Modal Logics of Negotiation and Preference

Ulle Endriss and Eric Pacuit

Institute for Logic, Language and Computation
University of Amsterdam

Abstract. We develop a dynamic modal logic that can be used to model scenarios where agents negotiate over the allocation of a finite number of indivisible resources. The logic includes operators to speak about both preferences of individual agents and deals regarding the reallocation of certain resources. We reconstruct a known result regarding the convergence of sequences of mutually beneficial deals to a Pareto optimal allocation of resources, and discuss the relationship between reasoning tasks in our logic and problems in negotiation. For instance, checking whether a given restricted class of deals is sufficient to guarantee convergence to a Pareto optimal allocation for a specific negotiation scenario amounts to a model checking problem; and the problem of identifying conditions on preference relations that would guarantee convergence for a restricted class of deals under all circumstances can be cast as a question in modal logic correspondence theory.

1 Introduction

Negotiation between autonomous agents over the allocation of resources has become a central topic in AI. In this paper, we present some first steps towards using (modal) logic to model negotiation scenarios. We explore to what extent known results about negotiation can be reconstructed in such a logic and whether it is possible to derive new insights about a negotiation framework by studying its formalisation in logic. The particular negotiation framework we are interested in here, which has recently been studied by several authors [1, 2, 3], involves a number of autonomous agents negotiating over the reallocation of a number of indivisible goods amongst themselves. Agents have preferences over the resources they hold, and they will only agree to take part in a deal if that deal would leave them with a preferred bundle of goods. That is, negotiation is driven by the rational interests of the participating agents. At the same time, we can observe different phenomena at the global level. For instance, it may or may not be the case that the sequence of deals implemented by the agents converges to a socially optimal allocation of resources (say, a Pareto optimal allocation).

Our aim in this paper is to show how such a negotiation setting can be formalised using modal logic. More specifically, we are developing a logic in the style of propositional dynamic logic (PDL) that allows us to speak both about the preferences of individual agents and the aggregated preferences of the society as a whole (to model Pareto improvements), as well as deals between agents involving the reassignment of specific resources to other agents. We show that

M. Fisher et al. (Eds.): JELIA 2006, LNAI 4160, pp. 138–150, 2006.

properties such as guaranteed convergence to a Pareto optimal allocation can be expressed in this logic, and we discuss how to apply logical reasoning techniques, such as model checking, to decision problems arising in the context of negotiation.

This work also fits in with the larger project of "social software" first discussed by Parikh [4]. The main idea of social software is that tools and techniques from computer science (in particular logic of programs) can be used to reason about *social procedures* (see [5] for a survey of the relevant literature). Much of the work on social software is concerned with developing logics intended to verify the "correctness" of social procedures [6]. There are often two key features of these logics. First, they should be expressive enough to capture the relevant concepts in order to state correctness conditions. Second, the logics should have well-behaved computational properties (for example, a decidable satisfiability problem and polynomial time model checking). The present paper will pay close attention to both of these issues.

The paper is organised as follows. In Section 2 we introduce a PDL-style logic for reasoning about negotiation settings, prove its decidability, and discuss some illustrative examples. Then we show in Section 3 how the language of this logic can express a property known as guaranteed convergence to a Pareto optimal allocation. Our discussion shows that this can be reduced to a statement about Pareto improvements alone; and we consequently introduce a second, more basic logic to reason about Pareto efficiency in Section 4. Section 5 concludes with an extensive discussion of further possibilities of linking reasoning tasks in our logic of negotiation spaces and questions arising in the context of negotiation. The appendix summarises relevant results about PDL and its extensions.

2 The Logic of Negotiation Spaces

In this section, we are going to develop a logic to describe negotiation scenarios of the following sort. There are a (finite) number of *agents* and a (finite) number of *resources*, which are indivisible and cannot be shared amongst more than one agent at a time. An *allocation* is a partitioning of the resources amongst the agents (each resource has to be assigned to exactly one agent). Agents have *preferences* over the bundles of resources they receive (but they are indifferent to what resources are being received by other agents; that is, we do not want to model allocative externalities). To improve their situation, agents can agree on *deals* to exchange some of the resources currently in their possession. In the most general case, we allow for any kind of multilateral deal. That is, a single deal may involve the reassignment of any number of resources amongst any number of agents. Agents are assumed to be *rational* in the sense of never accepting a deal that would leave them with a bundle that they like less than the bundle they did hold prior to that deal.

