i^j \nvdash ((\nu_{-j} \wedge \neg\nu_j) \to \neg g_j^{k'} \vee \neg\delta, \lambda)$. It follows that $\nu_{-j} \wedge \neg\nu_j \wedge \Psi_{k'} \wedge \delta$ is consistent, which means that there is some $\nu'_j \in \mathcal{S}_j(\nu_{-j})$ such that (ν_{-j}, ν'_j) satisfies $\gamma_j^{k'}$.

$\boxed{\Rightarrow \text{ of } (4)}$ Analogous to the proof of "$\Leftarrow$ of (3)".

$\boxed{\Leftarrow \text{ of } (4)}$ We prove directly that $\mathrm{BR}_i^{pos}(j, \nu) \geq \lambda$, i.e. $\exists k \in \{0, \ldots, p\}$ such that $\mathcal{K}_i^j \nvdash (\nu \to \neg g_j^{k+1}, \lambda^*)$ and $\forall \nu'_j \in \mathcal{S}_j(\nu_{-j}) : \mathcal{K}_i^j \nvdash ((\nu_{-j}, \nu'_j) \to g_j^k, \lambda^*)$. By assumption, there exists a Γ_j such that j plays a best response in ν and $\pi_i^j(\Gamma_j) \geq \lambda$. The former means that for some $k' \in \{0, \ldots, p\}$, $T_j \models \nu \to \gamma_{k'+1}^j$ and $\forall \nu'_j$ such that (ν_{-j}, ν'_j) satisfies δ it holds that $T_j \models (\nu_{-j}, \nu'_j) \to \neg\gamma_{k'}^j$. Since $T_j \models \gamma_l^j \leftrightarrow g_j^l$, it then holds that $T_j \models \nu \to g_j^{k'+1}$. Since by definition $\nu \not\models \bot$, $T_j \not\models \bot$ and $T_j \not\models \neg\nu$, it follows that $T_j \not\models \nu \to \neg g_j^{k'+1}$. The assumption that $\pi_i^j(\Gamma_j) \geq \lambda$ implies that $\forall(\varphi_l, \alpha_l) \in \mathcal{K}_i^j : T_j \not\models \varphi_l \Rightarrow \alpha_l \leq 1 - \lambda$. It follows that

$\mathcal{K}_i^j \not\vdash (\nu \to \neg g_j^{k'+1}, \lambda^*)$. Analogously, we can prove that the fact that $T_j \models (\nu_{-j}, \nu'_j) \to \neg\gamma_{k'}^j$ holds for every ν'_j such that (ν_{-j}, ν'_j) satisfies δ implies that $\forall \nu'_j \in \mathcal{S}_j(\nu_{-j}) : \mathcal{K}_i^j \not\vdash ((\nu_{-j}, \nu'_j) \to g_j^{k'}, \lambda^*)$. □

We now prove Equation (5) from Corollary 1. The proof of (6) is analogous and the rest of Corollary 1 follows immediately by minimizing over all agents in (5) and (6).

Proof of (5). By definition, $\min_{j\in N} \mathrm{BR}_i^{nec}(j,\nu) \geq \lambda$ iff $\mathrm{BR}_i^{nec}(j,\nu) \geq \lambda$ for every $j \in N$. We proved (Proposition 1) that the latter is equivalent with $\forall \Gamma_j \in \mathcal{G} : j$ no best response in $\nu \Rightarrow \pi_i^j(\Gamma_j) \leq 1-\lambda$. We first prove that this implies that for all $G' \in \mathcal{BG}$ it holds that $\pi_i^G(G') \leq 1-\lambda$ if ν is not a PNE in G'. By definition, this means that $N_i^G(\{G' \in \mathcal{BG} \mid \nu$ is a PNE in $G'\}) \geq \lambda$. Take an arbitrary G' such that ν is not a PNE in G'. Then there exists some j who plays no best response in ν if its goal base is $\Gamma_j^{G'}$. By assumption, this implies $\pi_i^j(\Gamma_j^{G'}) \leq 1-\lambda$, which implies $\pi_i^G(G') \leq 1-\lambda$ by definition. We now prove the opposite direction. Take an arbitrary j and Γ_j such that j plays no best response in ν with the goal base of j equal to Γ_j. Using Lemma 1, we can construct a $G' \in \mathcal{BG}$ such that $\Gamma_j^{G'} = \Gamma_j$ and $\pi_{j'}^i(\Gamma_{j'}^{G'}) = 1$ for every $j' \neq j$. Obviously ν is not a PNE in G' since j plays no best response. By assumption and by definition of N_i^G, it holds that $\pi_i^G(G') \leq 1-\lambda$. Since $\pi_{j'}^i(\Gamma_{G'}^{j'}) = 1$ for every $j' \neq j$, it follows that $\pi_i^j(\Gamma_j) \leq 1-\lambda$. Due to Proposition 1, we proved that $\mathrm{BR}_i^{nec}(j,\nu) \geq \lambda$. Since j is arbitrary, it follows that $\min_{j\in N} \mathrm{BR}_i^{nec}(j,\nu) \geq \lambda$. □

Example 11. Recall the scenario of Example 8. We define the corresponding Boolean game with incomplete information G. Since Bob has perfect knowledge of Alice's preferences, his goal-KB can be modelled as $\mathcal{K}_2^1 = \mathcal{K}_1^1 = \{((a \leftrightarrow b) \leftrightarrow g_1^1, 1), (((a \leftrightarrow b) \vee \neg a) \leftrightarrow g_1^2, 1)\}$. Alice is certain that Bob wants to attend a sports game together, or attend the game on his own if attending it together is not possible. This can be captured by the goal-KB $\mathcal{K}_1^2 = \{((a \wedge b) \leftrightarrow g_2^1, 1), (b \leftrightarrow g_2^2, 1)\}$. It is easy to see that π_1^G and π_2^G correspond to the possibility distributions π_1 and π_2 described in Example 8. Despite Alice's incorrect beliefs, Bob and Alice are both certain that attending a sports game together is a PNE, since $nec_G(\{G' \in \mathcal{BG} \mid \{a,b\}$ is a PNE in $G'\}) = PNE^{nec}(\{a,b\}) = 1$. Contrary to Alice, Bob knows that going to the theatre together is a PNE as well.

An interesting question is how the agents' beliefs can influence the proposals they can make in e.g. bargaining protocols. Suppose for instance that in the above example, Alice wants to make Bob a suggestion. Based on her beliefs, it would be rational to suggest to attend a sports game together, as she believes that neither one of them would have an incentive to deviate from this outcome. Bob would then rationally agree, based on his beliefs, as he is also convinced that neither one of them would deviate. We will study a simple bargaining protocol in Section 6.1, although that protocol will be based on Boolean beliefs (i.e. the setting from Section 5). Taking into account weighted beliefs in such protocols remains a topic for future work, where for instance strength of belief could be related to the degree of risk-aversion of an agent. Another interesting question for future work is how the actions of other agents can be used to modify one's beliefs about that agent. For example, if Bob were to make a proposal, he can choose between two rational suggestions: attending a sports game together or going to the theatre together. If he did the latter, Alice would know that her beliefs are incorrect, if she assumes Bob makes rational suggestions. Other research possibilities lie in manipulating Boolean games with incomplete information through communication, for instance through announcements, as investigated for Boolean games with environment variables [7]. Another option is to extend the framework of Boolean games with incomplete information, allowing agents to also reason about the beliefs of other agents, although this is likely to lead to an increase in computational complexity.

4.4. Computational Complexity

In this section two natural decision problems associated with the proposed possibilistic Nash equilibria are investigated.

Proposition 2. *Let G be a Boolean game with incomplete information and $\lambda \in]0,1]$. The following decision problems are Σ_2^{P}-complete:*

1. *Does there exist an outcome ν with $PNE^{nec}(\nu) \geq \lambda$?*
2. *Does there exist an outcome ν with $PNE^{pos}(\nu) \geq \lambda$?*

Proof. Hardness of 1 and 2 Both problems are Σ_2^{P}-hard since they contain the Σ_2^{P}-complete problem to decide whether a Boolean game has a PNE as a special case. Indeed, when G is a Boolean game, we can construct a Boolean game with incomplete information in which all agents have complete knowledge of each other's goals. Then $PNE^{nec}(\nu)$ and $PNE^{pos}(\nu)$ coincide and take values in $\{0,1\}$, depending on whether ν is a PNE or not. Consequently, G has a PNE iff there exists a ν with $PNE^{nec}(\nu) = PNE^{pos}(\nu) \geq \lambda$.

