It's all up to you:
A study of closed-world interaction and its optimal organization

Jan Broersen, Universiteit Utrecht, broersen@cs.uu.nl
Rosja Mastop, Universiteit Utrecht, rosja.mastop@phil.uu.nl
John-Jules Ch.Meyer, Univeristeit Utrecht, jj@cs.uu.nl
Paolo Turrini, Universiteit Utrecht, paolo@cs.uu.nl

Abstract

The aim of the work is to provide a deontic language to regulate closed-world interaction. To do so we use Coalition Logic enriched with a preference order over the outcomes of agents' choices. We take the perspective of a deontic language being agent-oriented, that is mandating choices that only belong to agents or coalitions. We formalize this intuition by identifying those interactions in which Nature does not play an active role. We apply the formal tools to games.

1 Introduction

Pauly's Coalition Logic has shown to be a sound formal tool to analyze the properties of strategic interactions. One issue left is to define in that language what the interesting properties of an interaction are, as possible for instance with regularity (it is never the case that a group of agents can determine that some variable p is true, while all the other agents can at the same time determine that p is false) or outcome monotonicity (if a coalition can force an outcome to lie in a set X, can also force an outcome to lie in all supersets of X).

If we think of a deontic logic as obligating agents to choose what it should ideally be the case, an intuitive property is that of *coherence*, a property of interaction that ensures players' abilities non to contradict one other and the empty coalition not to make active choices. With this property we can model a closed world interaction, such as those of a Coordination Game or of a Prisoner Dilemma, where all the outcomes are determined only by the choices of the agents that are present.

Our aim is to regulate multiagent interaction, mandating the optimal outcomes that result from the choices of the coalitions. By mandating we mean *the introduction of a normative constraint on individual and collective choices in a multiagent system.*

Row \ Column	White Dress	Black Dress
White Dress	$(3,3)$	$(0,0)$
Black Dress	$(0,0)$	$(3,3)$

Table 1. Clothing Conformity

We are specifically concerned with cases where the collective perspective is at odds with the individual perspective. That is, cases where we think that letting everybody pick their own best action regardless of other's interest gives a non-optimal result. The main question we are dealing with is then: how do we determine which norms, if any, are to be imposed?

To answer this question, the paper presents a language to talk about the conflict between coalitionally optimal and socially optimal choices in coherent interaction, and it expresses deontic notions referring to such circumstances.

1.0.1 Example

The toy example we would like to start with concerns conventional norms. Noms of this type are those in which players should conform to each other. In this situation (see Table 1), a legislator that wants to achieve the socially optimal state (players coordinate), should declare that discordant choices are forbidden, thereby labeling the combinations of moves (black, white), (white, black) as violations. As easy to see, these moves belong only to the set of agents taken together. A norm helping both players to reach an optimal outcome would be one that labels as violations combinations of discordant choices. However, in this kind of games Row will never know what is the best thing to choose, since the choice of Column is independent from his. In order to solve the problem a legislation should go beyond individual choice, by forcing the coalition made of Row and Column together to form and choose an efficient outcome.

Row \ Column	White Dress
White Dress	$(3, 3)$
Black Dress	$(0, 0)$

Table 2. Clothing Conformity Modified

1.1 Motivation

Provided the aim of regulating interactions, we ask ourselves whether it makes sense to construct a deontic logic for any type of game.

Suppose the environment (the coalition made by an empty set of agents) were active part of the game, and it could decide to transform the game of table 1 in the one of table 2.

What should then a legislator do? It is quite clear that imposing the agents to choose something should depend on the moves that are available to the players. But in a game in which Nature plays an active role, taking this statement serious would boil down to mentioning the environment in the deontic language, saying for instance "Nature should allow row to play only white" or "Nature should make it convenient for the grand coalition to form". If we think of a deontic language as a sort of "agent-oriented" language and as nature as a uncontrollable agent, the above mentioned statements do not make sense.

No legislator though would be in the condition of determining what moves Nature would play. Nature, unlike all the other players, does not have explicit preferences over the outcomes of the interaction and intuitively does not follow proper man made norms or orders. In order to have a regulation of the Multi Agent System, we need a proper agent-oriented deontic language and we should then avoid deontic statements that concern proper choices (i.e. those able to really modify the outcome of the game) to be carried out by Nature. This translates into ruling out all those interactions in which Nature plays an active role. In this paper we will pursue this idea formally, identifying all such interactions and axiomatizing their logic.

The paper is structured as follows: In the first part we introduce Coherent Coalition Logic, proving that Inability Of the Empty Coalition (IOEC) is not entailed by Pauly playable effectivity functions and it cannot even be defined in Coalition Logic. In the second part we discuss the axiomatization of the logic, giving a characterization of coherence in terms of global modality. In the third part we give application of the logic to the regulation of closed-world strategic interaction, constructing a deontic logic that tells coalitions how to behave in order to achieve socially desirable outcomes.