As outside observers, we are not actually interested in the preferences of individual agents, but we do care about the quality of allocations from a social point of view. In particular, we are going to be interested in allocations of resources that are *Pareto optimal* as well as in sequences of deals that lead to such Pareto optimal allocations. To describe such scenarios, we develop the logic $\mathcal{L}_{\langle \mathcal{A},\mathcal{R}\rangle}$, which is parametrised by a finite set of *agents* $\mathcal{A}$ and a finite set of *resources* $\mathcal{R}$.

2.1 Preliminaries

An *allocation* is a total function $A : \mathcal{R} \to \mathcal{A}$ specifying for each resource item which agent currently holds that item. As we shall see, the set $\mathcal{A}^{\mathcal{R}}$ of all allocations will be the "set of worlds" in the (intended) models of our logic. An *atomic deal* is of the form $(a \leftarrow r)$, for $a \in \mathcal{A}$ and $r \in \mathcal{R}$. It specifies that resource r is being reassigned to agent a (which agent held r before the deal is left unspecified). Each of these atomic deals induces a binary relation $R_{a\leftarrow r}$ over the set of allocations $\mathcal{A}^{\mathcal{R}}$: given two allocations x and y, we have $xR_{a\leftarrow r}y$ iff x and y are identical except possibly for the assignment of resource r which must be assigned to agent a in allocation y.

Each agent $i \in \mathcal{A}$ is equipped with a preference relation R_i over alternative *bundles* of resources: $R_i \subseteq 2^{\mathcal{R}} \times 2^{\mathcal{R}}$. We require preference relations to be *reflexive* and *transitive* (but not necessarily monotonic, for instance). Each R_i extends to a preference relation over alternative *allocations* of resources: for allocations $A, A' \in \mathcal{A}^{\mathcal{R}}$, we have $(A, A') \in R_i$ iff $(\{r \in \mathcal{R} \mid A(r) = i\}, \{r \in \mathcal{R} \mid A'(r) = i\}) \in R_i$. That is, agent i prefers allocation A' over allocation A iff they prefer the bundle they receive in A' over the bundle they receive in A. While the R_i are defined in terms of bundles, we are mostly going to use them in this derived form, as relations over allocations. Union ($\cup$), intersection ($\cap$), complement ($\overline{R}$), converse (R^{-1}), and iteration (R^*) of relations are defined in the usual manner.

2.2 Syntax

Atomic propositions. Let At be a finite or countable set of atomic propositions, including the special symbols H_{ij} for all $i \in \mathcal{A}$ and all $j \in \mathcal{R}$. The intended meaning of H_{ij} is that agent i holds resource j.

Relations and formulas. We first define the range of terms that can be used to index a modal operator, and then the set of formulas itself. We assume there is a set of atomic relation terms, one for each atomic deal relation and one for each preference relation. We will use the same symbol to represent both a relation term and the relation. We trust this abuse of notation will not cause any confusion. A relation term has the following syntactic form:

$$R \ ::= \ r \mid R \cup R' \mid R \cap R' \mid R^{-1} \mid \overline{R} \mid R^*,$$

where r is an atomic relation of the form $R_{a\leftarrow r}$ or R_i. Formulas have the following syntactic form:

$$\varphi \ ::= \ p \mid \neg\varphi \mid \varphi \vee \psi \mid \langle R\rangle\varphi,$$

where $p \in \mathsf{At}$ and R is a relation term. Further logical operators, such as conjunction, can be defined in terms of the above in the usual manner. The box-operator, in particular, is defined as the dual of the diamond: $[R]\varphi = \neg\langle R\rangle\neg\varphi$.

2.3 Semantics

Frames. A *frame* $\mathcal{F} = (\mathcal{A}, \mathcal{R}, \{R_i\}_{i \in \mathcal{A}})$ is a triple consisting of a set of agents $\mathcal{A}$, a set of resources $\mathcal{R}$, and a set of preference relations R_i over allocations, one for each agent. This would corresponds to the frame $(\mathcal{A}^{\mathcal{R}}, \{R_i\}_{i \in \mathcal{A}})$ in the standard Kripke semantics for a multi-modal logic; that is, the "worlds" in a frame are allocations of resources. Note that the deal relations $R_{a \leftarrow r}$ are fully specified by $\mathcal{A}$ and $\mathcal{R}$ already, so these need not be specified as relations of the frame.