Membership of 1 We can decide the problem by first guessing an outcome ν. Checking whether $PNE^{nec}(\nu) \geq \lambda$ means checking whether $\mathrm{BR}_i^{nec}(j,\nu) \geq \lambda$ for every $i,j \in N$. The latter involves checking possibilistic entailments, which can be done in constant time using an NP-oracle. Therefore, the decision problem is in Σ_2^{P}.

Membership of 2 We can decide the problem by first guessing an outcome ν. Checking whether $PNE^{pos}(\nu) \geq \lambda$ means checking whether $\mathrm{BR}_i^{pos}(j,\nu) \geq \lambda$ for every $i,j \in N$. To see that the latter can be reduced to checking a polynomial number of possibilistic entailments, we need to rewrite the condition that $\forall \nu_j' \in \mathcal{S}_j(\nu_{-j}) : \mathcal{K}_i^j \nvdash ((\nu_{-j}, \nu_j') \rightarrow g_j^k, \lambda^*)$. To this end, we define $\mathcal{K}^k$, for every $k \in \{1,\ldots,p\}$, as the knowledge base $\mathcal{K}_i^j$ in which all formulas defining necessary and/or sufficient conditions for g_j^k are preserved; all formulas with necessary conditions for g_j^l ($l \geq k$) are translated into necessary conditions for g_j^k by replacing $(\varphi \rightarrow g_j^l, \alpha)$ by $(\varphi \rightarrow g_j^k, \alpha)$; all formulas with sufficient conditions for g_j^l ($l \leq k$) are translated into sufficient conditions for g_j^k by replacing $(\varphi \leftarrow g_j^l, \alpha)$ by $(\varphi \leftarrow g_j^k, \alpha)$; all other formulas are removed. Then it holds

$$
\begin{aligned}
&\forall \nu_j' \in \mathcal{S}_j(\nu_{-j}) : \mathcal{K}_i^j \nvdash ((\nu_{-j}, \nu_j') \rightarrow g_j^k, \lambda^*) \\
&\Leftrightarrow \forall \nu_j' \in \mathcal{S}_j(\nu_{-j}) : \mathcal{K}^k \nvdash ((\nu_{-j}, \nu_j') \rightarrow g_j^k, \lambda^*) \\
&\Leftrightarrow \forall \nu_j' \in \mathcal{S}_j(\nu_{-j}) : \mathcal{K}_{1-\lambda}^k \nvdash ((\nu_{-j}, \nu_j') \rightarrow g_j^k)
\end{aligned}
$$

$$\Leftrightarrow \forall \nu'_j \in Int(\Phi_j) : \mathcal{K}^k_{1-\lambda} \nvdash ((\nu_{-j}, \nu'_j) \rightarrow (g^k_j \wedge \delta))$$
$$\Leftrightarrow \forall \nu'_j \in Int(\Phi_j) : \mathcal{K}^k_{1-\lambda} \text{ and } (\nu_{-j}, \nu'_j) \text{ and } \neg(g^k_j \wedge \delta) \text{ are consistent}$$
$$\Leftrightarrow \forall \nu'_j \in Int(\Phi_j) : \mathcal{K}^k_{1-\lambda} \text{ and } (\nu_{-j}, \nu'_j) \text{ and } (\neg g^k_j \vee \neg\delta) \text{ are consistent}$$
$$\Leftrightarrow \forall \nu'_j \in Int(\Phi_j) : \mathcal{K}'^k_{1-\lambda} \text{ and } \nu'_j \text{ and } (\neg g^k_j \vee \neg\delta') \text{ are consistent}$$

where $\mathcal{K}'^k_{1-\lambda}$ is obtained from $\mathcal{K}^k_{1-\lambda}$ by replacing each occurrence of $p \in \Phi \setminus \Phi_j$ by its truth value ($\top$ or $\bot$) in ν, and similar for δ'. The last condition is equivalent with $((\mathcal{K}'^k_{1-\lambda})_\top \wedge \neg\delta') \vee (\mathcal{K}'^k_{1-\lambda})_\bot$ being a tautology, where $(\mathcal{K}'^k_{1-\lambda})_\top$ is obtained from $\mathcal{K}'^k_{1-\lambda}$ by replacing each occurrence of g^k_j with $\top$ and similar for $(\mathcal{K}'^k_{1-\lambda})_\bot$. Checking whether the latter expression is indeed a tautology can be done with a SAT-solver, i.e. in constant time with an NP-oracle. Therefore, the decision problem is in Σ^{P}_2.

We can conclude that both problems are Σ^{P}_2-complete. $\square$

The result of Proposition 2 shows that the complexity for the introduced measures does not increase compared to PNEs of Boolean games, since deciding whether a Boolean game has a PNE is also Σ^{P}_2-complete. Moreover, given the experimental results reported in [26] for standard Boolean games, it seems plausible that a reduction to answer set programming would support an efficient computation of solutions for medium sized games.

5. Using GPL for Encoding Ignorance in Boolean Games

In this section, we introduce an alternative framework to represent incomplete information in Boolean games, using GPL to compactly encode what each agent knows about the preferences of each agent. In particular, we will consider GPL knowledge bases $\mathcal{K}^j_i$ to encode what agent i knows about the preferences of agent j. Each possibility distribution over Φ will correspond to a utility function, where those utility functions that i considers possible utility functions of agent j will correspond to the models of $\mathcal{K}^j_i$. Note that because i may be uncertain about the utility function of agent j, we need GPL (where knowledge bases may correspond to arbitrary sets of possibility distributions) rather than standard possibilistic logic (where knowledge bases are naturally associated with a particular possibility distribution).

Although Section 4 also uses possibilistic logic to model incomplete information in Boolean games, our motivation now differs: whereas the previous section uses possibilistic logic to encode graded beliefs about other agents' goals, we now want to use possibilistic logic to compactly describe agents' preferences. In contrast to the approach from Section 4, the weights associated with the formulas will now correspond to degrees of preference [28, 29] instead of degrees of certainty [3]. In particular the notion of uncertainty in this framework, i.e. the way we model incompleteness, is binary: an agent either considers a given utility function possible or impossible. Note that both the approach from Section 4 and the approach from this section model incomplete information w.r.t. the agents' preferences (namely partial knowledge about the goals of the agents) in the context of Boolean games. However, to avoid confusion when referring to either one of them, we use the term *Boolean* *game with incomplete information* for the framework of Section 4, and the term *Boolean game with incomplete preference-information* for the framework from this section.

Note that the approach presented in this section is clearly different from approaches that are aimed at modelling uncertainty in games, such as the Bayesian approaches that have been widely studied in the game theory literature (see e.g. [30]). Our use of GPL also differs from approaches such as CP-nets[4], which despite their similar aim to compactly model preferences, only capture a single preference structure and are thus less suitable for modelling incomplete information. To the best of our knowledge, this is the first research on using GPL to model incomplete information about other agents' preferences.

5.1. GPL Encoding of Incomplete Boolean Games

We now explain in detail how GPL can be used to model incomplete information about preferences in Boolean games with priorities. We use Example 1 as the running example of this section.

Definition 15 (BG with Incomplete Preference-Information). A Boolean game with incomplete preference-information is a tuple $G = (N,\ (\Phi_i)_{i\in N},\ \delta,\ (\Gamma_i)_{i\in N},\ (\mathcal{K}_i)_{i\in N})$ with $N, (\Phi_i)_{i\in N}, \delta$ and $(\Gamma_i)_{i\in N}$ as in Definition 1 and $\mathcal{K}_i = \{\mathcal{K}_i^1, \ldots, \mathcal{K}_i^n\}$, where $\mathcal{K}_i^j$ is a GPL knowledge base such that $\mathcal{K}_i^j \models \mathbf{N}_1(\delta)$ and $\mathit{Mod}(\mathcal{K}_i^i) = \{u_i\}$.

