On the Logic of Coalitional Games

Thomas Ågotnes
Department of Informatics, University of Bergen
PB. 7800, N-5020 Bergen, Norway
agotnes@ii.uib.no

Wiebe van der Hoek
Department of Computer Science, University of Liverpool
Liverpool L69 7ZF, UK
wiebe@csc.liv.ac.uk

Michael Wooldridge
Department of Computer Science, University of Liverpool
Liverpool L69 7ZF, UK
mjw@csc.liv.ac.uk

ABSTRACT

We develop a logic for representing and reasoning about coalitional games without transferable payoffs. Although a number of logics of cooperation have been proposed over the past decade (notably Coalition Logic [14] and Alternating-time Temporal Logic [1]), these logics focused primarily on the issue of strategic cooperative ability – what states a coalition can effectively enforce – and have tended to ignore the essential issue of the preferences that agents have over such states; in addition, the connection between such logics and coalitional games, in the sense of cooperative game theory, is left implicit. The Coalitional Game Logic (CGL) that we develop in this paper differs from such previous logics in two important respects. First, CGL includes operators that make it directly possible to represent an agent's preferences over outcomes. Second, we interpret formulae of CGL directly with respect to coalitional games without transferable payoff, thereby establishing an explicit link between formulae of the logic and properties of coalitional games. We show that these coalitional games cannot be seen directly as models for Coalition Logic. We give a complete axiomatization of CGL, prove that it is expressively complete with respect to coalitional games without transferable payoff, show that the satisfiability problem for the logic is NP-complete, and to illustrate its use, we show how the logic can be used to characterise axiomatically a number of well-known solution concepts for coalitional games, including for example non-emptiness of the core.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence-Multiagent Systems; I.2.4 [**Artificial Intelligence**]: Knowledge Representation Formalisms and Methods—*Modal logic*

General Terms

Theory

Keywords

Game theory, coalitional games, logic

AAMAS'06 May 8–12 2006, Hakodate, Hokkaido, Japan.

1. INTRODUCTION

If one aims to build software agents that must inhabit and act within some particular environment, then it is very natural to equip such agents with formal *representations* of the environment, in order to permit them to effectively select and execute actions within it. Previous work on knowledge representation in multi-agent systems may be broadly divided into two distinct categories. In the 1990s, considerable emphasis was placed on representations of the *cognitive structure* of agents – their beliefs, desires, intentions, and so forth. Key issues in such representations include for example the way in which communicative actions (speech acts) affect the mental states of conversation participants, and how we can characterise the mental states of agents engaged in teamwork [5]. More recently, an increasing body of work has focused instead on logics that make it possible to represent the strategic structure of multi-agent environments, and in particular, the powers that agents or groups of agents have in such environments [1, 13]. Such logics have proved to have many important applications, for example in the specification and verification of social choice mechanisms [13].

One significant feature of these cooperation logics is that they have a close link with formal games: the semantic models underpinning Coalition Logic (which corresponds to the next-time fragment of ATL [8]) can be understood as extensive games of almost perfect information [13, p.34]. Further links between logics of strategic ability and formal games have been established elsewhere – for example, van der Hoek and colleagues showed how solution concepts such as Nash equilibrium can be axiomatized in a variant of ATL [16]. However, as a general formalism for reasoning about cooperative games, ATL-like cooperation logics have limitations: their modal structures are only intended to support reasoning about cooperative ability, and they provide no direct mechanism for capturing preferences, which are of course a critical component of all but the very simplest multi-agent scenarios.

Our aim in this paper is to develop a logic that is intended to directly and transparently support reasoning about coalitional games – and more precisely, coalitional games without transferable payoff [11, p.268]. We develop a *Coalitional Game Logic* (CGL). Syntactically, CGL appears quite similar to ATL and Coalition Logic, in that it contains modal cooperation expressions. However, it differs from these previous logics in several important respects. First, CGL includes operators that make it *directly* possible to represent an agent's preferences over outcomes. Second, we interpret formulae of CGL directly with respect to coalitional games without transferable payoff, thereby establishing an explicit link between formulae of the logic and properties of coalitional games. Third, we show that a coalitional game indeed cannot be seen as a model for coalition logic with outcomes of the game as states.

Following the presentation of the syntax and semantics of the logic in the next section, we give a number of technical results relating to it, as follows. First, we prove that the logic is expressively complete with respect to coalitional games without transferable payoff, in the sense that for any two different coalitional games, there exists a formula of CGL that will be true in one game and false in the other. We then give an axiomatization of CGL, and show that it is complete. With respect to model checking and satisfiability, we show that while model checking for the logic is tractable, the satisfiability problem for CGL is NP-complete. We then compare the semantics of CGL and coalition logic. Finally, to illustrate the utility of the logic, we show how to axiomatically characterise a number of well-known solution concepts for coalitional games, including for example non-emptiness of the core.

2. A LOGIC FOR COALITIONAL GAMES

A *coalitional game* (without transferable payoff) is an $(m+3)$-tuple [11, p.268]:

$$\Gamma = \langle N, \Omega, V, \sqsupseteq_1, \ldots, \sqsupseteq_m \rangle \text{ where:}$$

- $N = \{1, \ldots, m\}$ is a non-empty set of *agents*;
- $\Omega = \{\omega_1, \ldots, \omega_n\}$ is a non-empty set of *outcomes*;
- $V : (2^N \setminus \emptyset) \to 2^\Omega$ is the *characteristic function* of Γ, which for every non-empty coalition C defines the choices $V(C)$, such that if $\omega \in V(C)$ then C can choose outcome ω; and
- $\sqsupseteq_i \subseteq \Omega \times \Omega$ is a complete, reflexive, and transitive *preference relation*, for each agent $i \in N$.

The above definition also includes the additional assumption, used in this paper, that the set of outcomes is finite.

Readers familiar with ATL [1] or Coalition Logic [14] may be tempted to interpret V as an effectivity function and the outcomes Ω as states of the world. This is *not* the intended interpretation, and indeed, we show in Section 6 that there is no direct mapping between the coalition models of [14] and coalitional games by interpreting outcomes Ω as states and V as an effectivity function. For example, it is perfectly consistent in a coalitional game that, at the same time, disjoint coalitions have the abilities to choose *different* outcomes.

We now define a logic for representing the properties of such games. The language is in two parts. First, given a set of outcome symbols Σ_Ω, we have an *outcome language* $\mathcal{L}_o$, defined by the grammar φ_o, below, which expresses the properties of outcomes. The outcome symbols themselves are the main constructs of this language; a formula such as $\omega_1 \vee \omega_2$ (where $\omega_1, \omega_2 \in \Sigma_\Omega$) means that the outcome corresponds to either ω_1 or ω_2. Next, given a set of agent symbols Σ_N and a set of coalition symbols Σ_C, we have a *cooperation language* $\mathcal{L}_c$, for expressing the properties of coalitional cooperation, and the preferences that agents have over possible outcomes. This language is generated by the grammar φ_c below. $\mathcal{L}_c$ has two main constructs. First, $\omega_1 \succeq_i \omega_2$ (where $\omega_1, \omega_2 \in \Sigma_\Omega$, $i \in \Sigma_N$) expresses the fact that agent i either prefers outcome ω_1 over outcome ω_2, or is indifferent between the two. Second, $\langle C \rangle \varphi$ (where $C \in \Sigma_C$) says that C can choose an outcome in which the formula φ will be true. This construct may seem syntactically similar to counterparts in ATL and Coalition Logic, but it stands here for a fundamentally different concept due to the semantic differences mentioned above. Recall that we construct formulae with respect to a set $\Sigma = \Sigma_N \cup \Sigma_C \cup \Sigma_\Omega$ of symbols, where Σ_N is a set of symbols for agents, Σ_C is a set of symbols for coalitions and Σ_Ω is a set of symbols for outcomes. Formally, we have the following:

$$\begin{array}{lcl} \varphi_o & ::= & \sigma_\omega \mid \neg\varphi_o \mid \varphi_o \vee \varphi_o \\ \varphi_c & ::= & (\sigma_\omega \succeq_{\sigma_i} \sigma_{\omega'}) \mid \langle \sigma_C \rangle \varphi_o \mid \neg\varphi_c \mid \varphi_c \vee \varphi_c \end{array}$$

where $\sigma_i \in \Sigma_N$ is an agent symbol, $\sigma_C \in \Sigma_C$ is a coalition symbol, and $\sigma_\omega, \sigma_{\omega'} \in \Sigma_\Omega$ are outcome symbols.