2 Coherent Interactions

We begin by defining the strategic abilities of agents and coalitions, introducing the concept of a dynamic Effectivity Function, adopted from [7]. Later on in the paper we will move from game forms to real games, by introducing the notion of preference.

Definition 2.1 [Dynamic Effectivity Function]
Given a finite set of agents Agt and a set of states W, a *dynamic Effectivity Function* is a function $E : W \rightarrow (2^{Agt} \rightarrow 2^{2^W})$.

◁

Any subset of Agt will henceforth be called a *coalition*.

For elements of W we use variables $u, v, w, \ldots$; for subsets of W we use variables $X, Y, Z, \ldots$; and for sets of subsets of W (i.e., elements of 2^{2^W}) we use variables $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \ldots$. The elements of W are called 'states' or 'worlds'; the subsets of Agt are called 'coalitions'; the sets of states $X \in E(w)(C)$ are called the 'choices' of coalition C in state w. The set $E(w)(C)$ is called the 'choice set' of C in w. The complement of a set $\overline{X}$ or of a choice set $\overline{\mathcal{X}}$ are calculated from the obvious domains.

A dynamic Effectivity Function assigns, in each world, to every coalition a set of sets of states. Intuitively, if $X \in E(w)(C)$ the coalition is said to be able to *force* or *determine* that the next state after w will be some member of the set X. If the coalition has this power, it can thus prevent that any state *not* in X will be the next state, but it might not be able to determine *which* state in X will be the next state. Possibly, some other coalition will have the power to refine the choice of C.

For studying closed-world interaction we isolate a set of minimally required properties, that constitute the class of *coherent Effectivity Functions*.

Definition 2.2 [Coherence]
For any world w, coalitions C,D and choice X, an Effectivity Function is *coherent* if it has the following properties:

1. coalition monotonicity: if $X \in E(w)(C)$ and $C \subseteq D$ then $X \in E(w)(D)$;
2. regularity: if $X \in E(w)(C)$ then $\overline{X} \notin E(w)(\overline{C})$;
3. outcome monotonicity: if $X \in E(w)(C)$ and $X \subseteq Y$ then $Y \in E(w)(C)$;
4. inability of the empty coalition (IOEC): $E(w)(\emptyset) = \{W\}$.

◁

The first property says that the ability of a coalition is preserved by enlarging the coalition. In this sense we do not allow new members to interfere with the preexistent capacities of a group of agents. The second property says that if a coalition is able to force the outcome of an interaction to lie in a particular set, then no possible combinations of moves by the other agents can prevent this to happen. We think that regularity is a key property to understand the meaning of ability. If an agent is properly able to do something this means that others have no means to prevent it. Outcome monotonicity is a property of all Effectivity Functions in CL, which is therefore a monotonic modal logic. It says that if a coalition is able to force the outcome of the interaction to lie in a particular set, then is also able to force the outcome to lie in all his supersets (see [7]). The last condition is IOEC, that forces the empty coalition relation to be universal. As noticed also in [2] with such a property the empty coalition cannot force non-trivial outcomes of a game.

One important class of Effectivity Functions are the *playable* ones, to which we will refer throughout the paper.

Definition 2.3 [Playability]

For any world w an Effectivity Function is *playable* if it has the following properties:

(1) $\emptyset \notin E(w)(C)$, for any C; (2) $W \in E(w)(C)$ for any C. (3) E is Agt-maximal, that is for any $X \subseteq W$, s.t. $W \setminus X \notin E(w)(\emptyset)$ implies $X \in E(w)(Agt)$ (4) E is superadditive, i.e. for $C \cap D = \emptyset$, if $X \in E(w)(C)$ and $Y \in E(w)(D)$ then $X \cap Y \in E(w)(C \cup D)$.

◁

The first condition imposes that games are nonempty, the second that coalitions can always choose the largest possible set, the third that the grand coalition of agents can do whatever not blocked by Nature, the fourth that coalitions can join their forces.

As proved in [7] [Theorem 2.27], nonempty strategic games exactly correspond to playable Effectivity Functions [1].

2.0.1 Playability and Coherence

What kind of interactions are coherent Effectivity Functions isolating?

[1]The proof involves the definition of strategic game as a tuple $\langle N, \{\Sigma_i | i \in N\}, o, S\rangle$ where N is a set of players, each i being endowed with a set of strategies σ_i from Σ_i, an outcome function that returns the result of playing individual strategies at each of the states in S; the definition of α-Effectivity Function for a nonempty strategic game G, $E^{\alpha}_{G} : \wp(N) \rightarrow \wp\wp(S)$ defined as follows: $X \in E^{\alpha}_{G}$ iff $\exists \sigma_C \forall \sigma_{\overline{C}} o(\sigma_C; \sigma_{\overline{C}}) \in X$. The above mentioned theorem establishes that $E^{\alpha}_{G} = E$ in case E is playable and G is a nonempty strategic game.

In this respect, it is interesting to compare playable and coherent Effectivity Function, in order to understand the types of interactions we are considering.