For every two agents i and j in N the GPL knowledge base $\mathcal{K}_i^j$ encodes what agent i knows about the preferences of j. Recall from Section 3.2.3 that $\mathit{Mod}(\mathcal{K})$ are the models of $\mathcal{K}$. These models are normalized possibility distributions over $\mathit{Int}(\Phi)$ which only take values from $\{0, \frac{1}{p}, ..., 1\}$, which means that we can view them as utility functions. To reinforce this view, throughout this section we will use the notation u to denote GPL interpretations, rather than the more common notation π. For a model u of $\mathcal{K}_i^j$ and an outcome ν, $u(\nu) = \frac{l}{p}$ then means that the utility of outcome ν for agent j is $\frac{l}{p}$, if we take u to be the actual utility function of j. Each model u of $\mathcal{K}_i^j$ thus implicitly encodes the prioritized goal base $\{\gamma_j^1, ..., \gamma_j^p\}$ where γ_j^l is equivalent to $\bigvee\{\nu \mid u(\nu) \geq p + 1 - l\}$ (i.e. $u(\nu) = \frac{l}{p}$ if the highest priority goal from j that is satisfied in ν is γ_j^{p+1-l}). In this way, the GPL knowledge bases $\mathcal{K}_i^1, \ldots, \mathcal{K}_i^n$ compactly encode the BGs with priorities that agent i considers possible representations of the actual game being played. Note that in contrast to the setting from Section 4, the link between the syntactic characterization from Definition 15 and the corresponding semantics in terms of sets of BGs with priorities is immediately clear.

The condition $\mathbf{N}_1(\delta) \in \mathcal{K}_i^j$ corresponds to the assumption that all agents are aware of the integrity constraint δ. Specifically, it expresses that agent i knows that agent j's utility is 0 for ν — the lowest possible payoff — if ν violates δ (i.e. if ν is not an outcome). Finally, the assumption that $\mathit{Mod}(\mathcal{K}_i^i)$ is a singleton $\{u_i\}$ in Definition 15 corresponds to the assumption that agent i knows its own utility. If the actual goal base of agent i is given by $(\gamma_i^1, ..., \gamma_j^p)$,

[4] CP = *ceteris paribus*

we can define $\mathcal{K}_i^i$ as follows:

$$\mathcal{K}_i^i = \{\mathbf{N}_{\frac{p-m+1}{p}}(\bigvee_{l=1}^{m} \gamma_i^l) \wedge \mathbf{\Delta}_{inv(\frac{m}{p})}(\gamma_i^m \wedge \delta) \,|\, m \in \{1, ..., p\}\} \cup \{\mathbf{N}_1(\delta)\}$$

Indeed, let $\nu \in Int(\Phi)$ and suppose that u is a model of $\mathcal{K}_i^i$. Suppose that $\nu \models \delta$ and let m be the highest priority goal satisfied in ν. Because $\mathcal{K}_i^i$ entails $\mathbf{\Delta}_{inv(\frac{m}{p})}(\gamma_i^m \wedge \delta)$ we have $u(\nu) \geq \frac{m}{p}$. If $m < p$ we furthermore have that $u(\nu) \leq \frac{m}{p}$ because $\mathcal{K}_i^i$ entails $\mathbf{N}_{\frac{p-m}{p}}(\bigvee_{l=1}^{m-1} \gamma_i^l)$ and ν is not a model of $\bigvee_{l=1}^{m-1} \gamma_i^l$. If $\nu \not\models \delta$, then $u(\nu) = 0$ because $\mathcal{K}_i^i$ contains the formula $\mathbf{N}_1(\delta)$. For every $\nu \in Int(\Phi)$ we thus have that the possibility degree $u(\nu)$ is completely determined by $\mathcal{K}_i^i$. In other words, $\mathcal{K}_i^i$ has a unique model u, which corresponds to the goal base $(\gamma_i^1, ..., \gamma_j^p)$.

GPL formulas can be used to encode knowledge about preferences in an intuitive way. For instance, $\mathcal{K}_i^j \models \mathbf{N}_\lambda(\alpha)$ means that agent i believes that, whenever an outcome does not satisfy α the utility of agent j can be at most $1-\lambda$. In other words, α is a necessary condition for agent j to reach utility higher than $1 - \lambda$. For example, in the context of Example 1, if Bob believes that Alice is unhappy without her dog, this can be encoded by $\mathbf{N}_1(d_A) \in \mathcal{K}_2^1$: whenever the dog is not with Alice, Alice's utility is 0. Similarly, $\mathcal{K} \models \mathbf{\Delta}_\lambda(\alpha)$ means that agent i believes that agent j's utility is at least λ whenever the outcome satisfies α. In other words, α is a sufficient condition for j to reach a utility of λ or more. For instance, in the context of Example 1, Bob can encode that Alice is at least partially happy when she is with the dog, regardless of whatever else happens, by adding $\mathbf{\Delta}_{\frac{1}{3}}(d_A)$ to his knowledge base $\mathcal{K}_2^1$. This states that Alice's utility is at least $\frac{1}{3}$ when the dog is with her. Finally, $\mathcal{K} \models \mathbf{\Pi}_\lambda(\alpha)$ encodes that agent i believes that there is some outcome that satisfies α in which agent j reaches utility λ. For instance, in the context of Example 1, when Bob's knowledge base $\mathcal{K}_2^1$ contains the formula $\mathbf{\Pi}_1(d_A \wedge (f_A \vee b_A))$, which encodes that Bob believes that Alice's first priority goal can be satisfied when she takes the dog to the beach or the forest.

Rather than encoding bounds on specific utility scores, in some applications it may be more natural to express *comparative preferences*, e.g. encoding that i believes that j prefers any/some outcome satisfying α over any/some outcome satisfying β. To conveniently express such knowledge, we will introduce some abbreviations for $\alpha, \beta \in L_\Phi$:

$$\begin{aligned} \beta \succeq \alpha &\equiv \bigvee_{m=1}^{p-1} (\neg\mathbf{\Pi}_{\frac{m+1}{p}}(\alpha) \wedge \mathbf{\Delta}_{\frac{m}{p}}(\beta)) \vee \neg\mathbf{\Pi}_{\frac{1}{p}}(\alpha) \vee \mathbf{\Delta}_1(\beta) \\ \beta \succ \alpha &\equiv \bigvee_{m=1}^{p} (\neg\mathbf{\Pi}_{\frac{m}{p}}(\alpha) \wedge \mathbf{\Delta}_{\frac{m}{p}}(\beta)) \end{aligned} \tag{7}$$

Intuitively, whenever $\mathcal{K}_i^j \models \beta \succeq \alpha$, agent i knows that the utility of agent j in any outcome that satisfies β is at least the utility of j in any outcome that satisfies α. Similarly, whenever $\mathcal{K}_i^j \models \beta \succ \alpha$, agent i knows that agent j strictly prefers any outcome in which β is true to any outcome in which α is true. Another useful abbreviation is:

$$\beta \succ_c \alpha \equiv \bigvee_{m=1}^{p} (\neg \mathbf{\Pi}_{\frac{m}{p}}(\alpha) \wedge \mathbf{\Pi}_{\frac{m}{p}}(\beta)) \tag{8}$$

Intuitively, whenever $\mathcal{K}_i^j \models \beta \succ_c \alpha$, agent i knows that there is some outcome satisfying β with a utility that is higher than that of any outcome satisfying β.

We illustrate the expressiveness of GPL for modelling preferences in the following example.

Example 12. Recall the context of Example 1 and suppose Alice knows that Bob's first priority goal can only be fulfilled without bringing the dog. This is encoded as $\mathbf{N}_{\frac{1}{3}}(\neg d_A) \in \mathcal{K}_1^2$. If Alice knows that Bob prefers going to the beach exclusively with her over going to the beach with her and the dog, this is encoded as $(b_A \wedge b_B \wedge \neg d_A) \succ (b_A \wedge b_B \wedge d_A) \in \mathcal{K}_1^2$. When Alice believes that Bob has at least utility $\frac{2}{3}$ when they both go to the beach, she can encode this as $\mathbf{\Delta}_{\frac{2}{3}}(b_A \wedge b_B)$. If Bob knows that Alice is unhappy without her dog, this is encoded as $\mathbf{N}_1(d_A) \in \mathcal{K}_2^1$. To encode that Alice is at least partially happy when she is with the dog, regardless of whatever else happens, Bob can add $\mathbf{\Delta}_{\frac{1}{3}}(d_A)$ to $\mathcal{K}_2^1$. Similarly, if Bob believes Alice's top priority is to go to the beach together with the dog, he can add $\mathbf{\Delta}_1(b_A \wedge b_B \wedge d_A)$ to his knowledge base.

Similar to Definition 13 and Definition 14, we define solution concepts which capture that agents believe that a certain outcome is a PNE.