We will usually exploit the direct correspondence between symbols for outcomes/agents and the outcomes/agents that appear in games. In this paper, we will henceforth assume a one-to-one correspondence between Σ_Ω and Ω, between Σ_N and N and between Σ_C and 2^N. So we assume that $\Sigma_\Omega = \{\sigma_\omega : \omega \in \Omega\}$, $\Sigma_N = \{\sigma_i : i \in N\}$ and $\Sigma_C = \{\sigma_C : C \subseteq N\}$. The languages are parameterised by the sets $\Sigma_\Omega, \Sigma_N, \Sigma_C$, i.e., as a consequence of the assumption, by some set N of agents and set Ω of outcomes. In the following, we assume that these two parameters — and thus the languages — are fixed.

An $\mathcal{L}_c$ formula γ is interpreted in a coalitional game Γ as follows, where $\Gamma \models \gamma$ means that γ is true in Γ. When we write $\Gamma \models \gamma$, it is implicitly assumed that the Γ corresponds to the parameters of the language in the way mentioned above, i.e., that the agents and outcomes in the model are exactly the sets used to parameterise our language. First, we define the satisfaction of a $\mathcal{L}_o$ formula α in an outcome ω of a game Γ, written $\Gamma, \omega \models \alpha$:

$\Gamma, \omega \models \sigma_{\omega'}$ iff $\omega = \omega'$

$\Gamma, \omega \models \neg\varphi$ iff not $\Gamma, \omega \models \varphi$

$\Gamma, \omega \models \varphi \vee \psi$ iff $\Gamma, \omega \models \varphi$ or $\Gamma, \omega \models \psi$

Satisfaction of γ in Γ is then defined as follows:

$\Gamma \models (\sigma_{\omega_1} \succeq_{\sigma_i} \sigma_{\omega_2})$ iff $(\omega_1 \sqsupseteq_i \omega_2)$

$\Gamma \models \langle \sigma_C \rangle \varphi$ iff $\exists \omega \in V(C)$ such that $\Gamma, \omega \models \varphi$

$\Gamma \models \neg\varphi$ iff not $\Gamma \models \varphi$

$\Gamma \models \varphi \vee \psi$ iff $\Gamma \models \varphi$ or $\Gamma \models \psi$

To simplify the text that follows, we abuse notation somewhat, and write ω for both an outcome (a semantic construct) and the corresponding symbol σ_ω in the language (a syntactic construct). Similarly, we will write i instead of σ_i for agents in the language, and C instead of σ_C for coalitions. So, we will just write $\langle C \rangle \omega$ for $\langle \sigma_C \rangle \sigma_\omega$, although the reader should be aware of the distinction between our object language $\mathcal{L}_c$ and the objects that live in the semantics: outcomes, agents and their preferences, and coalitions.

We will use the usual derived propositional connectives; $\varphi \wedge \psi$ for $\neg(\neg\varphi \vee \neg\psi)$, $\varphi \to \psi$ for $\neg\varphi \vee \psi$ and $\varphi \leftrightarrow \psi$ for $(\varphi \to \psi) \wedge (\psi \to \varphi)$, as well as $\varphi \nabla \psi$ for $(\varphi \vee \psi) \wedge \neg(\varphi \wedge \psi)$ (exclusive or) and $[C]\varphi$ for $\neg\langle C \rangle \neg\varphi$. We also write $(\omega_1 \succ_i \omega_2)$ to abbreviate $((\omega_1 \succeq_i \omega_2) \wedge \neg(\omega_2 \succeq_i \omega_1)$ and $(\omega_1 =_i \omega_2))$ to abbreviate $((\omega_1 \succeq_i \omega_2) \wedge (\omega_2 \succeq_i \omega_1))$.

Note that $\langle C \rangle \top$ iff C can at least bring about something: $V(C) \neq \emptyset$. $[C]\varphi$ means $\neg\langle C \rangle \neg\varphi$, i.e., every choice of C must involve φ. As an example, suppose $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and $V(C) = \{\omega_1, \omega_2\}$. Then:

$$\langle C \rangle \omega_1 \wedge \langle C \rangle (\omega_1 \vee \omega_3) \wedge \neg\langle C \rangle \omega_3 \wedge [C](\omega_1 \vee \omega_2) \wedge \neg[C](\omega_1 \vee \omega_3)$$

Note that if $\omega_1 \neq \omega_2$, then we can have $\langle C \rangle \omega_1 \wedge \langle C \rangle \omega_2$, but the formula $\langle C \rangle (\omega_1 \wedge \omega_2)$ can never be true.

A conjunction $\bigwedge_{\delta \in \emptyset} \varphi$ is, by convention, equal to $\top$, like $\bigvee_{\delta \in \emptyset} \varphi$ equals $\bot$, so that, indeed, we get $\Gamma \models [C]\bot$ iff $\Gamma \models \neg\langle C \rangle \top$ iff

$V(C) = \emptyset$. Let us, for any coalition C and set of outcome symbols Δ, suggestively write $\langle[C]\rangle\Delta$ for $\bigwedge_{\delta\in\Delta}\langle C\rangle\delta \wedge [C]\bigvee_{\delta\in\Delta}\delta$.[1] A formula of this form is said to *fully describe C's choices*. It is easy to see that we have the following. Let $\Delta \subseteq \Omega$.

$$\Gamma \models \langle[C]\rangle\Delta \quad \text{iff} \quad V(C) = \Delta$$

Exclusive disjunctions $\varphi\nabla\psi$ play an important role in our proofs. Note that the negation $\neg(\varphi\nabla\psi)$ is the same as $(\neg\varphi\wedge\neg\psi)\vee(\varphi\wedge\psi)$. Moreover, if Φ is a set of formulas, then we define $\bigtriangledown_{\varphi\in\Phi}\varphi$ to be true iff exactly one of the φ's is true. Formally: for any set $\Phi = \{\varphi_1, \ldots \varphi_k\}$,[2]

$$\bigtriangledown_{\varphi\in\Phi}\varphi \equiv (\bigvee_{i\leq k}\varphi_i \wedge \bigwedge_{j\neq i, j\leq k}\neg\varphi_j)$$

Note that $\Gamma \models [C](\omega_i \vee \omega_j) \leftrightarrow [C](\omega_i\nabla\omega_j)$ when $i \neq j$: using the definition of $[C]$ and contraposition this is the same as $\Gamma \models \langle C\rangle\neg(\omega_i\nabla\omega_j) \leftrightarrow \langle C\rangle\neg(\omega_i \vee \omega_j)$. Now, syntactically, $\langle C\rangle\neg(\omega_i\nabla\omega_j)$ is equivalent to $\langle C\rangle((\omega_i\wedge\omega_j)\vee\neg(\omega_i\vee\omega_j))$. But, inspecting the truth-definition of $\langle C\rangle$, this is again equivalent to $\langle C\rangle\neg(\omega_i\vee\omega_j)$ since the $\mathcal{L}_o$ formula $\omega_i\wedge\omega_j$ is never true.

So, which properties of a coalitional game can be expressed with our cooperation language? The answer, given by the following theorem, is "all".

THEOREM 1. *The logic* CGL *is expressively complete with respect to coalitional games. That is, for any two coalitional games Γ_1, Γ_2 such that $\Gamma_1 \neq \Gamma_2$, there exists a* CGL *formula ζ such that $\Gamma_1 \models \zeta$ and $\Gamma_2 \not\models \zeta$.*

PROOF. *Our proof is constructive. Given a game Γ, we define a formula ζ_Γ that* completely characterises Γ. *ζ_Γ is constructed from two conjuncts, Π_Γ, which characterises the preference relations of Γ, and Ξ_Γ, which characterises the cooperative properties of Γ. Let $\mathcal{C} = 2^N \setminus \emptyset$ collect all non-empty coalitions from N.*

$$\zeta_\Gamma \equiv \Pi_\Gamma \wedge \Xi_\Gamma$$

$$\Pi_\Gamma \equiv \bigwedge_{i\in N}(\bigwedge_{\substack{\omega,\omega'\in\Omega \\ \omega\sqsupseteq_i\omega'}}(\omega \succeq_i \omega') \wedge \bigwedge_{\substack{\omega,\omega'\in\Omega \\ \omega\not\sqsupseteq_i\omega'}}\neg(\omega\succeq_i\omega'))$$

$$\Xi_\Gamma \equiv \Big(\bigwedge_{C\in\mathcal{C}}(\bigwedge_{\omega\in V(C)}\langle C\rangle\omega) \wedge [C](\bigvee_{\omega\in V(C)}\omega)\Big)$$

By construction, for any Γ_1, we have $\Gamma_1 \models \zeta_{\Gamma_1}$. Moreover, for any coalitional game $\Gamma_2 \neq \Gamma_1$, we have that $\Gamma_2 \not\models \zeta_{\Gamma_1}$. $\Box$

From now on, $\mathcal{C}$ is the set of non-empty coalitions. Given a formula φ, then: let $coal(\varphi)$ denote the set of coalitions named in φ; let $ag(\varphi)$ denote the set of agents named in cooperation expressions in φ; and let $out(\varphi)$ be the set of outcomes named in φ.