Proposition 2.4 *Not all playable EF are coherent, and not all coherent EF are playable.*

Proof.

For the first part, take $W = \{x, y\}, Agt = \{i, j\}$ and the following Effectivity Function $E(\emptyset)(k) = E(\{i\})(k) = E(\{j\})(k) = E(Agt)(k) = \{W, W \backslash \{x\}\}$ for $k \in W$. Now it is just a matter of checking the conditions for playability.

For the second part take $W = \{x, y\}, Agt = \{i, j\}$ with $E(\emptyset)(k) = E(\{i\})(k) = E(\{j\})(k) = E(Agt)(k) = \{W\}$ for $k \in W$.

QED

Proposition 2.5 *Coherent Agt-maximal superadditive EF are playable.*

Proof.

It is a matter of checking the conditions of playability.

QED

3 On the axiomatization of Coherent Coalition Logic

In order to fully understand what sort of interactions we are investigating by using coherent effectivity functions we need to provide an axiomatization of their logic.

To do so we exploit some results due to Pauly and we adapt them to our framework. We recall first that Coalition Logic uses a modality $[C]\phi$ (to be read as "Coalition C can achieve ϕ") and it is interpreted in neighbourhood models with an outcome monotonic dynamic Effectivity Function as neighbourhood relation. The axioms of Coalition Logic extend propositional logic axiomatization with the Monotonicity axiom ($\phi \rightarrow \psi \Rightarrow [C]\phi \rightarrow [C]\psi$).

Consider the coalitional canonical model $C^* = ((W^*, E^*), V^*)$ and take $\overline{\phi} = \{w \in W^* | \phi \in w\}$, as the truth set of ϕ in the canonical model. The canonical relation (the rest is standard) is defined as

$$wE^*_C X \text{ iff } \exists \phi \text{ s.t } \overline{\phi} \subseteq X \text{ and } [C]\phi \in w$$

The set of formulas are closed under Modus Ponens and Monotonicity and the relation is easily proved to be monotonic. Moreover in [7] the following theorem [3.10] is proved: Every Coalition Logic Λ is sound and complete with respect to its canonical model C^*.

What we look for now is the a set of axioms and rules such that the corresponding maximally consistent sets generate a coherent Effectivity Function in the canonical models.

Nevertheless IOEC is not definable in Coalition Logic. To see this it is important to notice that Coalition Logic is monotonic multimodal logic, and frame validity of formulas of monotonic modal logics is closed under taking disjoint unions. This is proven for modal satisfaction in [4][Definition 4.1, Proposition 4.2].

Definition 3.1 [[4] 4.1]
Let $M_i = (W_i, N_i, V_i), i \in I$, be a collection of disjoint models. Then we define their *disjoint union* as the model $\oplus M_i = (W, N, V)$ where $W = \bigcup_{i \in I} W_i, V(p) = \bigcup_{i \in I} V_i(p)$ and for $X \subseteq W, w \in W_i$,

$$X \in N(w) \text{ iff } X \cap W_i \in N_i(w)$$

◁

Without loss of generality, we can simply think of the monotonic modal logic with only the box for the empty coalition, and take frames instead of models.

Consider the following monotonic frames $F_0 = (W_0, N_0)$ and $F_1 = (W_1, N_1)$, with a domain W_j and a relation $N_j \subseteq W_j \times 2^{W_j} (j \in \{0,1\})$. Take $W_0 = \{w_0\}$, $W_1 = \{w_1\}$, $N_0(w_0) = \{w_0\}$ and $N_1(w_1) = \{w_1\}$. Now suppose ϕ is some formula true at a world w in a model $M' = (W', N', V')$ of a monotonic frame F' iff $[[\top]]^M$ is neighbour of w (if $wN'[[\top]]$) and nothing else is ($wN'X \Rightarrow X = [[\top]]$). We see that $M_0, w_0 \models \phi$ and $M_1, w_1 \models \phi$ for arbitrary M_i inside F_i ($i \in \{0,1\}$). From [4] we construct the disjoint union $\oplus(F_0, F_1) = (W, N)$ as defined. We see clearly that our formula ϕ is not true in the disjoint union, because the neighbourhoods of the single models are copied in the disjoint union even if they are smaller than the unit. We observe moreover that the disjoint union is monotonic. The conclusion is that the formula expressing inability of the empty coalition is not definable is monotonic modal language.

At this point it is clear why $[\emptyset]\phi \rightarrow [\emptyset](\phi \vee \psi)$ or also $[\emptyset]\top$ would not be decent axioms for Coherent Coalition Logic. They would both ensure the presence of the unit in the neighbourhood of $\emptyset$, but they would not say anything about the absence of all the other sets. We will give to this intuition a formal characterization, stating that in fact the ability of the empty coalition in Coherent Coalition Logic *is* a global modality.