Definition 16. Let $i, j \in N$ be two agents in a Boolean game with incomplete preference-information G and let ν be an outcome of G. We say that agent i knows that agent j plays a best response in ν iff it holds that $\mathcal{K}_i^j \models \nu \succeq \nu_{-j}$. We say that agent i considers it possible that agent j plays a best response in ν iff it holds that $\mathcal{K}_i^j \not\models \nu_{-j} \succ_c \nu$.

At the semantic level, agent i knows that agent j plays a best response in ν iff this is the case for every utility function of j considered possible by i. Similarly, agent i believes it is possible that agent j plays a best response in ν iff this is the case for at least one utility function of j considered possible by i.

Definition 17. Let G be a Boolean game with incomplete preference-information. For every outcome $\nu \in \mathcal{S}$, we say that ν is a known PNE iff every agent in the game knows that every other agent plays a best response in ν. Similarly, we say that ν is a possible PNE iff every agent in the game considers it possible that every other agent plays a best response in ν.

Example 13. Recall the context of Example 12, where Alice's and Bob's knowledge bases w.r.t. each other are:

$$\mathcal{K}_1^2 = \{\mathbf{N}_1(\delta), \mathbf{N}_{\frac{1}{3}}(\neg d_A), (b_A \wedge b_B \wedge \neg d_A) \succ (b_A \wedge b_B \wedge d_A), \mathbf{\Delta}_{\frac{2}{3}}(b_A \wedge b_B)\}$$
$$\mathcal{K}_2^1 = \{\mathbf{N}_1(\delta), \mathbf{N}_1(d_A), \mathbf{\Delta}_{\frac{1}{3}}(d_A), \mathbf{\Delta}_1(b_A \wedge b_B \wedge d_A)\}$$

The possible PNEs correspond to the true PNEs: Bob and Alice go to the forest with the dog or Bob and Alice go to the beach with the dog. The three of them going to the beach is the unique known PNE.

Note that we have not demanded that the knowledge of the agents is correct, i.e. the true utility function is not ruled out as a possibility by other agents. However, if we do make the assumption that all knowledge is correct, then it is easy to verify that any true PNE is also a possible PNE. Also note that believing that some outcome ν is a possible PNE when it is not a true PNE is impossible when we assume that the agents know their own utility: when ν is not a true PNE, the agent who does not play a best response does not consider it possible that ν is a PNE.

5.2. Computational Complexity

In this section, we investigate the computational complexity of known and possible PNEs.

Proposition 3. *Let G be a Boolean game with incomplete preference-information. The following decision problems are Σ_2^P-complete:*

1. *Does there exist a known PNE?*
2. *Does there exist a possible PNE?*

Proof. $\boxed{\text{Hardness}}$ The problems are Σ_2^{P}-hard since they contain the Σ_2^{P}-complete problem to decide whether a Boolean game has a PNE as a special case. Indeed, when G is a Boolean game, we can construct a Boolean game with incomplete preference-information G' in which all agents have complete knowledge of each other's goals. In that case, known and possible PNEs in G' coincide with PNEs in G, thus G has a PNE iff G' has a known or possible PNE.
$\boxed{\text{Membership}}$ We can decide whether a known PNE exists by first guessing an outcome ν. Checking whether ν is a known PNE means checking whether agent i knows that agent j plays a best response in ν for every $i, j \in N$. The latter involves checking possibilistic entailments in GPL. Since the formula $\nu \succeq \nu_{-j}$ contains the modal operator $\mathbf{\Delta}$, the computational complexity of deciding $\mathcal{K}_i^j \models \nu \succeq \nu_{-j}$ is in Θ_2^P [31], i.e. the problem can be solved in polynomial time on a deterministic Turing maching, using a logarithmic number of calls to an NP-oracle. Since Σ_2^{P} represents the problems which can be solved in polynomial time on a non-deterministic Turing machine with an NP-oracle, we can conclude that deciding whether a known PNE exists is in Σ_2^{P}. Similarly, we can decide whether a possible PNE exists by first guessing an outcome ν. The latter involves checking possibilistic entailments in GPL. Since the formula $\nu_{-j} \succ_c \nu$ only contains the modal operator $\mathbf{\Pi}$, the computational complexity of deciding $\mathcal{K}_i^j \not\models \nu_{-j} \succ_c \nu$ is in NP [31]. Deciding whether $\mathcal{K}_i^j \not\models \nu_{-j} \succ_c \nu$ holds can thus be done using an NP-oracle. Consequently, deciding whether a possible PNE exists is in Σ_2^{P}.
We can conclude that both problems are Σ_2^{P}-complete. □

Similarly as for Proposition 2, the result of Proposition 3 shows that considering incomplete information about the preferences of other agents does not lead to an increase in the computational complexity of the main decision problems.

6. An Application to Negotiation

In this section we present an application of Boolean games with incomplete preference-information. In particular, we develop a multilateral negotiation protocol which allows the agents to use their knowledge about other agents' goals. We analyze how a lack of knowledge affects the agreement outcome. In particular, we show how knowledgeable agents can obtain a more desirable outcome than others.

Negotiating allows agents in a strategic setting to settle on an agreement outcome. A multilateral bargaining protocol in Boolean games with complete information has been investigated in [14], showing that, when the logical structure of the goals is restricted, the protocol is guaranteed to end in a Pareto optimal outcome, i.e. no agent can improve its position without another agent being worse off. In this section, we propose a protocol which converges to an acceptable agreement without restrictions on the game structure. Moreover, under complete information our protocol always results in a discrimin optimal outcome. Discrimin optimality refines Pareto optimality [18], and while the latter indeed ensures *efficiency*, it is often not sufficient to characterize desirable outcomes [32]. Suppose, for instance, that two agents are negotiating in a situation with two Pareto optimal outcomes, with utility vectors $(1, 0.2)$ and $(0.5, 0.6)$, where utility reflects the degree of satisfaction of the agents. A natural concept arising in negotiation is *fairness*: intuitively, the latter utility vector is more fair than the former in the sense that there is less inequality between the utilities of the two agents. In the literature, several notions of fairness apart from discrimin optimality have been introduced and studied; we refer the interested reader to [33] for an overview and discussion.

The literature on bargaining is extensive, and covers a wide range of possible settings, such as discrete versus continuous bargaining (e.g. prices [34]), bilateral [35, 36] versus multilateral bargaining, transferable and non-transferable utility, a limited versus unlimited number of responses, modeling incomplete knowledge through probability theory [34], through Cartesian products of so-called complete knowledge problems [37], or by means of possibilistic logic [35]. To the best of our knowledge, our work is the first research on negotiation that considers incomplete knowledge about the goals of other agents in a Boolean game setting.

This section is structured as follows. In Section 6.1 we propose a multilateral negotiation protocol for Boolean games with *complete* information and characterize the agreement outcomes. Then we generalize the negotiation protocol of Section 6.1 to Boolean games with *incomplete* information in Section 6.2. We characterize the agreement outcomes, linking back to those under complete information. Additionally, we show how knowledge is crucial for an agent to reach a satisfying agreement. To conclude, we discuss several future work directions.

6.1. Negotiating under Complete Information

We investigate multilateral negotiation in Boolean games with prioritized goal bases (see Definition 1). We use the following scenario as a running example.

Example 14. Recall Example 1 where Alice and Bob, who share a car, are planning their afternoon. Alice controls $\Phi_1 = \{b_A, f_A, d_A\}$ and Bob controls $\Phi_2 = \{b_B, f_B\}$, respectively expressing the actions of going to the beach, going to the forest and taking the dog (in Alice's case). The game is constrained by $\delta = \neg(b_B \wedge f_B) \wedge \neg(b_A \wedge f_A) \wedge (b_B \to \neg f_A) \wedge (b_A \to \neg f_B)$. The goal bases of Alice and Bob are:

$$\Gamma_1 = \{f_A \wedge f_B \wedge d_A; b_A \wedge b_B \wedge d_A; d_A\}$$
$$\Gamma_2 = \{b_B \wedge b_A \wedge \neg d_A; b_B \wedge b_A; f_B \wedge f_A\}$$

In the context of bargaining, it is natural that agents try to achieve an outcome that is, among others, efficient. A well-known efficiency concept is Pareto optimality. Recall that an outcome is Pareto optimal if no agent can be better off without another agent being worse off (see Definition 5). It is easy to see that every Boolean game has at least one Pareto optimal outcome. A well-known refinement of the Pareto ordering incorporating a notion of fairness is the discrimin ordering (see Definition 6). In Example 14, (b_A, b_B, d_A) is the unique discrimin optimal outcome, although it is not the only Pareto optimal outcome.