3. AXIOMS AND COMPLETENESS

We now present an axiomatic system for the language $\mathcal{L}_c$, and prove its soundness and completeness with respect to the class of all coalitional games without transferable payoff.

Table 1 summarizes the axioms and rules of our logic CGL. Formally, CGL is the set of all $\mathcal{L}_c$-formulas derivable under $\vdash$. In the axioms, $\vdash_c$ denotes derivability of classical logic, and $\varphi_o \in \mathcal{L}_o, \varphi, \psi \in \mathcal{L}_c$. The axiom *Taut* and rule *MP* guarantee that we extend classical logic. On top of that, the axiom K and rule *Nec* determine $[C]$ to be a normal necessity operator. Then, *Lin*, *Ref*, and *Trans* determine the preference of each i to be linear, reflexive and transitive, respectively. The only specific cooperation axiom, *Func*, says that whatever a coalition in the end will chose, it must be a unique alternative from Ω.

Taut	$\vdash \varphi$ φ an instance of a propositional tautology
Lin	$\vdash (\omega_1 \succeq_i \omega_2) \vee (\omega_2 \succeq_i \omega_1)$
Ref	$\vdash (\omega \succeq_i \omega)$
Trans	$\vdash (\omega_1 \succeq_i \omega_2) \wedge (\omega_2 \succeq_i \omega_3) \rightarrow (\omega_1 \succeq_i \omega_3)$
K	$\vdash [C](\varphi\rightarrow\psi) \rightarrow (([C]\varphi) \rightarrow ([C]\psi))$
Func	$\vdash [C](\bigtriangledown_{\omega\in\Omega}\omega)$
Nec	$\vdash_c \varphi_o \Rightarrow \vdash [C]\varphi_o$
MP	$\vdash \varphi, \vdash \varphi\rightarrow\psi \Rightarrow \vdash \psi$

Table 1: The logic CGL**.**

The following lemma tells us that in the scope of modal operators, disjunctions over *different* outcomes behave the same as exclusive disjunctions over outcomes. Note that this is in general not true for arbitrary disjunctions: $\langle C\rangle(\omega_1\vee\omega_1)$ is not the same as $\langle C\rangle(\omega_1\nabla\omega_1)$, the latter is equivalent to $\langle C\rangle\bot$.

LEMMA 1. *Let $\emptyset \neq C \subseteq N$ and $\Delta \subseteq \Omega$.*

1. *The following are equivalent, in* CGL*:* $(i)\langle C\rangle\top$, $(ii)\langle C\rangle\bigvee_{\omega\in\Omega}\omega$, *and* $(iii)\langle C\rangle\bigtriangledown_{\omega\in\Omega}\omega$
2. *In the scope of $\langle C\rangle$ and $[C]$ when exchanging arbitrary occurrences of $\bigtriangledown_{\omega\in\Delta}\omega$ with that of $\bigvee_{\omega\in\Delta}\omega$ in a formula φ, the result is equivalent to φ.*
3. $\vdash \bigwedge_{\omega\in\Delta}\neg\langle C\rangle\omega \rightarrow (\langle C\rangle\top \rightarrow \bigvee_{\omega\in\Omega\setminus\Delta}\langle C\rangle\omega)$
4. $\vdash (\bigwedge_{\delta\in\Delta}\neg\langle C\rangle\delta) \leftrightarrow [C]\bigvee_{\delta'\in\Omega\setminus\Delta}\delta'$

PROOF.

1. Since $\bigtriangledown_{\omega\in\Omega}\omega \Rightarrow \bigvee_{\omega\in\Omega}\omega \Rightarrow \top$, and the $\langle C\rangle$ is a normal diamond operator, we have $(iii) \Rightarrow (ii) \Rightarrow (i)$. By *Func*, we have $(i) \Rightarrow (iii)$.
2. With induction over φ. The only interesting case $\varphi = \langle C\rangle\psi$ follows from the previous item.
3. From axiom *Func* follows $\vdash \langle C\rangle\top \rightarrow \langle C\rangle\bigtriangledown_{\omega\in\Omega}\omega$. By item 1, we have $\vdash \langle C\rangle\top \rightarrow \langle C\rangle\bigvee_{\omega\in\Omega}\omega$. Since $\langle C\rangle$ is a normal diamond, we have $\vdash \langle C\rangle\top \rightarrow \bigvee_{\omega\in\Omega}\langle C\rangle\omega$. Applying *Taut* to this gives the desired property.
4. Follows directly from *Func* and some modal reasoning.

$\Box$

LEMMA 2.

1. *Let $C \neq \emptyset$ be a coalition. Then* $\vdash \bigvee_{\Delta\subseteq\Omega}\langle[C]\rangle\Delta$
2. *Let $C \neq \emptyset$ be a coalition. Then* $\vdash \bigtriangledown_{\Delta\subseteq\Omega}\langle[C]\rangle\Delta$
3. $\vdash \bigwedge_{C\in\mathcal{C}}\bigvee_{\Delta\subseteq\Omega}\langle[C]\rangle\Delta$
4. $\vdash \bigwedge_{C\in\mathcal{C}}\Delta_{\Delta\subseteq\Omega}\langle[C]\rangle\Delta$
5. $\vdash \bigvee_{\Delta_C\subseteq\Omega, C\in\mathcal{C}}\bigwedge_{C\in\mathcal{C}}\langle[C]\rangle\Delta_C$

PROOF.

[1]Note that the $\langle[C]\rangle$ modality plays the same role wrt. $[C]$ as Levesque's [9] *only knowing* operator plays wrt. the traditional belief operator.

[2]Note that $\bigtriangledown_{p,q,r}$ is not the same as $p\nabla(q\nabla r)$!

1. Note that $\langle C\rangle\omega$ is just an atom in the coalitional language. Let $M = \{1,\dots,n\}$. Then, even by propositional reasoning, $\vdash \bigvee_{I\cup J=M, I\cap J=\emptyset}(\bigwedge_{i\in I}\langle C\rangle\omega_i \wedge \bigwedge_{j\in J}\neg\langle C\rangle\omega_j)$. Now, take a fixed I and J with $I\cap J=\emptyset$ and $I\cup J = M$. If $I = \emptyset$, then $\bigwedge_{i\in I}\langle C\rangle\omega_i \wedge \bigwedge_{j\in J}\neg\langle C\rangle\omega_j$ equals $\bigwedge_{\omega\in\Omega}\neg\langle C\rangle\omega$, which is equivalent to $\langle[C]\rangle\emptyset$. With Lemma 1 item 4, we have, for fixed I and J, each disjunct in this is equivalent to $\langle[C]\rangle\Delta_I$, with $\Delta_I = \{\omega_i \mid i\in I\}$.
2. Note that, using the notation of the previous item, we even have $\vdash \Delta_{I\cup J=M, I\cap J=\emptyset}(\bigwedge_{i\in I}\langle C\rangle\omega_i \wedge \bigwedge_{j\in J}\neg\langle C\rangle\omega_j)$. From this the statement follows directly.
3. This is immediate from item 1: if for an arbitrary C we have $\vdash \langle[C]\rangle\varphi_C$, then also $\vdash \bigwedge_{C\subseteq Z}\langle[C]\rangle\varphi_C$, for any $Z\in\mathcal{C}$.
4. Follows from item 2 in the same way as 3 follows from 1.
5. This follows immediately from item 1 and propositional reasoning: note that for every two coalitions C_1 and C_2 we derive $\vdash \bigvee_{\Delta_1\subseteq\Omega}\langle[C_1]\rangle\Delta_1 \wedge \bigvee_{\Delta_2\subseteq\Omega}\langle[C_2]\rangle\Delta_2$, and use $(p\vee q)\wedge r \equiv (p\wedge r)\vee(q\wedge r)$.