3.1 Inability of the Empty Coalition is a global relation

We extend the language of Coalition Logic with a global modality, defined as follows:

$$M, w \models E\phi \Leftrightarrow \exists w' \in W \text{ s.t. } M, w' \models \phi$$

The dual $A\phi$ is defined as $\neg E \neg \phi$. We claim that in Coalition Logic plus the global modality IOEC is definable.

Proposition 3.2 *$A(\phi) \leftrightarrow [\emptyset]\phi$ defines IOEC. That is,* $\models_C A(\phi) \leftrightarrow [\emptyset]\phi \Leftrightarrow E(w)(\emptyset) = \{W\}$ *for every w in the coalitional frames C.*

Proof. ($\Rightarrow$) Assume that $\models_C A\phi \leftrightarrow [\emptyset]\phi$ while not $E(w)(\emptyset) = \{W\}$ for every w in any frame F in the class of Coalitional Frames C. Then there is an F in which there is a w such that $E(w)(\emptyset) \neq \{W\}$. Notice that both W and $E(w)(\emptyset)$ are nonempty. So there is a $W' \neq W$ s.t $W' \in E(w)(\emptyset)$ and $W' \subset W$. Take an atom p to be true in all $w' \in W'$ and false in $W \setminus W'$. Now we have model M based on a coalitional frame C for which $M \not\models Ap \leftrightarrow [\emptyset]p$. Contradiction.

($\Leftarrow$) Assume $E(w)(\emptyset) = \{W\}$ for a given w in an arbitrary model M of a coalition frame in C, and that $w \models A\phi$. Then $[[\phi]]^M = W$ and $w \models [\emptyset]\phi$ follows. Assume now that $w \models [\emptyset]\phi$. It has to be the case that $[[\phi]]^M = W$ by assumption. So also that $w \models A\phi$, which concludes the proof.

QED

3.2 Axiomatization for the Global Modality plus a new inclusion axiom

The global relation induces an equivalence class in the models, therefore it is axiomatizable by an $S5$ modality interpreted on a global relation.

However this does not ensure that the underlying relation - that we indicate with $R_\exists$ - is globally connected. Global connectedness is not definable in basic modal language [1] [2].

As suggested in [1][p.417-418], taken a set of maximally consistent formulae Σ^+ we can simply take a generated submodel of the canonical model in such a way that the formulae in Σ^+ are invariant and the relation is (it follows by construction) a global relation.

Taken the canonical model $M^* = ((W^*, E^*, R^*_\exists), V^*)$, its submodel

$M^{*'} = ((W^{*'}, E^{*'}, R^{*'}_\exists), V^*)$ generated by Σ^+ using the $R^*_\exists$ relation should ensure that $R^{*'}_\exists = W^{*'} \times W^{*'}$.

Nevertheless in taking the generated submodel we should ensure that the coalitional relation is not alterated. One way to do it is to guarantee that the canonical coalitional relation is included in the global relation and that the generated submodel for the second relation is also a generated submodel for the first.

We begin with some definitions:

[2] The reason is also the invariance under taking disjoint unions. This fact sheds light on the relation between IOEC and Global Relation, in fact now we see clearly that the ability of the empty coalition in Coherent Coalition Logic *is* a global modality.

Definition 3.3 [Generated Submodels for Basic Modal Language, [1]]

Let $M = (W, R, V)$ and $M' = (W', R', V')$ be two models; we say that M' is a submodel of M if $W \subseteq W'$, R' is the restriction of R to W', that is $R' = R \cap (W' \times W')$ and V' is the restriction of V to M'. We say that M' is a *generated submodel* of M ($M' \mapsto M$)if M' is a submodel of M and for all points the following closure condition holds:

$$\text{if } w \text{ is in } M' \text{ and } Rwv, \text{ then } v \text{ is in } M'$$

◁

Modal satisfaction is invariant under taking generated submodels [1].

Now the definition for monotonic modal logic.

Definition 3.4 [Generated Submodels for Monotonic Modal Language, [4]] Given a monotonic model M, M' is a submodel of M if $W' \subseteq W, V'(p) = V(p) \cap W'$ for p atomic, and $N' = N \cap (W' \times 2^{W'})$, that is

$$\forall s \in W' : N'(s) = \{X \subseteq W' | X \in N(s)\}$$

In neighbourhood semantics given M' submodel of M, M' is also a generated submodel of M if the identity mapping $i : W \rightarrow W'$ is a bounded morphism, that is, for all $w' \in W'$ and all $X \subseteq W$

$$i^{-1}[X] = X \cap W' \in N'(w') \text{ iff } X \in N(w')$$

◁

For all states of the generated submodels, truth of modal formulas is preserved [4].

Now the question is, is the submodel generated a maximally consistent set of formulas Σ^+ using the existential global modality relation (making the canonical model strongly connected with respect to this relation) also a generated submodel with respect to the coalitional relation?