We are interested in a negotiation protocol that is guaranteed to converge within a finite number of steps. Therefore, we want agents to make offers according to a negotiation rule, which ensures that every offered outcome is an improvement compared to the previous one. For instance, an agent might only be allowed to make a counteroffer if no agent is worse off than in the previous offer. Obviously, this rule will lead to Pareto optimal outcomes. However, the rule is so strict that the result can hardly be called fair: the first agent simply offers the outcome which yields its personal highest utility and no other agent is allowed to make a counteroffer which lowers the first agent's utility. Suppose, for instance, that there are two possible utility vectors: $(1, 0)$ and $(0.5, 0.5)$. If the first agent opens the negotiation with $(1, 0)$, the other agent would not be allowed to counter this offer with $(0.5, 0.5)$. To develop a fairer rule, we consider two properties that characterize a valid counteroffer. First of all, an agent is only interested in making a counteroffer if its own utility improves compared to the original offer. Second, the agents apply the *silver rule* or *ethic of reciprocity*, proposed by the Confucian Way of Humanity [38]:

> One should not treat others in ways that one would not like to be treated.

In our negotiation protocol, an agent reasons as follows: if I do not accept an offer of utility k, I should not lower another agent's utility to k or less in order to improve my own. Therefore, if an agent decides to lower other agents' utilities, it should offer more than k. We formally define the set $co(i, \nu)$ of agent i's legal counteroffers to the proposal ν as follows, where $i \in N$ and $\nu \in \mathcal{S}$:

$$co(i, \nu) = \{\nu' \in \mathcal{S} \mid u_i(\nu') > u_i(\nu) \wedge \forall j \in N : u_j(\nu') < u_j(\nu) \Rightarrow u_j(\nu') > u_i(\nu)\} \qquad (9)$$

We suggest the following negotiation protocol. In a given order, agents make proposals one by one. Without loss of generality, we assume that this order is $1, 2, \ldots, n$.
Algorithm 1 depends on a selection function to choose which $\nu' \in \mathcal{S}$ is made as the initial offer and which $\nu' \in co(i, \nu)$ is chosen each time as the counteroffer. The results discussed

Algorithm 1 Negotiation Protocol for Boolean game

$\nu \leftarrow \nu'$ with $\nu' \in \mathcal{S}$ % *Agent* 1 *proposes* ν'
accepted $\leftarrow 1$; $i \leftarrow 2$
while accepted $< n$ **do**
  if $co(i, \nu) == \emptyset$ **then**
    % *Agent* i *accepts the offer*
    accepted $\leftarrow$ accepted $+1$
  else
    % *Agent* i *rejects the offer and makes a counteroffer*
    $\nu \leftarrow \nu'$ with $\nu' \in co(i, \nu)$
    accepted $\leftarrow 1$
  end if
  $i \leftarrow (i == n\ ?\ 1 : i + 1)$
end while

in the paper hold regardless of this selection. The negotiation protocol ends if an offer ν is made such that no counteroffers can be made, i.e. $\forall i \in N : co(i, \nu) = \emptyset$.

If we apply Algorithm 1 in the context of Example 14, Alice is the first agent to make a proposal. Suppose she proposes to Bob the unique outcome corresponding to her first priority goal of going to the forest with Bob and her dog, i.e. $\nu_1 = \{f_A, f_B, d_A, \neg b_A, \neg b_B\}$. The associated utility vector is $(1, 0.33)$. Bob now looks into his possible counteroffers, i.e. $co(2, \nu_1)$. First of all, Bob has to be better off in his own counteroffer and should thus have utility 0.67 or greater. Therefore, the only candidates for $co(2, \nu_1)$ are *going to the beach exclusively with Alice* and *going to the beach with Alice and the dog*. The second condition for $co(2, \nu_1)$ regards Alice's utility: if Alice's utility in Bob's counteroffer is lower than in ν_1, then her utility in the counteroffer should be strictly greater than 0.33 (i.e. Bob's utility in Alice's proposal), or it should thus be at least 0.67. Given that Alice has utility 0 if she is not with her dog, the only possible counteroffer for Bob is *going to the beach with Alice and the dog*, i.e. $\nu_2 = \{b_B, b_A, d_A, \neg f_A, \neg f_B\}$. In other words, $co(2, \nu_1) = \{\nu_2\}$. Bob thus makes the counteroffer ν_2 with associated utility vector $(0.67, 0.67)$. If Alice now wants to make another counteroffer, then by definition of $co(1, \nu_2)$ both Alice and Bob should have utility 1 in this counteroffer, or Alice should have utility 1 and Bob should keep utility 0.67. Clearly no such counteroffers exist, i.e. $co(1, \nu_2) = \emptyset$. Alice therefore accepts Bob's offer ν_2 and the protocol ends. Bob and Alice have agreed to go to the beach with the dog.

We now prove that, whenever an offer is rejected, the new offer is fairer according to the discrimin ordering.

Proposition 4. *For $\nu \in \mathcal{S}$, $i \in N$ and $\nu' \in co(i, \nu)$, it holds that $\nu' >_{discr} \nu$.*

Proof. According to Definition 6 we need to prove:

$$\min_{j \notin eq(\mathcal{U}(\nu), \mathcal{U}(\nu'))} u_j(\nu') > \min_{j \notin eq(\mathcal{U}(\nu), \mathcal{U}(\nu'))} u_j(\nu)$$

For every $j \notin eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))$ either $u_j(\nu') > u_j(\nu)$ holds or $u_j(\nu') < u_j(\nu)$ holds by definition of $eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))$. For every $j \notin eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))$ such that $u_j(\nu') > u_j(\nu)$ holds, we have $u_j(\nu') > \min_{j' \notin eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))} u_{j'}(\nu)$ by definition of the minimum. For every $j \notin eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))$ such that $u_j(\nu') < u_j(\nu)$ holds, the definition of $co(i,\nu)$ (9) implies that $u_j(\nu') > u_i(\nu)$. Moreover, (9) implies that $i \notin eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))$. Consequently, we have $u_j(\nu') > \min_{j' \notin eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))} u_{j'}(\nu)$. Combining all parts, we see that the inequality $u_j(\nu') > \min_{j' \notin eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))} u_{j'}(\nu)$ holds for every $j \notin eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))$, thus implying what needed to be proven. □

It immediately follows that every discrimin optimal outcome is accepted.

Corollary 2. *For $\nu \in DO$ and $i \in N$ it holds that $co(i,\nu) = \emptyset$.*

Conversely, we can also show that only discrimin optimal outcomes will be overall accepted, i.e. accepted by every agent.

Proposition 5. *For $\nu \in \mathcal{S} \setminus DO$ there is an agent $i \in N$ with $co(i,\nu) \neq \emptyset$.*

Proof. Let $\nu \in \mathcal{S} \setminus DO$, then by definition of DO there exists a $\nu' \in \mathcal{S}$ such that $\nu' >_{discr} \nu$. By Definition 6 this implies that

$$\min_{j \notin eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))} u_j(\nu') > \min_{j \notin eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))} u_j(\nu)$$

Since the number of agents is finite, there exists at least one agent i outside $eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))$ such that $u_i(\nu) = \min_{j \notin eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))} u_j(\nu)$. In other words: of all agents for which the choice between ν and ν' matters, agent i has the lowest utility in ν. In particular, since $\nu' >_{discr} \nu$, it holds that $u_i(\nu') > u_i(\nu)$. Moreover, since $\min_{j \notin eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))} u_j(\nu') > u_i(\nu)$, it holds that $u_j(\nu') > u_i(\nu)$ for every $j \notin eq(\mathcal{U}(\nu),\mathcal{U}(\nu'))$. In particular, $u_j(\nu') > u_i(\nu)$ for every $j \in N$: $u_j(\nu') < u_j(\nu)$. Therefore, it immediately follows that $\nu' \in co(i,\nu)$. □

Note that since there are only a finite number of offers that can be made, and because each offer must strictly improve the previous offer in terms of the discrimin ordering, we know that the negotiation protocol always ends. From Corollary 2 and Proposition 5 we moreover know that the possible agreement outcomes at the end of the negotiation protocol are exactly the discrimin optimal outcomes. This result implies that the first offering agent still has a strong advantage, as this agent can select the discrimin optimal outcome that yields its highest personal utility, which no agent is allowed to reject. For instance, if the only discrimin optimal outcomes have utility vectors $(1, 0.5)$ and $(0.5, 1)$, agent 1 should propose the former and agent 2 has no choice but to accept. If the first agent follows this strategy, the negotiation ends within one step.