□

DEFINITION 1.

1. *For any agent i, we say that a formula is a preference literal for i if it is either $\omega \succeq_i \omega'$ or $\neg(\omega\succeq_i\omega')$, for some ω and ω'. We say that π_i fully describes i's preferences, if*

$$\pi_i \text{ is of the form } \bigwedge_{\omega,\omega'\in\Omega}(\neg)(\omega\succeq_i\omega')$$

We then say that $\pi\in PossPref(i)$.

2. *Given that we have m agents, a conjunction $\Pi = (\pi_1\wedge\cdots\wedge\pi_m)$, (where each π_i fully describes i's preferences) is said to fully describe the preferences of all the agents. Similarly for coalitions.*
3. *Recall that $\langle[C]\rangle\Delta$, where C is a coalition and Δ is a set of (atoms for) outcomes $\omega_1,\dots,\omega_u$, is said to fully describe C's choices. Now let $\mathcal{C} = 2^N\setminus\emptyset$, and let, for each $C\in\mathcal{C}$, Δ_C be a set of outcomes. Then $\bigwedge_{C\in\mathcal{C}}\langle[C]\rangle\Delta_C$ is said to fully describe all of N's choices. (Similarly for subsets of N.) We often will denote such a full description Ξ.*

LEMMA 3.

1. *Let $\Pi_1,\Pi_2,\dots,\Pi_d$ be all full descriptions of N's preferences. Then: $\vdash\bigvee_{k\le d}\Pi_k$*
2. *Let $\Pi_1,\Pi_2,\dots,\Pi_d$ be all full descriptions of N's preferences. Moreover, let $\mathcal{C}=2^N\setminus\emptyset$ and let*

$$(\bigwedge_{C\in\mathcal{C}}\langle[C]\rangle\Delta_C)_1,\dots,(\bigwedge_{C\in\mathcal{C}}\langle[C]\rangle\Delta_C)_z$$

enumerate all possible full descriptions of all choices of all coalitions (note that $z=(2^n)^m$). Then:

$$\vdash\bigvee_{k\le d, t\le z}(\Pi_k\wedge\bigwedge_{C\in\mathcal{C}}\langle[C]\rangle\Delta_C)_t$$

THEOREM 2. *Let φ be a formula of the cooperation language.*

1. *Let N be the set of agents, Ω the set of outcomes, and let $\mathcal{C} = 2^N\setminus\emptyset$. Then φ is equivalent to a formula of the form*

$$\bigvee_{k\in K, t\in T}(\Pi_k\wedge(\bigwedge_{C\in\mathcal{C}}\langle[C]\rangle\Delta_C)_t)$$

where each $\Pi_k = (\pi_1\wedge\cdots\wedge\pi_m)$ fully describes N's preferences, i.e., each π_i fully describes i's preferences, and each $(\bigwedge_{C\in\mathcal{C}}\langle[C]\rangle\Delta_C)_t$ describes fully what N can choose.

2. *The same holds if we take $\Omega = out(\varphi)$, $N = ag(\varphi)$ and we let C range over all $coal(\varphi)$.*

Complex as it may appear, our normal form is nothing more than an enumeration of possible full preferences combined with full descriptions of choices. The range of these possibilities is determined by the index sets K and T, which act as a kind of constraints: the smaller those index sets, the smaller the possible models for the formula. As a reading guide, note that

$$\bigvee_{k\in\{1,2\}, t\in\{a,b\}}(A_k\wedge B_t)$$

equals $(A_1\wedge B_a)\vee(A_1\wedge B_b)\vee(A_2\wedge B_a)\vee(A_2\wedge B_b)$.

PROOF. Note that the theorem is semantically obvious, the point is that we should be able to *syntactically prove* it from the sole axioms. This is done by induction over φ.

If $\varphi = \omega_1\succeq_k\omega_2$. Let us say that a π_k is k-compatible with $(\omega_1\succeq_k\omega_2)$ if the latter occurs as a conjunct in π_k. We then write $(\omega_1\succeq_k\omega_2)\in\pi_k$ Then, φ is (in propositional logic) equivalent to

$$\bigvee_{\pi_i\in PossPref(i)(i\neq k),(\omega_1\succeq_k\omega_2)\in\pi_k}(\pi_1\wedge\cdots\wedge\pi_k\wedge\cdots\wedge\pi_m)$$

This if of the form $\bigvee_{k\in K}\Pi_k$. Since by Lemma 2, item 5 we have that $\top$ is equivalent to $\bigvee_{\Delta_C\subseteq\Omega, C\in\mathcal{C}}\bigwedge_{C\in\mathcal{C}}\langle[C]\rangle\Delta_C$ we see that φ is equivalent to

$$\bigvee_{k\in K}\Pi_k\wedge\bigvee_{\Delta_C\subseteq\Omega, C\in\mathcal{C}}\bigwedge_{C\in\mathcal{C}}\langle[C]\rangle\Delta_C$$

which, by propositional reasoning, is equivalent to

$$\bigvee_{k\in K,\Delta_C\subseteq\Omega, C\in\mathcal{C}}\Pi_k\wedge\bigwedge_{C\in\mathcal{C}}\langle[C]\rangle\Delta_C$$

Suppose $\varphi = \langle E\rangle\varphi_0$. Formula φ_0 regards outcomes, and is equivalent to a disjunction $\bigvee\alpha$ where each α is of the form

$$\alpha = ((\neg)\omega_1\wedge(\neg)\omega_2\wedge\cdots\wedge(\neg)\omega_n)$$

Using rule *Nec* and axiom K, we derive $\vdash\langle E\rangle\varphi_0\leftrightarrow\langle E\rangle\bigvee\alpha$. Using that $\langle\cdot\rangle$ is a diamond, we then obtain that $\vdash\langle E\rangle\varphi_0\leftrightarrow\bigvee\langle E\rangle\alpha$. Now we use axiom *Func* to get rid of every α that contains more than one positive literal ω_i: let β range over all the those α's with at most one positive literal. Then $\vdash\langle E\rangle\varphi_0\leftrightarrow\bigvee\langle E\rangle\beta$. Now, again in propositional logic, note that every disjunction $\bigvee_{i\in M}\psi_i$ is equivalent to

$$\bigvee_{I\cup J=M, I\cap J=\emptyset}(\bigwedge_{i\in I}\psi_i\wedge\bigwedge_{j\in J}\neg\psi_j).$$

In our case, letting the β's range over $\beta_1,\dots,\beta_M$,

$$\vdash\langle E\rangle\varphi_0\leftrightarrow\bigvee_{I\cup J=M, I\cap J=\emptyset}(\bigwedge_{i\in I}\langle E\rangle\beta_i\wedge\bigwedge_{j\in J}\neg\langle E\rangle\beta_j)$$

But, for every fixed $I, J\subseteq M$, $(\bigwedge_{i\in I}\langle E\rangle\beta_i\wedge\bigwedge_{j\in J}\neg\langle E\rangle\beta_j)$ is equivalent to $\langle[E]\rangle\Delta_I$, with $\Delta_I = \{\omega_i \mid i\in I\}$. Hence, we find that $\vdash\langle E\rangle\varphi_0\leftrightarrow\bigvee_{I\subseteq M}\langle[E]\rangle\Delta_I$. This only limits the abilities of C, and not those of the others: using Lemma 2 item 5 once again:

$$\bigvee_{I\subseteq M}\langle[E]\rangle\Delta_I\leftrightarrow\bigvee_{I\subseteq M, D\neq E,\Delta_D\subseteq\Omega}\langle[E]\rangle\Delta_I\wedge\bigwedge_{D\in\mathcal{C}}\langle[D]\rangle\Delta_D$$

of which the r.h.s. is of the form $\bigvee_{C\in\mathcal{C}, t\in T}(\langle[C]\rangle\Delta_C)_t$, for some index set T. Since $\top$ is provably equivalent with $\bigvee_{k\in K}\Pi_k$ (where now K gives all possible full preference descriptions), the result follows.