The answer is: it depends on the extra axioms. Usually when we have a K and a global modality it is sufficient to include the diamond relation in the global modality relation. But we cannot simply have:

$$[C]\phi \rightarrow E\phi$$

because the coalitional canonical relation may cross $S5$ equivalence classes. Instead the good candidate for our attempt is just the following:

$$A\phi \leftrightarrow [\emptyset]\phi$$

We claim that taking a generated submodel with respect to the global relation, given this axiom, ensures the condition of taking also a generated submodel with respect to the neighbourhood modality.

This is easy to see, because all the neighbourhoods of all coalitions are of the form $X \subseteq W$ and W is covered by the global modality.

Proposition 3.5 *The axiom $A\phi \leftrightarrow [\emptyset]\phi$ guarantees inclusion of the canonical relation in the global relation*

Proof.

Take a maximally consistent set of formulas Σ^+ that extends a consistent set of formulas Σ according to the axioms and the rules that we have just defined (for the global and the coalitional modality). Suppose now $A\phi$ is in Σ^+ for some ϕ. This means that $W^* = [[\phi]]^{C^*}$. Now take a given $[C]\psi$ in the same maximally consistent set of formulas. This means that $[[\phi]]^{C^*} \in E^*(\Sigma^+)(C)$. But by definition, $[[\phi]]^{C^*} \subseteq W^*$ which proves that all neihbourhoods are covered by the global modality relation.

QED

Now, let us take a generated submodel, as described in [1] for basic modal logic, using the maximally consistent set Σ^+ looking only at the global modality.

Proposition 3.6 *The generated canonical submodel under Σ^+ preserves both global modality and monotonic Coalition Logic formulas satisfaction.*

Proof.

It is just a matter of verifying that the generated submodel for the global relation is also a generated submodel for the coalitional relation.

QED

It follows that we have an axiomatization for the Coherent Coalition Logic.

3.3 A sound and complete axiomatization

Take now the maximally consistent sets $w \in W^*$, closed under the proof system depicted in the table.

We take the following conditions to describe coherence of the Effectivity Function on the canonical relation.

- $wE^*_C X$ iff $\exists \overline{\phi} \subseteq X : [C]\phi \in w$ and $\forall \overline{\psi} \subseteq (W^* \backslash X) : [\overline{C}]\psi \notin w$ (for $C \neq \emptyset$)
- $E^*_C \subseteq E^*_D$ (for $C \subseteq D$)
- $wE^*_C X$ iff $X = W^*$ (for $C = \emptyset$)
- $wR_\exists v$ iff $w, v \in W^*$

Proposition 3.7 *The canonical Coherent frame for Coalition Logic with $A\phi \leftrightarrow [\emptyset]\phi$ as axiom has the property that $E(w)(\emptyset) = \{W^*\}$ for any MCS w and $A\phi \leftrightarrow [\emptyset]\phi$ is valid in the class of frames with that property.*

It is a consequence of the previous propositions and the canonical relation definition.

Proposition 3.8 *The set of axioms and rules in the table are sound and complete with respect to Coherent Coalition Frames*

Proof.

We need just to check the statement with respect to $M^{*'}$. We omit the detailed proof.

QED

Proof System	
A1	$[C]\phi \rightarrow [D]\phi$ (for $C \subseteq D$)
A2	$[C]\phi \rightarrow \neg[\overline{C}]\neg\phi$
A3	$A\phi \leftrightarrow [\emptyset]\phi$
A4	$\phi \rightarrow E\phi$
A5	$EE\phi \rightarrow E\phi$
A6	$\phi \rightarrow AE\phi$
A7	$A(\phi \rightarrow \psi) \rightarrow (A\phi \rightarrow A\psi)$
R1	$\phi \wedge (\phi \rightarrow \psi) \Rightarrow \psi$
R2	$\phi \rightarrow \psi \Rightarrow [C]\phi \rightarrow [C]\psi$
R3	$\phi \Rightarrow A\phi$

3.4 On Agt-Maximal Coherent Games

Notice that if we add Agt-maximality to Coherent Games, the following holds:

$$M, w \models [Agt]\phi \leftrightarrow E\phi$$

This suggest, at the expressivity level, that Coherent Coalition Logic is powerful enough to reason on global properties of the models. These results are useful to apply the language to the study of multiagent interactions.

4 A Deontic Logic for Efficient Interactions

Any deontic language comes along with an idea of how a certain world state should be.

Once we view a deontic language as regulating a Multi Agent System, we can say that a set of commands promote a certain interaction (or social state), prohibiting certain others. Following this line of reasoning it is possible, given a notion of optimality or efficiency, to construct a deontic language that requires this notion to hold.

If we want to consider what it is socially optimal, as we do here, we can see obligations and prohibitions as resulting from one general norm saying that all actions of coalitions that do not take into account the interests of the society as a whole, are forbidden.

From the practical point of view, one way to view our logic is to say that it can be used to derive obligations, permission and prohibitions from conflicting group preferences, and use these as *suggestions* for norm introduction in the society.