Models. A model $\mathcal{M} = (\mathcal{F}, V)$ is a pair consisting of a frame $\mathcal{F} = (\mathcal{A}, \mathcal{R}, \{R_i\}_{i \in \mathcal{A}})$ and a valuation function V mapping atomic propositions to subsets of $\mathcal{A}^{\mathcal{R}}$. Intuitively, $V(p)$ will be the set of allocations at which the proposition p is true. V has to respect the condition $V(H_{ij}) = \{A \in \mathcal{A}^{\mathcal{R}} \mid A(j) = i\}$. That is, H_{ij} is true in exactly those allocations where agent i holds resource j.

Truth in a model. Truth of a formula φ at a world w (an allocation) in a given model $\mathcal{M}$ is defined as follows:

(1) $\mathcal{M}, w \models p$ iff $w \in V(p)$ for atomic propositions p;
(2) $\mathcal{M}, w \models \neg\varphi$ iff not $\mathcal{M}, w \models \varphi$;
(3) $\mathcal{M}, w \models \varphi \vee \psi$ iff $\mathcal{M}, w \models \varphi$ or $\mathcal{M}, w \models \psi$;
(4) $\mathcal{M}, w \models \langle R \rangle \varphi$ iff there is a $v \in \mathcal{A}^{\mathcal{R}}$ such that wRv and $\mathcal{M}, v \models \varphi$.

For instance, $\langle R_i \rangle \varphi$ means that φ is true in some allocation that agent i prefers over the current allocation. Notions such as validity and satisfiability are defined in the usual manner [7, 8]. The formula $[R_{a \leftarrow r_1} \cup R_{a \leftarrow r_2}]\varphi$, for instance, expresses that in every allocation that we can reach by giving either item r_1 or item r_2 to agent a satisfies φ.

2.4 Decidability

Next we are going to show that the logic $\mathcal{L}_{\langle \mathcal{A}, \mathcal{R} \rangle}$ is decidable. This may seem surprising at first, given the close connection of our logic to PDL extended with the complement operator, which is known to be undecidable (see appendix). In short, the reason why $\mathcal{L}_{\langle \mathcal{A}, \mathcal{R} \rangle}$ is decidable is that, for this logic, fixing the language of formulas involves fixing the set $\mathcal{A}$ of agents and the set $\mathcal{R}$ of resources. This is turn amounts to fixing the set of possible worlds of our models.

Proposition 1 (Decidability). *The logic $\mathcal{L}_{\langle \mathcal{A}, \mathcal{R} \rangle}$ is decidable.*

Proof. A formula φ in the language of $\mathcal{L}_{\langle \mathcal{A}, \mathcal{R} \rangle}$ is valid iff it is true at every world in every model of $\mathcal{L}_{\langle \mathcal{A}, \mathcal{R} \rangle}$. The number of frames of $\mathcal{L}_{\langle \mathcal{A}, \mathcal{R} \rangle}$ is finite: $\mathcal{A}$ and $\mathcal{R}$ are fixed and the number of choices for each preference relation R_i is bound above by the square of the number of bundles of resources from $\mathcal{R}$. The definition of the valuation function over atomic propositions not appearing in φ is not relevant, so we only need to consider a finite number of valuation functions, and hence a finite number of models. Each of these models is itself finite, and checking whether φ is true at a given world in a given model is a decidable problem. Hence, checking validity amounts to deciding a finite number of decidable problems, so it must be a decidable problem itself. $\square$

2.5 Examples

We are now going to give a couple of examples that demonstrate what can be expressed in our logic $\mathcal{L}_{\langle \mathcal{A},\mathcal{R}\rangle}$ of negotiation spaces.

Describing bundles and allocations. Formulas of the following form completely specify the bundle held by agent i (there is one such formula for each $X \subseteq \mathcal{R}$):

$$\textsc{bun}_i^X = \bigwedge_{j \in X} H_{ij} \ \wedge \bigwedge_{j \in \mathcal{R}\setminus X} \neg H_{ij} \tag{1}$$

Conjunctions of such BUN-formulas (with one conjunct for each $i \in \mathcal{A}$) completely specify an allocation. Let $\langle X_1, \ldots, X_n\rangle$ be a partitioning of the set of resources $\mathcal{R}$. The following formula identifies the corresponding allocation:

$$\textsc{alloc}_{\langle X_1,\ldots,X_n\rangle} = \bigwedge_{i=1}^{n} \textsc{bun}_i^{X_i} \tag{2}$$

Given our semantics, any such ALLOC-formula will be true in exactly one world (by definition); that is, these formulas have a similar role as *nominals*, familiar from hybrid logic [7]. In fact, an alternative approach would have been to introduce a nominal for each allocation, and to define the propositions H_{ij} in terms of these nominals, rather than giving the H_{ij} a special status.