Suppose φ is of the form $\varphi_1 \vee \varphi_2$. We can assume that $\varphi_i \equiv \bigvee_{k_i \in K_i, t_i \in T_i} (\Pi_{k_i} \wedge \Xi_{t_i})$, then

$$\varphi \equiv \bigvee_{k \in K_1 \cup K_2, t \in T_1 \cup T_2} (\Pi_k \wedge \Xi_t)$$

which is of the required form.

Let $\varphi = \neg\varphi_1$. Then $\varphi_1 = \bigvee_{k \in K, t \in T} (\Pi_k \wedge \Xi_t)$, for some index sets K and T. From Lemma 3, item 2, we know that $\vdash \bigvee_{k \leq d, t \leq z} (\Pi_k \wedge \Xi_t)$. In words: we know that a big disjunction is valid, but also that φ excludes some of them. Then, using propositional reasoning again, we obtain:

$$\varphi \equiv \bigvee_{k \leq d, k \notin K, t \leq z, t \notin T} (\Pi_k \wedge \Xi_t)$$

□

THEOREM 3 (COMPLETENESS). *We have, for all $\varphi \in \mathcal{L}_c$ over N and Ω: $\models \varphi \Rightarrow \vdash \varphi$*

PROOF. Suppose $\not\vdash \varphi$, i.e., $\neg\varphi$ is consistent. We know that

$$\vdash \neg\varphi \leftrightarrow \bigvee_{k \in K, t \in T} (\Pi_k \wedge (\bigwedge_{C \in \mathcal{C}} \langle [C] \rangle \Delta_C)_t)$$

for some index sets K and T. Call the righthandside of this equivalence φ'. Let X be a maximal consistent set around φ'. By virtue of maximal consistent sets, we know that for some k and t, $\varphi'' = (\Pi_k \wedge (\bigwedge_{C \in \mathcal{C}} \langle [C] \rangle \Delta_C)_t) \in X$. But now we can read off the game $\Gamma = \langle N, \Omega, V, \sqsupseteq_1, \ldots, \sqsupseteq_m \rangle$ from φ'' immediately:

1. N and Ω are already given;
2. Let $V(C) = \Delta$, where $\langle [C] \rangle \Delta$ is part of φ'
3. every $\sqsupseteq_i$ relation is immediately read off from the component π_i for Π_k in φ''.

Now, it is easy to see, for every subformula ψ of φ'':

$$\Gamma \models \psi \Leftrightarrow \psi \in X$$

□

4. MODEL CHECKING & SATISFIABILITY

It is trivial to see that the *model checking problem* for CGL (i.e., the problem of determining, for any given Γ and φ, whether or not $\Gamma \models \varphi$ [4]) may be solved in deterministic polynomial time: an obvious recursive algorithm for this problem can be directly extracted from the semantic rules of the language.

The *satisfiability problem* is the problem of checking whether or not, for any given φ there exists a game Γ such that $\Gamma \models \varphi$. For most modal logics, the corresponding satisfiability problem has a trivial NP-hard lower bound, since such logics subsume propositional logic, for which satisfiability is the defining NP-complete problem [3, p.374]. However, our logic is specialised for reasoning about coalitional games, and it is not so obvious that it subsumes propositional logic, since we do not have primitive propositions. We must therefore prove NP-hardness from first principles.

For the proof, we need a few additional constructions. A *partial coalitional game* is a structure $\langle N, \Omega, V, \sqsupseteq_1, \ldots, \sqsupseteq_m \rangle$ where all the components are as in regular coalitional games, except that V is a *partial* function, i.e., it is not required to be defined for every possible coalition. Given a partial game $\Gamma = \langle N, \Omega, V, \sqsupseteq_1, \ldots, \sqsupseteq_m \rangle$, we can use the semantic rules for CGL to interpret some formulae (although because V is not defined for all coalitions, we cannot necessarily interpret all formulae over N, Ω). Where Γ is a partial game and φ is a formula, let us write $\Gamma \models_p \varphi$ to mean that (i) it is possible to evaluate φ with respect to Γ, and (ii) φ is true under this evaluation. Now, we can prove the following.

LEMMA 4. *A CGL formula φ is satisfiable iff there exists a partial game $\Gamma = \langle N, \Omega, V, \sqsupseteq_1, \ldots, \sqsupseteq_m \rangle$ such that:*

1. *$N = ag(\varphi)$,*
2. *$|\Omega| = |out(\varphi)| + 1$ and $out(\varphi) \subseteq \Omega$;*
3. *dom V = coal(φ), and*
4. *$\Gamma \models_p \varphi$.*

PROOF. *The right-to-left direction is obvious, so consider the left-to-right direction, and let $\Gamma = \langle N, \Omega, V, \sqsupseteq_1, \ldots, \sqsupseteq_m \rangle$ be a game such that $\Gamma \models \varphi$. Let $A = \Omega \setminus out(\varphi)$, i.e., A is the set of outcomes in Γ not named in φ. Let ω^* be an outcome such that $\omega^* \notin \Omega$, and define a partial game $\Gamma^* = \langle N^*, \Omega^*, V^*, \sqsupseteq_1^*, \ldots, \sqsupseteq_o^* \rangle$ as follows:*

- *$N^* = ag(\varphi)$;*
- *$\Omega^* = out(\varphi) \cup \{\omega^*\}$;*
- *The relation $\sqsupseteq_i^*$ is obtained by first restricting $\sqsupseteq_i$ to $out(\varphi)$, and then defining $\omega^* \sqsupseteq_i \omega$ for all $\omega \in out(\varphi)$;*
- *V^* is the partial function such that V^* is only defined for coalitions named in φ (i.e., $C \in$ dom V^* iff $C \in coal(\varphi)$);*
- $V^*(C) = \begin{cases} V(C) & \text{if } V(C) \subseteq out(\varphi) \\ (V(C) \setminus A) \cup \{\omega^*\} & \text{otherwise.} \end{cases}$

Notice that Γ^* satisfies conditions (1)–(3) of the lemma. We now prove that Γ^* satisfies condition (4). More precisely, we show that for all sub-formulae ψ of φ: $\Gamma \models \psi$ iff $\Gamma^* \models \psi$. The inductive base is where $\varphi = (\omega_1 \succeq_i \omega_2)$, and is obvious, since $i \in ag(\varphi)$ and $\{\omega_1, \omega_2\} \subseteq out(\varphi)$, and hence $\omega_1 \sqsupseteq_i^* \omega_2$ iff $\omega_1 \sqsupseteq_i \omega_2$.

For the inductive assumption, assume the result is proved for all sub-formulae; in the inductive step, the significant case is where $\varphi = \langle C \rangle \psi$. If $\Gamma \models \langle C \rangle \psi$ then $\exists \omega \in V(C)$ such that $\Gamma, \omega \models \psi$. There are two possibilities: either $\omega \in out(\varphi)$ (in which case $V^*(C) = V(C)$, and the result is obvious), or else $\omega \notin out(\varphi)$. In the latter case, $V^*(C) = (V(C) \setminus A) \cup \{\omega^*\}$; we claim that $\Gamma^*, \omega^* \models \psi$. To see this, assume w.l.o.g. that ψ is in Conjunctive Normal Form. Now, since $\omega \notin out(\varphi)$, then no positive literals can be satisfied by Γ, ω: only negative literals. But such literals must also be satisfied by Γ^*, ω^*, and so $\Gamma^*, \omega^* \models \psi$.

The case for $\Gamma \not\models \langle C \rangle \psi$ implies $\Gamma^* \not\models \langle C \rangle \psi$ is similar. □

Given this, we can prove:

THEOREM 4. *The satisfiability problem for CGL formulae is NP-complete, even for CGL formulae φ such that $|ag(\varphi)| = 1$.*

PROOF. For membership of NP, we know that φ is satisfiable iff it has a "certificate" for this in the form of a partial game Γ as in Lemma 4. This partial game is of size linear in the size of the formula φ. Since we can check whether $\Gamma \models_p \varphi$ in polynomial time, we conclude that CGL satisfiability is in NP.