Remark 1. In our protocol, it is irrelevant which agent controls which atoms. The dependence of actions implied by the constraint δ forces agents to negotiate about what actions they will undertake. In Example 14, Alice and Bob cannot individually decide to go out. However, Alice can decide to stay with the dog without violating δ. Moreover, both Alice and Bob can decide to stay at home without restricting the other agent's options w.r.t. the constraint. Thus Alice is able to reach a utility of 0.33 without negotiating, and Bob will be stuck with a utility of 0. The utility vector $(0.33, 0)$ can be viewed as the *disagreement point* [39], i.e. the utility the agents would receive if they fail to reach an agreement. This information could be added to the framework: Alice rejects everything with a lower utility than 0.33, ergo Bob should not make such offers during the negotiation. Note that we can incorporate this info in the constraint δ, demanding that the utility of every agent is greater than its *disagreement utility.*

6.2. Negotiating under Incomplete Information

In this section, we consider Boolean games in which the agents are uncertain about the preferences of the other agents. For example, recall the context of Example 14 where Bob and Alice are planning their Sunday afternoon: they can go to the beach or the forest, or they can stay at home, and Alice can bring the dog or leave it at home. Since being a couple does not imply having identical preferences nor knowing exactly each other's preferences, Bob and Alice will have to compromise under incomplete information. However, they might not be completely ignorant about each other's goals; for instance, Bob knows that Alice loves the dog and Alice knows that Bob loves the beach. In this section, we show how they can use such knowledge to reach an agreement through negotiation. Moreover, we explore the link between having information about other agents' goals and obtaining a satisfactory agreement. To the best of our knowledge, our process is the first multilateral negotiation protocol for Boolean games that takes uncertainty w.r.t. the other agents' goals into account.

We first define a set of possibilistic discrimin optimal outcomes in $\mathcal{S}$. Intuitively, an outcome ν is optimal if for any outcome ν' which dominates ν according to the discrimin ordering, the agents who are better off in ν' than in ν are not aware that ν' is a valid counteroffer in the sense of (9). Recall that the models of a generalized possibilistic knowledge base $\mathcal{K}_i^j$ are possibility distributions.

Definition 18 (Possibilistic Discrimin). We define the set of possibilistic discrimin optimal outcomes:

$$DO^p = \{\nu \in \mathcal{S} \,|\, \forall \nu' \in \mathcal{S} : (\nu' >_{discr} \nu) \Rightarrow \big(\forall i \in N : (u_i(\nu') > u_i(\nu)) \Rightarrow (\exists j \in N, \exists u_i^j \in Mod(\mathcal{K}_i^j) : (u_i^j(\nu') < u_i^j(\nu)) \wedge (u_i^j(\nu') \leq u_i(\nu))))\big)\}$$

It is easy to see that $DO \subseteq DO^p$. In particular, when each agent has full knowledge, i.e. $Mod(\mathcal{K}_i^j) = \{u_j\}$ for every $i, j \in N$, DO and DO^p coincide.

We now analyze negotiation in Boolean games with incomplete preference-information. The protocol remains as specified in Algorithm 1: agents take turns in responding to an offer, by accepting it or making a counteroffer. However, the set of legal counteroffers $co(i, \nu)$ might be unknown to agent i. Indeed, determining the allowed counteroffers requires – possibly unknown – information about the other agents' utility. Therefore, we replace $co(i, \nu)$ by $co^p(i, \nu)$, which intuitively contains every outcome $\nu' \in \mathcal{S}$ for which agent i has enough information to derive that ν' is indeed a legal counteroffer to ν:

$$co^p(i, \nu) = \{\nu' \in \mathcal{S} \,|\, (\mathcal{K}_i^i \models \nu' \succ \nu) \wedge \forall j \in N : \mathcal{K}_i^j \models (\mathbf{\Delta}_{u_i(\nu)+\frac{1}{p}}(\nu') \vee (\nu' \succeq \nu))\}$$

Recall that an outcome ν, when used as a formula, represents the conjunction of the literals in ν. Moreover, as defined in (7), $\nu' \succ \nu$ and $\nu' \succeq \nu$ are abbreviations of GPL formulas. When agent i's knowledge base $\mathcal{K}_i^j$ models $\nu' \succ \nu$, this means that agent i knows that the utility of agent j in ν' is higher than its utility in ν. Similarly, $\mathcal{K}_i^j \models \nu' \succeq \nu$ encodes that agent i knows that the utility of agent j in ν' is at least as high as its utility in ν. If $\mathcal{K}_i^j$ models the GPL formula $\mathbf{\Delta}_{u_i(\nu)+\frac{1}{p}}(\nu')$, this means that agent i knows that agent j's utility in ν' is higher than $u_i(\nu)$. As before, an outcome ν is agreed upon iff $co^p(i,\nu) = \emptyset$ for every $i \in N$.

We can prove that every possibilistic discrimin optimal outcome is generally accepted. To this end, we first prove the following link between the sets of valid counteroffers under complete and incomplete information, assuming the agents have correct beliefs about the preferences of others. By the latter we mean that the beliefs of an agent i concerning the utility of an agent j do not rule out the true utility function of an agent j.

Proposition 6. *For every $i \in N$ such that $\forall j \in N : u_j \in Mod(\mathcal{K}_i^j)$ it holds that $co^p(i,\nu) \subseteq co(i,\nu)$ for every $\nu \in \mathcal{S}$.*

Proof. Let ν' be an arbitrary element of $co^p(i,\nu)$. By definition, it holds that $u_i(\nu') > u_i(\nu) \wedge \forall j \in N : \mathcal{K}_i^j \models (\mathbf{\Delta}_{u_i(\nu)+\frac{1}{p}}(\nu') \vee (\nu' \succeq \nu))$. Let $j \in N$. It remains to prove that $u_j(\nu') < u_j(\nu) \Rightarrow u_j(\nu') > u_i(\nu)$ or equivalently $u_j(\nu') \geq u_j(\nu) \vee u_j(\nu') > u_i(\nu)$. It is assumed that u_j is a model of $\mathcal{K}_i^j$. Consequently, it holds that u_j satisfies either $\mathbf{\Delta}_{u_i(\nu)+\frac{1}{p}}(\nu')$ or $\nu' \succeq \nu$. In the first case, it holds that $u_j(\nu') \geq u_i(\nu) + \frac{1}{p}$ or thus $u_j(\nu') > u_i(\nu)$. In the second case, it holds that $u_j(\nu') \geq u_j(\nu)$. In any case it holds that $u_j(\nu') \geq u_j(\nu) \vee u_j(\nu') > u_i(\nu)$. $\square$

We can prove that an agent always accepts a possibilistic discrimin optimal outcome, assuming it has correct beliefs.

Proposition 7. *For $\nu \in DO^p$ and $i \in N$ such that $\forall j \in N : u_j \in Mod(\mathcal{K}_i^j)$ it holds that $co^p(i,\nu) = \emptyset$.*

Proof. Let $\nu \in DO^p$ and $i \in N$. Suppose there exists some $\nu' \in co^p(i,\nu)$. If $\neg(\nu' >_{discr} \nu)$, then Proposition 4 and Proposition 6 imply that $\nu' \notin co^p(i,\nu)$, which is a contradiction. Now assume that $\nu' >_{discr} \nu$. Since $\nu' \in co^p(i,\nu)$ and $Mod(\mathcal{K}_i^i) = \{u_i\}$ it follows that $u_i(\nu') > u_i(\nu)$. Because $\nu \in DO^p$, we know that there exists some $j \in N$ and $u_i^j \in Mod(\mathcal{K}_i^j)$ such that $u_i^j(\nu') < u_i^j(\nu)$ and $u_i^j(\nu') \leq u_i(\nu)$. Consequently, u_i^j does not satisfy $\mathbf{\Delta}_{u_i(\nu)+\frac{1}{p}}(\nu')$, nor $\nu' \succeq \nu$. This contradicts the fact that $\mathcal{K}_i^j \models (\mathbf{\Delta}_{u_i(\nu)+\frac{1}{p}}(\nu') \vee (\nu' \succeq \nu))$. $\square$

Conversely, we can also show that only possibilistic discrimin optimal outcomes will be generally accepted, i.e. accepted by all agents.