This last part of the paper is devoted to formalize this derivation. Here we will introduce a notion of preference in the strategic interaction scenario, to be lifted to coalitional choice, in order to define what it is best for a society to choose. We will then move to study the property of the enriched language focusing on the regulation of coherent interactions. We will show that Nature can be obliged to do something when and only when it is not avoidable, that is it will be assigned only trivial obligations.

4.1 Preference

As already noticed by von Wright, the notion of preference can be understood and modeled in many ways [9]. This is especially true in strategic interaction, in which players, in order to choose what is best to do, need to have preferences over the possible outcomes of the game. Thus those are the preferences that constitute our main concern.

The claim is thus that players do have a fixed ordering over the domain of discourse (what we call *preferences*), and that generate their strategic preference considering where the game may end (called *choices domination*, or simply *domination*).

We start from a preference relation for individuals over states working our way up to preferences for coalitions over sets. A similar view is taken in [3].

Definition 4.1 [Individual preferences for states] A preference ordering $(\geq_i)_{i\in Agt}$ consists of a partial order (reflexive, transitive, antisymmentric) $\geq_i \subseteq W \times W$ for all agents $i \in Agt$, where $v \geq_i w$ means that v is 'at least as nice' as w for agent i. The corresponding strict order is defined as usual: $v >_i w$ if, and only if, $v \geq_i w$ and not $w \geq_i v$. ◁

Definition 4.2 [Individual preferences for sets of states] Given a preference ordering $(\geq_i)_{i\in Agt}$, we lift it to an ordering on nonempty sets of states by means of the following principles.

1. $\{v\} \geq_i \{w\}$ iff $v \geq_i w$; (Singletons)
2. $(X \cup Y) \geq_i Z$ iff $X \geq_i Z$ and $Y \geq_i Z$; (Left weakening)

3. $X \geq_i (Y \cup Z)$ iff $X \geq_i Y$ and $X \geq_i Z$. (Right weakening)

$\lhd$

These are some properties that seem minimally required for calling some relation a preference relation. The first ensures that preferences are copied to possible choices. The properties of left and right weakening ensure a lifting from singletons to sets.

The lifting enables us to deal with preference under uncertainty or indeterminacy. The idea is that if an agent were ever confronted with two choices X, Y he would choose X over Y provided $X >_i Y$. preferences do not consider any realizability condition, they are simply basic aspirations of individual players, on which to construct a more realistic order on the possible outcomes of the game, which are by definition dependent on what all the agents can do together.

Out of agents' preferences, we can redefine on choices the classical notion of Pareto Efficiency.

Definition 4.3 [Strong Pareto efficiency] Given a choice set $\mathcal{X}$, a choice $X \in \mathcal{X}$ is *Strongly Pareto efficient* for coalition C if, and only if, for no $Y \in \mathcal{X}$, $Y \geq_i X$ for all $i \in C$, and $Y >_i X$ for some. When $C = Agt$ we speak of *Strong Pareto Optimality*. $\lhd$

We will use the characterization of Pareto Efficiency and Optimality to refer to the notions we have just defined, even though the classical definitions (compare [5]) are weaker [3].

We now construct a preference relation on choices. To do so we first need to look at the interaction that agents' choices have with one another.

Definition 4.4 [Subchoice] If E is an Effectivity Function, and $X \in E(w)(\overline{C})$, then the X-*subchoice set* for C in w is given by $E^X(w)(C) = \{X \cap Y \mid Y \in E(w)(C)\}$. $\lhd$

Considering subchoices allows to reason on a restriction of the game and to consider possible moves looking from a coalitional point of view, i.e. what is best for a coalition to do provided the others have already moved.

When agents interact therefore they make choices on the grounds of their own preferences. Nevertheless the moves at their disposal need not be all those that the grand coalition has. We can reasonably assume that preferences are filtered through a given coalitional Effectivity Function. That is we are going to consider what agents prefer among the things they can do.

Definition 4.5 [Domination] Given an Effectivity Function E, X is *undominated* for C in w (abbr. $X \rhd_{C,w}$) if, and only if, (i) $X \in E(w)(C)$ and (ii) for all $Y \in E(w)(\overline{C})$, $(X \cap Y)$ is Pareto efficient in $E^Y(w)(C)$ for C. $\lhd$

The idea behind the notion of domination is that if X' and X'' are both members of $E(w)(C)$ then, in principle, C will not choose X'', if X' dominates X''. This property ensures that a preference takes into account the possible moves of the other players. This resembles the notion of Individual Rationality in Nash solutions [5], according to which an action is chosen reasoning on the possible moves of the others.

If we take the Coordination Game previously discussed, we have the following cases:

- $(White_R, White_C) \rhd_{Agt,w}$ for any w.
- $(Black_R, Black_C) \rhd_{Agt,w}$ for any w.
- not $(Black_C) \rhd_{C,w}$

The preceding three definitions capture the idea that 'inwardly' coalitions reason Pareto-like, and 'outwardly' coalitions reason strategically, in terms of strict domination. A coalition will choose its best option given all possible moves of the opponents. Looking at the definition of Optimality we gave, we can see that undomination collapses to individual rationality when we only consider individual agents, and to Pareto efficiency when we consider the grand coalition of agents.