For NP-hardness, we reduce SAT, the problem of determining whether a formula $\varphi(x_1, \ldots, x_k)$ of propositional logic, over Boolean variables $x_1, \ldots, x_k$, has a satisfying assignment [12]. The basic idea is to map variables x_i to outcomes ω_i, to introduce an additional outcome $\omega_\perp$ to correspond to the truth value "false", so that $(\omega_x \succ_1 \omega_\perp)$ will mean "x takes the value 'true' ". Formally, let $\varphi^\#$ denote the CGL formula obtained from the propositional logic formula φ by systematically replacing every Boolean variable p by the corresponding CGL expression $(\omega_p \succ_1 \omega_\perp)$. Now, we claim that $\varphi^\#$ is CGL satisfiable iff the input SAT instance φ is a satisfiable formula of propositional logic.

For the $\Rightarrow$ direction, assume $\varphi^\#$ is CGL satisfiable, and consider the associated preference relation $\sqsupseteq_1$ in any Γ such that $\Gamma \models \varphi$. From this relation, extract a valuation ξ for the variables $x_1, \ldots, x_k$

as follows: each variable x_i is true under ξ if $\omega_{x_i} \sqsupset_1 \omega_\perp$, and false otherwise. The interpretation ξ is consistent, since we cannot have both $\omega_{x_i} \sqsupset_1 \omega_\perp$ and $\omega_\perp \sqsupseteq_1 \omega_{x_i}$. The interpretation ξ satisfies φ by a trivial induction on the structure of φ.

For $\Leftarrow$, assume φ is a satisfiable formula of propositional logic, and let ξ be a valuation that satisfies φ. Then we can reconstruct a game Γ_ξ such that $\Gamma_\xi \models \varphi^\#$, as follows. Γ_ξ contains a single agent, (agent 1), and an outcome ω_{x_i} for each variable x_i appearing in φ. We also define an additional outcome $\omega_\perp$. The preference relation $\sqsupseteq_i$ is then defined as follows:

- For each Boolean variable p such that p is true under ξ, define $\omega_p \sqsupset_i \omega_\perp$.
- For each pair of Boolean variables p_1, p_2 such that p_1 and p_2 are both true or both false under ξ, define $\omega_{p_1} =_i \omega_{p_2}$.
- For each Boolean variable p such that p is false under ξ, define $\omega_\perp \sqsupseteq_i \omega_p$.
- For each pair of Boolean variables p_1, p_2 such that p_1 is true (respectively, false) and p_2 is false (respectively, true) under ξ, define $\omega_{p_1} \sqsupset_i \omega_{p_2}$ (respectively, $\omega_{p_2} \sqsupset_i \omega_{p_1}$).

An induction on $\varphi^\#$ proves that $\Gamma_\xi \models \varphi^\#$. $\Box$

5. CHARACTERISING COALITIONAL GAMES

We characterise three solution concepts from the theory of coalitional games, viz. *the core* [7], *stable sets* [10] and *the bargaining set* [2], in CGL. We use the formulations of these solution concepts in [11]; there the two latter solution concepts are however defined only for games with real numbered payoffs and transferable utility and below we translate the definitions to the more general games with preference relations over general outcomes and non-transferable utility.

Henceforth, a *C-feasible outcome* is an outcome which can be chosen by the coalition C and a *feasible outcome* is a N-feasible outcome where N is the set of players in the game. We start by looking at the core, which is a, possibly empty, set of outcomes.

DEFINITION 2 (CORE). *The core of a coalitional game is the feasible outcomes ω for which there is no coalition C with a C-feasible outcome ω' such that $\omega' \succ_i \omega$ for all $i \in C$.*

We write $CM(\omega)$ to mean that ω is in the core.

$$CM(\omega) \equiv \langle N \rangle \omega \wedge \neg \left[\bigvee_{C \subseteq N} \bigvee_{\omega' \in \Omega} (\langle C \rangle \omega') \wedge \bigwedge_{i \in C} (\omega' \succ_i \omega) \right]$$

CNE will then mean that the core is non-empty:

$$CNE \equiv \bigvee_{\omega \in \Omega} CM(\omega)$$

THEOREM 5. *The core of Γ is non-empty iff $\Gamma \models CNE$.*

A *stable set* is a set of outcomes. A coalitional game may have several stable sets, but must not necessarily have any. We characterize stable sets in terms of *imputations* and *objections*. An imputation is a feasible outcome that for each agent i is as least as good as any outcome the singleton coalition $\{i\}$ can choose on his own. The CGL formula $IMP(\omega)$ is true whenever ω is an imputation:

$$IMP(\omega) \equiv \langle N \rangle \omega \wedge \bigwedge_{\omega' \in \Omega} \bigwedge_{i \in N} (\langle \{i\} \rangle \omega' \rightarrow \omega \succeq_i \omega')$$

An imputation ω is a C-objection to an imputation ω' if every agent in C prefers ω over ω' and the coalition C can choose an outcome which for every agent in C is as least as good as ω. ω is an objection to ω' if ω is a C-objection to ω' for some coalition C. Next, $OBJ(\omega, \omega', C)$ expresses that outcome ω is an C-objection to outcome ω', when both ω and ω' are imputations:

$$OBJ(\omega, \omega', C) \equiv (\bigwedge_{i \in C} \omega \succ_i \omega') \wedge \bigvee_{\omega'' \in \Omega} (\langle C \rangle \omega'' \wedge \bigwedge_{i \in C} \omega'' \succeq_i \omega)$$

DEFINITION 3 (STABLE SET). *A set of imputations Y is a stable set if it satisfies:*

Internal stability *If $\omega \in Y$, there is no objection to ω in Y.*

External stability *If $\omega \notin Y$, there is an objection to ω in Y.*

Given a set of outcomes $Y \subseteq \Omega$, the CGL formula $STABLE(Y)$ expresses the fact that Y is a stable set:

$$\begin{aligned} &STABLE(Y) \equiv \\ &\quad \textstyle\bigwedge_{\omega \in Y} IMP(\omega) \\ &\quad \wedge \left(\textstyle\bigwedge_{\omega \in Y} \bigwedge_{C \subseteq N} \bigwedge_{\omega' \in Y} \neg OBJ(\omega', \omega, C) \right) \\ &\quad \wedge \left(\textstyle\bigwedge_{\omega \in \Omega \setminus Y} IMP(\omega) \rightarrow \left(\bigvee_{C \subseteq N} \bigvee_{\omega' \in Y} OBJ(\omega', \omega, C) \right) \right) \end{aligned}$$

THEOREM 6. *Y is a stable set of Γ iff $\Gamma \models STABLE(Y)$.*

PROOF. Given a coalitional game, let $\mathcal{I}$ denote the set of all imputations. First, we argue that $IMP(\omega)$ and $OBJ(\omega, \omega', C)$ have the correct meaning. Every $\omega \in Y$ is an imputation iff $\omega \in V(N)$ (feasibility) and $\omega \succeq_i \omega'$ for all i and $\omega' \in V(\{i\})$ which is exactly when $IMP(\omega)$ holds. If $\omega, \omega' \in \mathcal{I}$ and $C \subseteq N$, ω is a C-objection to ω' iff $\omega \succ_i \omega'$ for every $i \in C$ and there is a $\omega'' \in V(C)$ such that $\omega'' \succeq_i \omega$ for every $i \in C$, which is exactly when $OBJ(\omega, \omega', C)$ holds.

For the main proof, let $Y \subseteq \Omega$. If there is an ω in Y which is not an imputation, Y is not a stable set and $IMP(\omega)$ is not true and we are done, so assume that Y is a set of imputations. Let

$$\hat{Y} = \{\omega \in \mathcal{I} : \text{ there is no objection to } \omega \text{ in } Y\}$$

It is easy to see that if Y is a stable set iff $Y = \hat{Y}$. We argue that the second and third main conjuncts of the formula $STABLE(Y)$ is true whenever $Y \subseteq \hat{Y}$ and $\hat{Y} \subseteq Y$ hold, respectively, and the theorem follows (the first conjunct is true under the assumption that Y are imputations). The second conjunct is true exactly when for every member of Y there is no C-objection to ω in Y for any C, which is exactly when $Y \subseteq \hat{Y}$ holds. The third conjunct is true iff every imputation which is not in Y has an objection in Y or, contrapositively, that every imputation which does not have an objection in Y is included in Y which is the same as $\hat{Y} \subseteq Y$. $\Box$

Existence of a stable set can then be expressed as:

$$ES \equiv \bigvee_{Y \subseteq \Omega} STABLE(Y)$$

COROLLARY 1. *Γ has a stable set iff $\Gamma \models ES$.*

Finally, we focus on the notion of a bargaining set of a coalitional game which is, like a stable set, a set of imputations, but, unlike a stable set, is unique and always exists. The bargaining set of a game can be defined in terms of *objections* and *counterobjections*, but the former concept is not the same as in the definition of stable sets. Let ω be an imputation:

Objection: A pair consisting of a coalition C and a C-feasible outcome ω' is an objection of an agent $i \in C$ against an agent $j \notin C$ to ω if every agent in C prefers ω' over ω.