Proposition 8. *For $\nu \in \mathcal{S} \setminus DO^p$ there is an agent $i \in N$ with $co^p(i,\nu) \neq \emptyset$.*

Proof. For $\nu \in \mathcal{S} \setminus DO^p$ there exists a $\nu' \in \mathcal{S}$ such that $\nu' >_{discr} \nu$ and there exists an agent $i \in N$ such that $u_i(\nu') > u_i(\nu)$, and for every $j \in N$ and $u_i^j \in Mod(\mathcal{K}_i^j)$ it holds that

$u_i^j(\nu') \geq u_i^j(\nu)$ or $u_i^j(\nu') > u_i(\nu)$. Since $Mod(\mathcal{K}_i^i) = \{u_i\}$, it follows that $\mathcal{K}_i^i \models \nu' \succ \nu$. Now let j be an arbitrary agent in N. For every model u_i^j of $\mathcal{K}_i^j$ such that $u_i^j(\nu') \geq u_i^j(\nu)$ it holds that u_i^j models $\nu' \succeq \nu$. For every model u_i^j of $\mathcal{K}_i^j$ such that $u_i^j(\nu') > u_i(\nu)$ it holds that u_i^j models $\mathbf{\Delta}_{u_i(\nu)+\frac{1}{p}}(\nu')$. Consequently, we have $\mathcal{K}_i^j \models (\mathbf{\Delta}_{u_i(\nu)+\frac{1}{p}}(\nu') \vee (\nu' \succeq \nu))$, and thus $\nu' \in co^p(i, \nu)$. $\square$

Note that since the number of possible offers is finite and because each offer must strictly improve the previous offer in terms of the discrimin ordering, the negotiation protocol always ends when we assume correct beliefs. Note that incorrect beliefs, however, can lead to an infinite loop: if agent 1 believes ν is a valid counteroffer to ν' and agent 2 believes that ν' is a valid counteroffer to ν, then a negotiation between the both of them can loop infinitely from ν to ν' and vice versa. From Proposition 7 and Proposition 8 we know that the possible agreement outcomes at the end of the negotiation protocol are exactly the possibilistic discrimin optimal outcomes, assuming the agents have correct beliefs.

From $DO \subseteq DO^p$ and Proposition 7, it follows that any discrimin optimal offer is overall accepted under incomplete information. However, Example 15 shows that the opposite does not hold, i.e. a non-discrimin optimal outcome might be accepted under incomplete information.

Example 15. Suppose, in the context of Example 14, that Alice has absolutely no information concerning Bob's goals. If Bob may make the first offer and suggests to go to the beach together without the dog, Alice's utility is 0. Although this outcome is discrimin dominated by going to the beach with the dog, Alice is unable to make this counteroffer, because she does not know whether Bob's utility is at least 0.33 in that case or whether Bob's utility is at least the same as in his first offer.

Note that, in contrast to a fully informed agent, an agent with limited knowledge might not be able to open with a discrimin-optimal solution. It is clear that having no information leaves an agent in a very weak position. Indeed, if agent i knows nothing about the preferences of another agent, it holds that $co^p(i, \nu) = \emptyset$ for every $\nu \in \mathcal{S}$, hence agent i is obliged to accept every offer. In contrast, an agent who has full knowledge knows all valid counteroffers and may be able to achieve a better outcome than in any discrimin optimal outcome, cfr. Bob in Example 15. Note that an agent with full knowledge can either use a safe or a risky selection function. Suppose for instance that there are only three possible utility vectors: $(0.6, 0.4)$, $(0.4, 0.6)$ and $(1, 0.2)$. If agent 1 proposes $(0.6, 0.4)$, it is certain that agent 2 accepts. Alternatively, if agent 1 proposes $(1, 0.2)$ and agent 2 does not know that there exists a valid counteroffer, agent 1 can get away with an unfair agreement, yielding a higher utility than in any fair outcome. However, if agent 2 knows that $(0.4, 0.6)$ is a valid counteroffer, the negotiations end in $(0.4, 0.6)$, leaving agent 1 worse off than if it had proposed $(0.6, 0.4)$ right away. This discussion shows that an interesting extension of the framework would be to allow agents to reason about the knowledge of others. Such knowledge can be encoded using multi-agent extensions of modal logics for epistemic reasoning, although we are then forced to express knowledge about preferences at the propositional level (e.g. by introducing variables g_i^m to denote the m^{th} most preferred goal of agent i as in Section 4). This extension would allow agents to act based on their knowledge of how other agents would react to various counteroffers, as is common in the field of epistemic game theory [40].

Remark 2. In our protocol, the order of the agents plays an important role, which is natural in hierarchical contexts (e.g. leader-follower type setting, where followers can only question proposals by leaders if they can prove their unfairness). Alternatively, the power of agents [41] can be used to deduce a sensible ordering in which agents are allowed to make offers: the most powerful agent can make the initial offer. Note, however, that the use of GPL for encoding knowledge about the preferences of others is independent of the negotiation protocol. Consequently, future research w.r.t. alternative negotiation protocols e.g. for settings in which agents have equal status can also rely on our GPL framework.

Even though the negotiation model we have discussed in this section is rather simple, it offers a rich basis from which we can study a wide variety of settings. Interesting extensions could include the use of agents who expand their knowledge base during the protocol, by drawing conclusions from the offering behavior of other agents [42]. Another option is to use different negotiation rules, e.g. an agent could be allowed to make a counteroffer ν if it does not know that ν is an illegal counteroffer. However, if the offer turns out to be illegal, the agent must pay a penatly. In that case, agents need to weigh the potential gain of such an offer against the risk of paying a penalty, which brings them in a standard setting for decision making under uncertainty (see e.g. [43]). Alternatively, we can allow 'third party' agents to protest against offers, in case they know that the offer is illegal. This can also be employed in case the assumption of correct beliefs is violated. Recall that the characterization of the agreement outcomes under incomplete information relies on this assumption. When an agent has incorrect beliefs, it is possible that the agent makes a counteroffer that violates the original bargaining rule, while believing it does not. For instance, suppose Bob suggests to Alice to attend a sports game together. Now assume Alice mistakingly believes Bob is indifferent between attending a sports game together and going to the theatre together, while Alice prefers the latter. She therefore believes going to the theatre together is a valid counteroffer. If she makes this offer, Bob can deduce that Alice's beliefs concerning his preferences are wrong. He could protest against the counteroffer and Alice could update her beliefs. However, protesting against an unfair proposal requires the revelation of knowledge, which might also weaken the bargaining power of the agent. Hence, it is not straightforward that protesting is always in the protester's advantage, even if it initially leads to a higher utility. Other options for alternative protocols include the addition of time constraints [44] or the use of arguments to support an offer [35].

7. Conclusion

We introduced the first Boolean game frameworks that allow agents to be uncertain about the other agents' goals. Moreover, we approached the incompleteness of information in two distinct ways. In Section 4 we have argued that such a scenario can naturally be modelled by associating with each agent a possibility distribution over the universe of all possible games (given the considered action variables and constraints). While this allows us to define a variety of solution concepts in a natural way, definitions at the semantic level are not directly useful in practice, due to the exponential size of these possibility distributions. Therefore, we also proposed a syntactic characterization, which avoids exponential representations by relying on standard possibilistic logic inference, and can be implemented by reduction to answer set programming. Our main result is that this syntactic characterization indeed corresponds to the intended semantic definitions. We furthermore showed that the computational complexity of the introduced solution concepts remains at the second level of the polynomial hierarchy.

In Section 5 we have developed an alternative framework for Boolean games with incomplete information, using GPL to compactly represent agents' knowledge about the preferences of others. In contrast to the previous approach, here the possibility distributions correspond to utility functions, hence the weights of formulas reflect preference instead of certainty. An agent then considers a set of possibility distributions as being the possible utility functions of another agent. We illustrated how the syntax of GPL allows us to easily model intuitive notions: not only can we capture necessary and sufficient conditions for reaching subgoals, it is also straightforward to encode e.g. comparative preferences. Moreover, in contrast to the framework from Section 4, the GPL-based model does not require additional variables in the logical language. However, the approach from Section 4 allows different degrees of certainty, whereas the GPL-based model corresponds to binary certainty: either an agent completely rules a scenario out or it considers a scenario completely possible. To the best of our knowledge, our frameworks are the first models for Boolean games with incomplete information regarding the agents' goals.

We introduced new solution concepts, which are appropriate for this context. They reflect whether an outcome is known to be a PNE or whether it is considered possible of being a PNE. For instance, in case of a known PNE, all agents believe that no other agent has the incentive to deviate. Moreover, we investigated the associated computational complexity and showed that this complexity does not increase compared to PNEs in Boolean games with complete information.