Proposition 4.6

$X \rhd_{Agt,w}$ iff X is a standard Pareto Optimal Choice in w.

$X \rhd_{i,w}$ iff X is a standard Dominating Choice in w for i.

Proof For the first, notice that since $E(w)(\emptyset) = \{W\}$, then X is undominated for Agt in w iff it is Pareto efficient in $E(w)(Agt)$ for Agt (i.e., it is Pareto optimal in w). The second is due to the restriction of undomination to singleton agents.Q.E.D.

4.1.1 Violation

A way to impose normative constraints in a Multi Agent System is to look at the optimality of the strategic interaction of such system. In particular the presence of possible outcomes in which agents could not unanimously improve (Pareto Efficient) can be a useful guide line for designing a new set of norms to be imposed.

Following this line we define a a set of violation sets as the set of those choices that are not a Pareto Efficient interaction.

Definition 4.7 [Violation] If E is an Effectivity Function and $C \subseteq C'$, then the choice $X \in E(w)(C)$ is a C'-violation in w ($X \in VIOL_{C,C',w}$) iff there is a $Y \in$

[3] The last definition is clearer when we consider the case $\mathcal{X} = E(w)(C)$. But it is formulated in a more abstract way in order to smoothen the next two definitions.

$E(w)(C' \setminus C)$, $(X \cap Y)$ that is not undominated for C' in w. ◁

In words, X is a violation if it is not safe for the other agents, in the sense that not all the moves at their disposal yield an efficient outcome.

We indicate with $VIOL_{C,w}$ the set violations by C at w towards Agt.

5 Logic

We now introduce the syntax of our logic, an extension of the language of Coalition Logic [7] with modalities to talk about ideal states in a closed-world interaction.

5.1 Language

Let Agt be a finite set of agents and $Prop$ a countable set of atomic formulas. The syntax of our Logic is defined as follows:

$$\phi ::= p|\neg\phi|\phi \vee \phi|[C]\phi|E\phi|P(C,\phi)|F(C,\phi)|O(C,\phi)|[rational_C]\phi$$

where p ranges over $Prop$ and C ranges over the subsets of Agt. The other boolean connectives are defined as usual. The informal reading of the modalities is: "Coalition C can choose ϕ", "There is a state that satisfies ϕ","It is permitted (/forbidden/obligated) for coalition C to choose ϕ", "It is rational for coalition C to choose ϕ".

5.2 Structures

Definition 5.1 [Models] A *model* for our logic is a tuple

$$(W, E, R_\exists, \{\geq_i\}_{i\in Agt}, V)$$

where:

- W is a nonempty set of states;
- $E : W \longrightarrow (2^{Agt} \longrightarrow 2^{2^W})$ is a Coherent Effectivity Function.
- $R_\exists = W \times W$ is a global relation.
- $\geq_i \subseteq W \times W$ for each $i \in Agt$, is the preference relation. Out of this relation we define the undomination relation $\rhd \subseteq 2^{Agt} \times W \times 2^W \times 2^{2^W}$ as previously specified.
- $V : W \longrightarrow 2^{Prop}$ is a valuation function.

◁

5.3 Semantics

The satisfaction relation of modal formulas (the rest is standard) with respect to a pointed model M, w is defined as follows:

$$\begin{array}{rll}
M,w \models [C]\phi & \text{iff} & [[\phi]]^M \in E(w)(C) \\
M,w \models E\phi & \text{iff} & \exists v \text{ s.t. } M,v \models \phi \\
M,w \models [rational_C]\phi & \text{iff} & \forall X(X \rhd_{C,w} \Rightarrow X \subseteq [[\phi]]^M) \\
M,w \models P(C,\phi) & \text{iff} & \exists X \in E(w)(C) \text{ s.t.} \\
 & & X \in \overline{VIOL}_{C,w} \text{ and } X \subseteq [[\phi]]^M \\
M,w \models F(C,\phi) & \text{iff} & \forall X \in E(w)(C)(X \subseteq [[\phi]]^M \Rightarrow \\
 & & X \in VIOL_{C,w}) \\
M,w \models O(C,\phi) & \text{iff} & \forall X \in E(w)(C)(X \in \overline{VIOL}_{C,w} \Rightarrow \\
 & & X \subseteq [[\phi]]^M)
\end{array}$$

In this definition, $[[\phi]]^M =_{def} \{w \in W \mid M, w \models \phi\}$.

The modality for coalitional ability is standard from Coalition Logic [7]. The modality for rational action requires for a proposition ϕ to be rational (wrt a coalition C in a given state w) that all undominated choices (for C in w) be in the extension of ϕ. This means that there is no safe choice for a coalition that does not make sure that ϕ will hold. It is still possible for a coalition to pursue a rational choice that may be socially not rational. The deontic modalities are defined in terms of the coalitional abilities and preferences. A choice is permitted whenever is safe, forbidden when it may be unsafe (i.e. when it contains an inefficient choice), and obligated when it is the only safe.