Counterobjection: A pair consisting of a coalition D and a D-feasible outcome v is a counterobjection to an objection (ω', C) of i against j to ω, if D includes j but not i, every agent in $D \setminus C$ thinks v is as least as good as ω and every agent in $D \cap C$ thinks v is as least as good as ω'.

The CGL formula $OBJB(\omega', C, \omega)$ means that (ω', C) is an objection of any $i \in C$ against any $j \notin C$ to ω.

$$OBJB(\omega', C, \omega) \equiv \langle C \rangle \omega' \wedge \bigwedge_{k \in C} \omega' \succ_k \omega$$

$ECO(\omega', C, i, j, \omega)$ means that there exists a counterobjection to the objection (ω', C) of i against j to ω.

$$ECO(\omega', C, i, j, \omega) \equiv \bigvee_{v \in \Omega} \bigvee_{D' \subseteq N \setminus \{i\}} (\langle D' \cup \{j\} \rangle v \wedge \Big(\Big(\bigwedge_{k \in (D' \cup \{j\}) \setminus C} v \succeq_k \omega \Big) \wedge \Big(\bigwedge_{k \in (D' \cup \{j\}) \cap C} v \succeq_k \omega' \Big) \Big))$$

DEFINITION 4 (BARGAINING SET). *The bargaining set of a coalitional game is the set of all imputations ω such that there exists a counterobjection to every objection of any player i against any player j to ω.*

$INBARG(\omega)$ means that outcome $\omega \in \Omega$ is in the bargaining set:

$$INBARG(\omega) \equiv IMP(\omega) \wedge \bigwedge_{C \subseteq N} \bigwedge_{i \in C} \bigwedge_{j \in N \setminus C} \bigwedge_{\omega' \in \Omega} [OBJB(\omega', C, \omega) \rightarrow ECO(\omega', C, i, j, \omega)]$$

THEOREM 7. *ω is a member of the bargaining set of Γ iff $\Gamma \models INBARG(\omega)$.*

PROOF. It is easy to see that when ω is an imputation, C is a coalition, $i \in C, j \notin C$ and ω' an outcome, $\Gamma \models OBJB(\omega', C, \omega)$ iff (ω', C) is an objection of i against j to ω. To see that there exist a counterobjection to the objection (ω', C) of i against j to ω iff $\Gamma \models ECO(\omega', C, i, j, \omega)$, observe that (v, D) is a counterobjection iff the disjunct given by v and $D' = D \setminus \{j\}$ is true. The theorem follows immediately. □

We can now define $BS(Y)$, $Y \subseteq \Omega$, to express the fact that Y is the bargaining set.

$$BS(Y) = \bigwedge_{\omega \in Y} INBARG(\omega) \wedge \bigwedge_{\omega \in \Omega \setminus Y} \neg INBARG(\omega)$$

COROLLARY 2. *Y is the bargaining set of Γ iff $\Gamma \models BS(Y)$*

COROLLARY 3. *$\neg BS(\emptyset)$ is a theorem of* CGL.

PROOF. For any coalitional game without transferable payoff Γ, the bargaining set is non-empty [6, 15]. By Corollary 2, $\Gamma \models \neg BS(\emptyset)$ for any Γ, and by Theorem 3 $\vdash \neg BS(\emptyset)$. □

6. RELATION TO COALITION LOGIC

As we noted in section 2, it is rather tempting to believe that the outcomes of coalitional games can be interpreted as states, and that the characteristic function can be interpreted as an effectivity function, and that as a consequence Coalition Logic (CL) [14] could be interpreted directly in coalitional games. In this section, we compare the semantics of CL and CGL, and show that in fact there is a fundamental difference between the two approaches.

We first give a very brief review of some of the concepts of CL. A *coalition model* for agents N over a set of atomic propositions Φ_0 is a triple $M = (S, E, \pi)$, where S is a nonempty set of states, $\pi : \Phi_0 \rightarrow 2^S$ an assignment and E gives a function of type $2^N \rightarrow 2^{2^S}$ for each state $s \in S$. It is required that each $E(s)$ is a *playable effectivity function*, satisfying the following conditions: i) $\forall_{C \subseteq N} \emptyset \notin E(s)(C)$, ii) $\forall_{C \subseteq N} S \in E(s)(C)$, iii) for any X, if $S \setminus X \notin E(s)(N \setminus C)$ then $X \in E(s)(C)$ (N*-maximality*), iv) for all $X \subseteq X' \subseteq S$ and all C, if $X \in E(s)(C)$ then $X' \in E(s)(C)$ (*outcome monotonicity*) and v) if $C_1 \cap C_2 = \emptyset$, $X_1 \in E(s)(C_1)$ and $X_2 \in E(s)(C_2)$, then $X_1 \cap X_2 \in E(s)(C_1 \cup C_2)$ (*superadditivity*). Formulae of coalition logic, and their satisfaction in states s of coalition models M, are defined as follows[3]:

$$M, s \models \langle C \rangle \psi \quad \text{iff } \psi^M \in E(s)(C)$$

where $\psi^M = \{s \in S : M, s \models \psi\}$. If we take the set of atomic propositions to be $\Phi_0 = \Omega \cup \{\omega \succeq_i \omega' : \omega, \omega' \in \Omega, i \in N\}$, then we can read every formula in $\mathcal{L}_o \cup \mathcal{L}_c$ as a formula of coalition logic. Thus we can interpret $\mathcal{L}_c$ formulae in both a game Γ and in a pointed coalition model (M, s), and a $\mathcal{L}_o$ formula in both a pointed game (Γ, ω) and in a pointed model (M, s). Coalitional games and coalition models have many similarities. The former have "outcomes" while the latter have "states". An interesting question is: given a coalitional game Γ, does there exist an equivalent coalition model M with states corresponding to the outcomes of Γ, maybe in addition to a designated "initial" state t? Equivalence here means that Γ and the pointed coalition model (M, t) agree on $\mathcal{L}_c$ formulae and (Γ, ω) and (M, ω) agree on $\mathcal{L}_o$ formulae for any outcome ω. We can say that Γ and M then are *outcome-equivalent*.

In other words, a coalitional game Γ and a coalition model M, defined over atomic propositions Φ_0 above and having states $\Omega \cup \{t\}$, are outcome-equivalent iff for any $\varphi_1 \in \mathcal{L}_c$, any $\varphi_0 \in \mathcal{L}_o$ and any $\omega \in \Omega$: we have both (a) $\Gamma \models \varphi_1$ iff $M, t \models \varphi_1$, and (b) $\Gamma, \omega \models \varphi_0$ iff $M, \omega \models \varphi_0$.

A natural question then is: given a game, does there exist an outcome-equivalent coalition model?

The answer, given by the following theorem, is "no", except for for certain special cases of games. The latter is the class of games where $V(C) = \{\omega\}$ for all coalitions $C \neq N$, for some fixed outcome $\omega \in \Omega$. To give them a name, we will call such games *limited games*, since, first, most games are not of this kind and, second, they are not very interesting. The only coalition in a limited game which possibly can select an outcome different from the fixed outcome ω is the grand coalition.

THEOREM 8. *No non-limited coalitional game with more than one player has an outcome-equivalent coalition model.*

PROOF. Let $\Gamma = \langle N, \Omega, V, \sqsupseteq_1, \ldots, \sqsupseteq_m \rangle$ be a coalitional game, and assume that $M = (S, E, \pi)$ with $S = \Omega \cup \{t\}$ is a coalition model outcome-equivalent to Γ. We argue, by using the properties i) – v) of a playable effectivity function given above, that Γ must be limited. First, observe that for any $\mathcal{L}_o$ formula φ and coalition $C \neq \emptyset$, $\Gamma \models \langle C \rangle \varphi$ iff, by (a), $M, t \models \langle C \rangle \varphi$, i.e.

$$\varphi^M \in E(t)(C) \Leftrightarrow \exists_{\omega \in V(C)} \Gamma, \omega \models \varphi \tag{1}$$

for any $\varphi \in \mathcal{L}_0$, $C \subseteq N$, $C \neq \emptyset$.