No externalities. In our definition of the preference relations R_i we have stipulated that they should be free of externalities by defining them as being induced by preferences over bundles. Next we are going to see that this could in fact also be defined syntactically; that is, we may define the R_i as preference relations over allocations and additionally impose axioms that exclude the option of externalities, when this is desired. We first define a modality that allows us to move to any world in the model from any given starting point. This is possible, because all worlds (allocations) can be reached by a sequence of atomic deals (as long as no conditions on the acceptability of a deal are being imposed).

$$[*]\varphi = [(\bigcup_{a\in\mathcal{A}, r\in\mathcal{R}} R_{a\leftarrow r})^*]\varphi \tag{3}$$

Since any two states of our model are connected via a finite sequence of deals, $[*]$ is a universal modality. That is, $[*]\varphi$ is true at a state provided φ is true at *every* state in the model.

Intuitively, the preferences depend only on the bundles if, whenever there is a situation in which agent i prefers bundle Y over bundle X, then whenever the agent has bundle X, then the agent prefers a situation in which it has bundle Y. With the help of the universal modality we can express this as follows:

$$(\textsc{bun}_i^X \wedge \langle R_i\rangle \textsc{bun}_i^Y) \rightarrow [*](\textsc{bun}_i^X \rightarrow \langle R_i\rangle \textsc{bun}_i^Y) \tag{4}$$

The conjunction of the above type of implication for all bundles $X, Y \in 2^{\mathcal{R}}$ would then describe the fact that preferences only depend on bundles (no externalities).

3 Convergence to a Pareto Optimal Allocation

A central question in negotiation concerns *convergence* [1, 2, 3]: under what circumstances can we be sure that any sequence of deals negotiated by the agents will eventually lead to an allocation with certain desirable properties? Such "desirable properties" are usually expressed in terms of an aggregation of the preferences of the individual agents. A fundamental criterion for economic efficiency is the concept of Pareto optimality: an allocation of resources is *Pareto optimal* iff there is no other alternative that would be strictly better for one agent without being worse for any of the others [9]. In this paper, we are going to be interested under what circumstances a sequence of deals can be guaranteed to converge to a Pareto optimal allocation of resources. More specifically, in this section, we are going to reconstruct a result of [2], which may be paraphrased as stating that any sequence of deals that are beneficial for all the agents involved and that are not subject to any structural restrictions (say, on the number of agents involved in a single deal), will eventually result in a Pareto optimal allocation.

We are now going to formalise this result as a formula of $\mathcal{L}_{\langle \mathcal{A},\mathcal{R}\rangle}$. This formula will have the following general structure: $[\Phi^*]\langle\Phi^*\rangle\textsc{opt}$. Here Φ stands for the union of all deals that are possible and $\textsc{opt}$ is a formula describing that the allocation in question is "optimal". So the formula says that for any initial allocation, if we implement any sequence of Φ-deals, we can always reach an optimal allocation by implementing a further such sequence (or we are already at the optimal allocation).

To instantiate this template to a concrete formula, we first need to say what it means for a deal (a move to another allocation) to be "beneficial" (or *rational*) for everyone involved. For this we use the notion of Pareto improvement. We first need to define an agent's strict preference. Given any preference R_i, we can define its strict version, R_i^s as follows. For allocations w and v, say that $wR_i^s v$ if $wR_i v$ and it is not the case that $vR_i w$. Thus,

$$R_i^s = R_i \cap \overline{R_i^{-1}} \tag{5}$$

Thus the intended interpretation of $\langle R_i^s\rangle\varphi$ is that φ is true at an alternative which agent i *strictly* prefers to the current state.

We can now define a relation, denoted $\textsc{par}$, with intended interpretation of $\langle\textsc{par}\rangle\varphi$ being that φ is true at an alternative which is a Pareto improvement to the current alternative. Formally, we define $\textsc{par}$ as follows:

$$\textsc{par} = \bigcap_{i\in\mathcal{A}} R_i \cap \bigcup_{i\in\mathcal{A}} R_i^s \tag{6}$$

Now if $\mathcal{M}, w \models [\textsc{par}]\bot$, then w is an "end-state" with respect to the $\textsc{par}$ relation. Thus, there is no state which is a Pareto improvement over w. In other words, w is *Pareto efficient*.