To illustrate how the proposed frameworks could be used in practice, we presented an application to negotiation in the context of Boolean games with incomplete information. Our multilateral negotiation protocol uses an intuitive negotiation rule based on the ethic of reciprocity principle and is guaranteed to converge within a finite number of steps. We characterized the set of possible outcomes of the negotiation process, confirming the intuition that incomplete knowledge may lead to negotiation inefficiency, i.e. the agreement outcome may not be fair or efficient.

The presented frameworks lead to several interesting avenues for future work. First, the approaches could be generalized for taking into account prior beliefs about the likely behaviour of other players (e.g. for modelling collusion) and/or for modelling situations where agents may be uncertain about the actions that are being played by other agents. Moreover, it seems of interest to analyse the effect of adding communication to the framework, by allowing agents to strategically ask questions or make proposals to each other in order to reduce uncertainty or as part of a negotiation process.

[1] M. Osborne, A. Rubinstein, A Course in Game Theory, MIT press, 1994.

[2] J. C. Harsanyi, Games with incomplete information, The American Economic Review 85 (3) (1995) 291–303.

[3] D. Dubois, J. Lang, H. Prade, Possibilistic logic, in: Handbook of Logic for Artificial Intelligence and Logic Programming, Vol. 3, Oxford University Press, 1994, pp. 439–513.

[4] D. Dubois, H. Prade, S. Schockaert, Reasoning about uncertainty and explicit ignorance in generalized possibilistic logic, in: Proc. ECAI, 2014, pp. 261–266.

[5] S. De Clercq, S. Schockaert, M. De Cock, A. Nowé, Possibilistic Boolean games: Strategic reasoning under incomplete information, in: Proc. JELIA, 2014, pp. 196–209.

[6] S. De Clercq, S. Schockaert, A. Nowé, M. De Cock, Multilateral negotiation in Boolean games with incomplete information using generalized possibilistic logic, in: Proceedings of IJCAI, AAAI Press, 2015, pp. 2890–2896.

[7] J. Grant, S. Kraus, M. Wooldridge, I. Zuckerman, Manipulating Boolean games through communication, in: Proc. IJCAI, 2011.

[8] T. Ågotnes, P. Harrenstein, W. van der Hoek, M. Wooldridge, Verifiable equilibria in Boolean games, in: Proc. IJCAI, 2013.

[9] T. Ågotnes, P. Harrenstein, W. van der Hoek, M. Wooldridge, Boolean games with epistemic goals, in: Logic, Rationality, and Interaction, Vol. 8196 of LNCS, Springer, 2013, pp. 1–14.

[10] D. Dubois, H. Prade, Epistemic entrenchment and possibilistic logic, Artif. Intell. 50 (2) (1991) 223–239.

[11] A. Brandenburger, On the existence of a ‘complete’ possibility structure, in: M. Basili, N. Dimitri, I. Gilboa (Eds.), Cognitive Processes and Economic Behavior, Routledge, 2003, pp. 30–34.

[12] J. Chen, S. Micali, R. Pass, Tight revenue bounds with possibilistic beliefs and level-k rationality, Econometrica 83 (4) (2015) 1619–1639.

[13] D. Dubois, H. Prade, Possibility Theory, Plenum Press, New York (NY), 1988.

[14] P. Dunne, W. van der Hoek, S. Kraus, M. Wooldridge, Cooperative Boolean games, in: Proc. AAMAS, Vol. 2, IFAAMAS, 2008, pp. 1015–1022.

[15] E. Bonzon, M.-C. Lagasquie-Schiex, J. Lang, Compact preference representation for Boolean games, in: Proc. PRICAI, Springer-Verlag, 2006, pp. 41–50.

[16] E. Marchioni, M. Wooldridge, Łukasiewicz games: A logic-based approach to quantitative strategic interactions, ACM Transactions on Computational Logic 16 (4) (2015) 33:1–33:44.

[17] E. Bonzon, M.-C. Lagasquie-Schiex, J. Lang, Effectivity functions and efficient coalitions in Boolean games, Synthese 187 (2012) 73–103.

[18] D. Dubois, H. Fargier, H. Prade, Beyond min aggregation in multicriteria decision: (ordered) weighted min, discri-min, leximin, in: The ordered weighted averaging operators, Springer, 1997, pp. 181–192.

[19] D. Dubois, H. Prade, Possibilistic logic: a retrospective and prospective view, Fuzzy Sets and Systems 144 (2004) 3–23.

[20] H. J. Levesque, All I know: A study in autoepistemic logic, Artificial Intelligence 42 (1990) 263 – 309.

[21] T. Schelling, The strategy of conflict, Oxford University Press, 1960.

[22] J. C. Harsanyi, Games with incomplete information played by bayesian players part i, Management Science 14 (1968) 159–182.

[23] J. C. Harsanyi, Games with incomplete information played by bayesian players part ii, Management Science 14 (1968) 320–334.

[24] J. C. Harsanyi, Games with incomplete information played by bayesian players part iii, Management Science 14 (1968) 486–502.

[25] S. Schockaert, H. Prade, Solving conflicts in information merging by a flexible interpretation of atomic propositions, Artif. Intell. 175 (11) (2011) 1815–1855.

[26] S. De Clercq, K. Bauters, S. Schockaert, M. Mihaylov, A. Nowé, M. De Cock, Exact and heuristic methods for solving boolean games, Autonomous Agents and Multi-Agent Systems 31 (1) (2017) 66–106.

[27] S. Benferhat, S. Kaci, Logical representation and fusion of prioritized information based on guaranteed

possibility measures: Application to the distance-based merging of classical bases, Artif. Intell. 148 (2003) 291–333.
[28] D. Dubois, H. Fargier, H. Prade, Possibility theory in constraint satisfaction problems: Handling priority, preference and uncertainty, Appl. Intell. 6 (4) (1996) 287–309.
[29] S. Benferhat, D. Dubois, H. Prade, Towards a possibilistic logic handling of preferences, Appl. Intell. 14 (3) (2001) 303–317.
[30] J. Harsanyi, Games with incomplete information played by "Bayesian players", I-III, Management Science 14 (3) (1967) 159–182.
[31] D. Dubois, H. Prade, S. Schockaert, Generalized possibilistic logic: Foundations and applications to qualitative reasoning about uncertainty.
[32] S. Bouveret, J. Lang, Efficiency and envy-freeness in fair division of indivisible goods: Logical representation and complexity, J. Artif. Intell. Res. (2008) 525–564.
[33] B. Tungodden, Egalitarianism: Is leximin the only option?, Econ. Philos. 16 (2000) 229–245.
[34] S. Fatima, M. Wooldridge, N. Jennings, Optimal negotiation strategies for agents with incomplete information, in: Intelligent Agents VIII, Springer, 2002, pp. 377–392.
[35] L. Amgoud, H. Prade, A possibilistic logic modeling of autonomous agents negotiation, in: Progress in Artificial Intelligence, Springer, 2003, pp. 360–365.
[36] X. Luo, N. Jennings, N. Shadbolt, H. Leung, J. Lee, A fuzzy constraint based model for bilateral, multi-issue negotiations in semi-competitive environments, Artif. Intell. 148 (1) (2003) 53–102.
[37] W. Bossert, H. Peters, Minimax regret and efficient bargaining under uncertainty, Games and Economic Behavior 34 (1) (2001) 1–10.
[38] J. Hertzler, On golden rules, Int. J. Ethics 44 (4) (1934) 418–436.
[39] K. Binmore, A. Rubinstein, A. Wolinsky, The Nash bargaining solution in economic modelling, The RAND Journal of Economics (1986) 176–188.
[40] A. Perea, Epistemic game theory: reasoning and choice, Cambridge University Press, 2012.
[41] J. Ben-Naim, E. Lorini, Evaluating power of agents from dependence relations in Boolean games, in: Proc. AAMAS, 2014, pp. 853–860.
[42] P. Cramton, Bargaining with incomplete information: An infinite-horizon model with two-sided uncertainty, The Review of Economic Studies 51 (4) (1984) 579–593.
[43] D. Dubois, H. Fargier, H. Prade, R. Sabbadin, A survey of qualitative decision rules under uncertainty, Decision-making Process: Concepts and Methods (2009) 435–473.
[44] S. Kraus, J. Wilkenfeld, G. Zlotkin, Multiagent negotiation under time constraints, Artif. Intell. 75 (2) (1995) 297–345.