6 Properties

It is now interesting to look at what we can say within our system.

Some Validities	
1	$P(C,\phi) \rightarrow \neg O(C, \neg\phi)$
2	$F(C,\phi) \leftrightarrow \neg P(C,\phi)$
3	$P(C,\phi) \vee P(C,\psi) \rightarrow P(C, \phi \vee \psi)$
4	$O(C,\phi) \rightarrow [C]\phi \rightarrow P(C,\phi)$
5	$[rational_C]\phi \wedge [rational_{Agt}]\neg\phi \rightarrow F(C,\phi)$
6	$O(C,\top)$
7	$O(\emptyset,\phi) \leftrightarrow [\emptyset]\phi$

The first validity says that permissions are consistent with obligations (the converse does not hold in general). The second that prohibition and permission are interdefinable. The third says that permission is monotonic. The fourth that the obligation to choose ϕ for an agent plus the ability to do something entails the permission to carry out ϕ. The validity number 5 says that the presence of a safe state that is rational for the grand coalition of agents is a norm for every coalition, even in case of conflicting preferences, i.e. in case of conflict the interest of the grand coalition prevails. The sixt one that there are no empty normative systems. The last validity says that obligations for Nature coincides with its ability. Notice that in Coherent Coalition Logic this means that obligation for Nature can only be a trivial choice.

6.1 Back to the Game

Norms of Conformity Consider the model M_c of the Game of Conformity described in Table 1.

Nature is obligated only a trivial choice:

$$M_c \models O(\emptyset, \phi) \leftrightarrow A\phi$$

What is interesting also is that also players are individually permitted only nontrivial choices:

$M_c \models \neg P(R, white_R) \wedge \neg P(R, black_R)$
$\wedge \neg P(C, black_C) \wedge \neg P(C, black_C)$.

But as coalition they are:

$$M_c \models P(\{R, C\}, white_{R,C}) \wedge M_c \models P(\{R, C\}, black_{R,C})$$

No precise indication of the choices is given by the resulting obligation:

$$M_c \models O(\{R, C\}, (white_R, C) \vee (black_{R,C}))$$

This is revealing of the form of the game: no equilibrium can be achieved by the agents acting independently, but only as a coalition [4]. As a matter of fact, looking at the obligations for this game tells us more than just a static fact about coalitional choice. In Coordination games only the grand coalition can make an optimal choice, which suggest that the grand coalition is in fact obligated to form.

7 Conclusion and Future Work

In this paper we studied those interactions in which Nature does not play an active role, and we proposed a deontic logic to indicate their optimal solutions. We provided an axiomatization of the resulting logic, switching from gameform interactions to interactions with preferences in order to analyze gametheoretical examples like Coordination Game. The work here described allows for several developments. Among the most interesting ones is the study of the relation between imposed outcomes and steady states that describe where the game will actually end up (i.e. Nash Solution, the Core etc...). As suggested by the last example, some obligations say something about the convenient dynamics to achieve a socially optimal outcome. One idea is to talk explicitly about such dynamics. Conversely another feature that is worth studying is those structures in which Pareto Efficiency is not always present. Agents will reckon some actions as optimal even though there is no social equilibrium that can be ever reached. This can be achieved by talking explicitly about preferences in the language as done for instance in [8]. The study of the interaction between choices and preferences has shown to have an interesting connection with deontic logic that, viewed in a multiagent perspective, allows to talk about those desirable properties that an interaction should have. As system designers, our aim is at last to construct efficient social procedures that can guarantee a socially desirable property to be reached. We think that normative system design is at last a proper part of the Social Software enterprise [6].

[4]Notice that we have no way of detaching from this choice a more precise command: $O(C, \phi \vee \psi) \rightarrow ((O(C, \phi) \vee O(C, \psi)))$ is not a validity.

References

[1] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge Tracts in Theoretical Computer Science, 2001.

[2] S. Borgo. Coalitions in action logic. In *Proc. of IJCAI*, pages 1822–1827, 2007.

[3] P. Gardenfors. Rights, games and social choice. *Nous*, 15:341–56, 1981.

[4] H.H. Hansen. *Monotonic Modal Logics*. Master Thesis, ILLC, 2001.

[5] M. Osborne and A. Rubinstein. *A course in Game Theory*. The MIT Press, 1994.

[6] R. Parikh. Social software. *Synthese*, 132(3):187–211, 2002.

[7] M. Pauly. *Logic for Social Software*. ILLC Dissertation Series, 2001.

[8] J. van Benthem and F.Liu. Dynamic logic of preference upgrade. *Journal of Applied Non-Classical Logics*, 14, 2004.

[9] G.H. von Wright. *The logic of preference*. Edimburgh University Press, 1963.