Requiring deals to be rational is one way of restricting the range of possible deals. Another form of restriction are *structural* constraints. For instance, a particular negotiation protocol may only permit agents to negotiate bilateral deals (deals involving only two agents each), or there may be an upper limit on the number of resources that can be reassigned in a single deal. Let D be the set of deals licensed by our negotiation protocol. For instance, D could be the set of all *atomic* deals:

$$D = \bigcup_{a \in \mathcal{A}, r \in \mathcal{R}} R_{a \leftarrow r} \tag{7}$$

Another option would be to define D as the set of *all* deals (observe that every deal can be implemented as a sequence of atomic deals):

$$D = (\bigcup_{a \in \mathcal{A}, r \in \mathcal{R}} R_{a \leftarrow r})^* \tag{8}$$

We should note that, of course, not every restriction of interest can be expressed using our language for describing deals. This is due to the fact that we define atomic deals in terms of a single resource and the agent receiving that resource, but we do not specify from which other agent that resource is being taken.

The set of deals that are both rational and subject to the structural constraints defining D are given by the intersection $D \cap \textsc{par}$. Sequences of such deals belong to $(D \cap \textsc{par})^*$. We can now state the convergence property:

$$[(D \cap \textsc{par})^*]\langle (D \cap \textsc{par})^* \rangle [\textsc{par}]\bot \tag{9}$$

This formula expresses that any sequence of deals that are rational and belong to D will either lead to a Pareto optimal allocation, or to an allocation from which a Pareto optimal allocation is still reachable by means of such a sequence. In case we also know that any such sequence is bound to *terminate*, then this reduces to every sequence of rational D-deals eventually resulting in a Pareto optimal allocation of resources. For D being the full set of deals (without any structural restrictions), this has been proved to hold in [2]. Hence, formula (9) with D being the full set of deals must be valid in our logic $\mathcal{L}_{\langle \mathcal{A}, \mathcal{R} \rangle}$.

We can see this also as follows. If D is the full set of deals, *i.e.* D is defined by equation (8), then D is a universal relation, linking any two allocations in $\mathcal{A}^{\mathcal{R}}$. Hence the intersection $D \cap \textsc{par}$ is actually just the relation $\textsc{par}$. It is not difficult to see (and we are going to explain precisely why in the following section), that $\textsc{par}$ must be a transitive relation. Hence, $\textsc{par}^*$ is just the reflexive closure of $\textsc{par}$. Thus formula (9) reduces to the formula $[\textsc{par}^*]\langle \textsc{par}^* \rangle [\textsc{par}]\bot$. Observe that this formula is valid on a given frame iff the following is:

$$[\textsc{par}]\bot \vee \langle \textsc{par} \rangle [\textsc{par}]\bot \tag{10}$$

That is, either we are already at a Pareto efficient state or there is a $\textsc{par}$-path that leads to a Pareto efficient state. Thus our convergence theorem reduces to a statement purely about Pareto improvements, which can be expressed in a fragment of our logic in which the modalities contain only preference relation symbols. Since this logic may be of independent interest, we treat it in detail in the next section.

4 The Logic of Pareto Efficiency

The goal of this section is to develop a logic of Pareto efficiency. We start with an arbitrary set of alternatives W and assume each agent has a (reflexive and transitive) preference over W. This is the setting of a recent paper by van Benthem *et al.* [10]. In fact, studying preferences from a logical perspective has been studied by a number of different authors (cf. Hansson [11]). Of course, since each R_i is assumed to be reflexive and transitive, the class of all preference models is axiomatized by multi-agent **S4**. Van Benthem *et al.* [10] show that taking the above language as a starting point, a number of different game-theoretic notions, such as the Nash equilibrium and the backward induction solution, can be expressed and studied from a modal preference logic point of view. To that end, standard tools from extended modal logic, such as nominals, dynamic epistemic operators, and the universal modality, are used. The logic presented in this section continues this line of thinking.