Observe that $V(C) \neq \emptyset$ for any coalition $C \neq \emptyset$. This follows from ii) and (1): $S = \top^M \in E(t)(C)$, where $\top$ is some tautology in the $\mathcal{L}_0$ language (e.g., $\omega' \vee \neg\omega'$), so $\exists_{\omega \in V(C)} \Gamma, \omega \models \top$, which ensures that $V(C)$ is non-empty.

We show that for any coalition $C \neq \emptyset$ and any $\omega \in \Omega$

$$\omega \in V(C) \Leftrightarrow \{\omega\} \in E(t)(C) \tag{2}$$

[3]We only give the coalitional clause. Pauly [14] uses $[C]$ where we use $\langle C \rangle$; here we use the latter notation for easier comparison.

For the direction to the right, assume that $\omega \in V(C)$. $\Gamma, \omega \models \omega$ (note the dual role of ω as both an outcome and a formula), so $\omega^M \in E(t)(C)$ by (1). By (b), $M, \omega \models \omega$, so $\omega \in \omega^M$. If $\omega' \in \omega^M$, then $M, \omega' \models \omega$, $\Gamma, \omega' \models \omega$ by (b) and $\omega' = \omega$. Thus, $\{\omega\} = \omega^M \in E(t)(C)$. For the direction to the left, let $\{\omega\} \in E(t)(C)$. Since, again, $\omega^M = \{\omega\}$, by (1) there is a $\omega' \in V(C)$ such that $\Gamma, \omega' \models \omega$. This is only the case when $\omega = \omega' \in V(C)$.

For any non-empty disjoint coalitions C_1 and C_2:

$$(\omega_1 \in V(C_1) \text{ and } \omega_2 \in V(C_2)) \Rightarrow \omega_1 = \omega_2 \quad (3)$$

$$V(C_1) = V(C_2) \quad (4)$$

$$|V(C_1)| = 1 \quad (5)$$

We prove (3)–(5):

(3) Assume otherwise, that $\omega_1 \in V(C_1)$ and $\omega_2 \in V(C_2)$ and $\omega_1 \neq \omega_2$. By (2), $\{\omega_1\} \in E(t)(C_1)$ and $\{\omega_2\} \in E(t)(C_2)$, and by superadditivity it must be the case that $\emptyset = \{\omega_1\} \cap \{\omega_2\} \in E(t)(C_1 \cup C_2)$, but this contradicts i). Thus, (3) must hold.

(4) Assume that $\omega \in V(C_1)$; we show that $\omega \in V(C_2)$. Since $V(C_2)$ is non-empty, let $\omega' \in V(C_2)$. By (3), $\omega' = \omega$. Thus $\omega \in V(C_2)$. By a symmetric argument, $\omega \in V(C_2)$ implies that $\omega \in V(C_2)$.

(5) Since $V(C_1)$ is non-empty, there is an $\omega_1 \in V(C_1)$. If $\omega_2 \in V(C_1)$, then $\omega_2 \in V(C_2)$ by (4) and $\omega_1 = \omega_2$ by (3). Thus, $V(C_1) = \{\omega_1\}$.

Let $a, b \in N$ such that $a \neq b$ (existence is ensured by the assumption of more than one player). By (4) and (5) there is an ω_1 such that $V(\{a\}) = V(\{b\}) = \{\omega_1\}$. For any $d \in N$ such that $d \neq a$ and $d \neq b$, $\{a\}$ and $\{d\}$ are disjoint and we again get that $V(\{d\}) = V(\{a\}) = \{\omega_1\}$. Thus, $V(\{d\}) = \{\omega_1\}$ for any $d \in N$. Let $C \subset N$ be a coalition different from the grand coalition. There is a $d \in N$ such that C and $\{d\}$ are disjoint, so by (4) $V(C) = V(\{d\}) = \{\omega_1\}$, which shows that Γ is limited. □

Thus, in general, a coalitional game is *not* simply a coalition model with outcomes as states. Even though the language of Coalition Logic is similar to the language of our logic, it follows from Theorem 8 that we cannot use the semantic rules of Coalition Logic "directly" to say whether a formula is true or not in a coalitional game. The main reason is that a difference between outcomes in coalitional games and states in coalition models is that an outcome is *local* to the coalition which chooses it, while states are *global*. As a consequence, while it is perfectly possible in a coalitional game that both a coalition C can choose outcome ω ($\omega \in V(C)$) *and* a coalition C', C' and C disjoint, can choose outcome ω' ($\omega' \in V(C')$) when $\omega' \neq \omega$, it is not possible in a coalition model that both C is effective for $\{\omega\}$ and C' is effective for $\{\omega'\}$. The proof of Theorem 8 shows that in general there is no playable effectivity function corresponding, in the sense of (2), to a characteristic function.

7. CONCLUSIONS AND FUTURE WORK

We introduced a knowledge representation language for coalitional games without transferable payoffs, and a logic specifically intended for reasoning about such games. We presented the logic, gave a complete axiomatization for it, showed it was expressively complete with respect to coalitional games without transferable payoff, showed that the satisfiability problem was NP-complete, showed how the logic could be used to capture a range of solution concepts for coalitional games, and finally, showed formally why the logic was fundamentally different to existing cooperation logics. Other solution concepts for coalitional games [11], for which the definitions in terms of payoffs cannot be trivially extended to the more general case of preferences and non-transferable utility, are the *the kernel*, *the nucleoulus* and *the Shapley value*. Logical characterisation of these concepts, in addition to further comparison with Coalition logic, could be interesting future work.

Acknowledgements

The research reported in this paper was carried out when the first author was visiting the Department of Computer Science, University of Liverpool, supported by grant 166525/V30 from the Norwegian Research Council.

8. REFERENCES

[1] R. Alur, T. A. Henzinger, and O. Kupferman. Alternating-time temporal logic. *Journal of the ACM*, 49(5):672–713, September 2002.

[2] R. J. Aumann and M. Maschler. The bargaining set for cooperative games. *Advances in Game Theory (Annals of Mathematics Studies 52)*, pages 443–467, 1964.

[3] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press: Cambridge, England, 2001.

[4] E. M. Clarke, O. Grumberg, and D. A. Peled. *Model Checking*. The MIT Press: Cambridge, MA, 2000.

[5] P. R. Cohen and H. J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42:213–261, 1990.

[6] M. Davis and M. Maschler. Existence of stable payoff configurations for cooperative games. *Bulletin of the American Math. Soc.*, 69:106–108, 1963.

[7] D. B. Gillies. Solutions to general non-zero-sum games. *Contribution to the Theory of Games (Annals of Mathematics Studies, 40)*, IV:47–85, 1959.

[8] V. Goranko. Coalition games and alternating temporal logics. In J. van Benthem, editor, *Proceeding of the Eighth Conference on Theoretical Aspects of Rationality and Knowledge (TARK VIII)*, pages 259–272, Siena, Italy, 2001.

[9] H. J. Levesque. All I know: a study in autoepistemic logic. *Artificial Intelligence*, 42:263–309, 1990.

[10] J. Von Neumann and O. Morgenstern. *Theory of Games and Economic Behaviour*. Princeton University Press: Princeton, NJ, 1944.

[11] M. J. Osborne and A. Rubinstein. *A Course in Game Theory*. The MIT Press: Cambridge, MA, 1994.

[12] C. H. Papadimitriou. *Computational Complexity*. Addison-Wesley: Reading, MA, 1994.

[13] M. Pauly. *Logic for Social Software*. PhD thesis, University of Amsterdam, 2001. ILLC Dissertation Series 2001-10.

[14] M. Pauly. A modal logic for coalitional power in games. *Journal of Logic and Computation*, 12(1):149–166, 2002.

[15] B. Peleg. Existence theorem for the bargaining set $m_1^{(i)}$. *Bulletin of the American Math. Soc.*, 69:109–110, 1963.

[16] W. van der Hoek, W. Jamroga, and M. Wooldridge. A logic for strategic reasoning. In *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2005)*, pages 157–153, Utrecht, The Netherlands, 2005.