Let At be a finite or countable set of atomic propositions. The language of the logic $\mathcal{L}_{\mathsf{Pareto}}$ of Pareto efficiency is defined as follows (with $p \in \mathsf{At}$):

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \langle R_i\rangle\varphi \mid \langle R_i^s\rangle\varphi \mid \langle \text{PAR}\rangle\varphi$$

The standard boolean connectives and the operators $[R_i]$, $[R_i^s]$ and $[\text{PAR}]$ are defined as usual. Truth in a model is defined as usual. Here we are working in a multi-modal language interpreted over standard Kripke structures in which the accessibility relation for each $\langle R_i^s\rangle$ and the $\langle\text{PAR}\rangle$ modal operator are defined in terms of the R_i relations. This is analogous to working in a multi-agent epistemic logic with a common knowledge operator (in this case, the accessibility for the common knowledge operator is defined to be the reflexive transitive closure of the union of the individual accessibility relations). Recall the definitions of R_i^s and PAR from the previous section. Putting everything together, a *preference model* is a tuple $(W, \{R_i\}_{i\in\mathcal{A}}, V)$ where each R_i is reflexive and transitive, and the R_i^s and PAR relations are defined as above.

For issues of decidability and axiomatization it will be convenient to interpret the above language as a fragment of PDL with converse, intersection and complement operators. In this case, each R_i is an atomic program, and the modalities $\langle R_i^s\rangle$ and $\langle\text{PAR}\rangle$ can be defined by the appropriate operations on the R_i. See the appendix for a discussion of the relevant issues. We end this section with two simple observations.

Observation 1. *If each R_i is transitive, then* PAR *is transitive.*

Proof. Suppose that $w\text{PAR}v$ and $v\text{PAR}z$. By transitivity of the R_i, it is easy to see that $(w,z) \in \bigcap_i R_i$. Since $v\text{PAR}z$, there is some agent i such that vR_iz but not zR_iv. Our claim is that not zR_iw. Suppose that zR_iw. Then by transitivity of R_i, since zR_iw and wR_iv, zR_iv which contradicts our assumption. □

Consider again formula (10): $[\text{PAR}]\bot \vee \langle\text{PAR}\rangle[\text{PAR}]\bot$. Intuitively, this formula will be true at an alternative w provided either w is Pareto efficient or there is a Pareto improvement v that is. That is, $\mathcal{M}, w \models [\textsc{par}]\bot \vee \langle\textsc{par}\rangle[\textsc{par}]\bot$ just in case either there is no v such that $w\textsc{par}v$ or $w\textsc{par}v$ and v is an "end state". Our last observation is that assuming W is finite, this formula is valid.

Observation 2. *Suppose that W is finite and $\mathcal{M} = (W, \{R_i\}_{i\in\mathcal{A}}, V)$ is a preference model. Then for each $w \in W$, we have $\mathcal{M}, w \models [\textsc{par}]\bot \vee \langle\textsc{par}\rangle[\textsc{par}]\bot$.*

Proof. The proof follows easily from the fact that $\textsc{par}$ is irreflexive and W is assumed to be finite. Under these assumptions it is easy to see that for each state $w \in W$, if w is not an $\textsc{par}$ end state, then it is $\textsc{par}$ accessible to an $\textsc{par}$ end state. That is, for each $w \in W$, either there is no state v such that $w\textsc{par}v$ or there is a state $v \in W$ such that $w\textsc{par}v$ and for each $v' \in W$, it is not the case that $v\textsc{par}v'$. This is precisely what it means to say that $\mathcal{M}, w \models [\textsc{par}]\bot \vee \langle\textsc{par}\rangle[\textsc{par}]\bot$. □

From a modal logic perspective, these observations are easy exercises. However, from the perspective of this paper, they demonstrate that modal logic, and in particular variants of PDL, can provide an interesting perspective on negotiation.

5 Discussion

In this section we are going to explore further connections between different types of reasoning tasks in our logic $\mathcal{L}_{\langle\mathcal{A},\mathcal{R}\rangle}$ and questions arising in the context of negotiation.

5.1 Necessity of Complex Deals and Satisfiability

Besides convergence, another important property of negotiation systems that has been studied in the literature concerns the *necessity* of specific deals [1, 2]. A given deal or class of deals, characterised by structural constraints (rather than rationality conditions), is said to be necessary in view of reaching an allocation with a certain desired property (such as being Pareto optimal) by means of rational deals iff there are an initial allocation and individual preference relations such that any path leading to such a desirable allocation would have to involve that particular deal. A known result [2] states that if you do not allow all structural types of deals, but do require rationality, then you cannot guarantee Pareto optimal outcomes in all cases. In this section, we are going to discuss what this result corresponds to in our logic $\mathcal{L}_{\langle\mathcal{A},\mathcal{R}\rangle}$.

Consider again our convergence formula (9). The claim is that, if the set of deals D excludes even a single deal, then formula (9) will cease to be valid. In other words, its negation will become satisfiable:

$$\neg[(D \cap \textsc{par})^*]\langle(D \cap \textsc{par})^*\rangle[\textsc{par}]\bot \tag{11}$$

The proof of the necessity theorem given in [2] amounts to giving a general algorithm for constructing individual preference relations and an initial allocation such that the one deal not included in D will be the only deal taking us from the initial allocation to the (only) allocation that Pareto-dominates the initial allocation. This constructive element of the proof would correspond to giving a general method for proving satisfiability of formula (11). Vice versa, the known necessity theorem shows that formula (11) must be satisfiable for any given set of deals D that is not the full set of complex deals.

The discussion of necessity theorems highlights the fact that the exact form of presentation chosen for specifying deals can lead to somewhat different results. In [2] deals are represented as pairs of allocations, which amounts to a more fine-grained representation than we have opted for in this paper. For example, the deal $R_{a \leftarrow r}$ does in fact represent n different deals: for any of the n agents (including a itself), that agent could have owned r before the deal. If the more fine-grained representation is chosen, then certain deals need to be excluded from the statement of the theorem: a deal that is *independently decomposable* (meaning there are two groups of agents involved in the deal, but not a single resource is changing group) is not necessary for convergence, but can always be decomposed into two smaller deals. If deals are specified in terms of reassignments, as in this paper, however, each such deal does in fact correspond to a class of deals involving both independently decomposable deals and deals that are not independently decomposable. Hence, excluding that whole class from the negotiation protocol will always cause a problem, and therefore any such deal must be necessary.

5.2 Reachability Properties and Model Checking

Recall the formulation of the convergence property as given by formula (9). It states that any sequence of rational D-deals will eventually result in a Pareto optimal allocation (or in an allocation from which a Pareto optimal allocation is still accessible by means of such a sequence). We have seen that the formula is valid if D is the full set of deals, and that it is not valid if D is any subset of the full set of deals (that is, every single deal is necessary).

Dunne and colleagues [3, 12] have studied the complexity of deciding whether a given negotiation scenario allows for convergence to an optimal allocation by means of a structurally restricted class of (rational) deals. To be precise, these authors have concentrated on a framework where agent preferences are represented using utility functions (rather than ordinal preference relations) and where an allocation is considered optimal if it maximises the sum of individual utilities (so-called utilitarian social welfare [9]), a notion that is stronger than Pareto optimality. Nevertheless, conceptually there are interesting parallels to be explored.

This problem of deciding whether a given negotiation scenario admits convergence for a given restricted class of deals amounts to a *model checking* problem in our logic. This is interesting for at least two reasons. Firstly, model checking as a well-developed algorithmic technique may turn out to be a useful tool for deciding such questions in practice. Secondly, it may be of interest to compare and relate complexity results for negotiation frameworks and PDL model checking. A discussion of the latter may be found in the appendix. As shown by

Lange [13], model checking is PTIME-complete for all conceivable extensions of PDL (e.g. with intersection). It is important to note, however, that such complexity results must be understood with respect to the number of worlds in a model. In our case (as in many other applications), this will be an exponential number. Dunne and Chevaleyre [12] have recently shown that deciding whether a given negotiation scenario admits convergence by means of rational atomic deals is PSPACE-complete for the "numerical" version of the problem (with utility functions). A deeper understanding of the exact relationship between the two problems may allow us to obtain complexity results for model checking in our logic expressed in terms of the numbers of agents and resources (rather than the exponential number of allocations).

5.3 Guaranteed Convergence and Correspondence Theory

While Dunne *et al.* [3] have concentrated on establishing complexity results for deciding when convergence is possible, another line of work has attempted to establish general conditions (on the preferences of individual agents) that would guarantee that convergence by means of structurally simple deals is always possible [2, 14]. These results mostly relate to the numerical negotiation framework (with utility functions, monetary side payments, and maximal utilitarian social welfare as the chosen notion of optimality). Also, these results are either very simple (for instance, if all agents use modular utility functions, then convergence to an optimal allocation can be guaranteed by rational atomic deals alone) or require an overly complex specification of conditions. Here the logic-based representation of the problem promises to offer some real help in identifying further interesting cases of guaranteed convergence.

This kind of question can be cast as a question in modal logic *correspondence theory* [7]. Suppose we want to identify suitable conditions on agent preferences that would allow us to guarantee convergence by means of rational deals all belonging to a class of deals D. Then we have to identify a class of frames on which formula (9) would be valid. Again, this is an issue we put forward for detailed investigation in the future.

References

1. Sandholm, T.W.: Contract types for satisficing task allocation: I Theoretical results. In: Proc. AAAI Spring Symposium: Satisficing Models. (1998)
2. Endriss, U., Maudet, N., Sadri, F., Toni, F.: Negotiating socially optimal allocations of resources. Journal of Artificial Intelligence Research **25** (2006) 315–348
3. Dunne, P.E., Wooldridge, M., Laurence, M.: The complexity of contract negotiation. Artificial Intelligence **164**(1–2) (2005) 23–46
4. Parikh, R.: Social software. Synthese **132** (2002) 187–211
5. Pacuit, E., Parikh, R.: Social interaction, knowledge, and social software. In: Interactive Computation: The New Paradigm. Springer-Verlag (forthcoming)
6. Pauly, M., Wooldridge, M.: Logic for mechanism design: A manifesto. In: Proc. 5th Workshop on Game-theoretic and Decision-theoretic Agents. (2003)

7. Blackburn, P., de Rijke, M., Venema, Y.: Modal Logic. Cambridge University Press, Cambridge (2002)
8. Harel, D., Kozen, D., Tiuryn, J.: Dynamic Logic. MIT Press, Boston (2000)
9. Moulin, H.: Axioms of Cooperative Decision Making. Cambridge University Press, Cambridge (1988)
10. van Benthem, J., van Otterloo, S., Roy, O.: Preference logic, conditionals and solution concepts in games. In: Modality Matters: Twenty-Five Essays in Honour of Krister Segerberg. University of Uppsala (2006)
11. Hansson, S.O.: Preference logic. In: Handbook of Philosophical Logic. 2nd edn. Kluwer Academic Publishers (2001)
12. Dunne, P.E., Chevaleyre, Y.: Negotiation can be as hard as planning: Deciding reachability properties of distributed negotiation schemes. Technical Report ULCS-05-009, Department of Computer Science, University of Liverpool (2005)
13. Lange, M.: Model checking propositional dynamic logic with all extras. Journal of Applied Logic **4**(1) (2005) 39–49
14. Chevaleyre, Y., Endriss, U., Lang, J., Maudet, N.: Negotiating over small bundles of resources. In: Proc. 4th International Joint Conference on Autonomous Agents and Multiagent Systems, ACM Press (2005)
15. Passy, S., Tinchev, T.: An essay in combinatory dynamic logic. Information and Computation **93**(2) (1991) 263–332
16. Balbiani, P., Vakarelov, D.: Iteration-free PDL with intersection: A complete axiomatization. Fundamenta Informaticae **45** (2001) 1–22
17. Fischer, M.J., Ladner, R.E.: Propositional dynamic logic of regular programs. Journal of Computer and System Sciences **18**(2) (1979) 194–211
18. Pratt, V.R.: Semantical considerations on Floyd-Hoare logic. In: Proc. 17th Annual Symposium on Foundations of Computer Science, IEEE (1976) 109–121
19. Lutz, C., Walther, D.: PDL with negation of atomic programs. Journal of Applied Non-Classical Logics **15**(2) (2005) 189–214
20. Danecki, R.: Non-deterministic propositional dynamic logic with intersection is decidable. In: Proc. 5th Workshop on Computation Theory. Springer-Verlag (1985)
21. Lange, M., Lutz, C.: 2-EXPTIME lower bounds for propositional dynamic logics with intersection. Journal of Symbolic Logic **70**(4) (2005) 1072–1086

A PDL and Its Extensions

In this short appendix we list the relevant results surrounding propositional dynamic logic and its extensions. Much of this information can be found in the textbook *Dynamic Logic* by Harel, Kozen and Tiuryn [8]. The reader is also referred to Passy and Tinchev [15] for more information.

Let At be a set of atomic propositions and Pr a set of atomic programs. Formulas and programs have the following syntactic form ($p \in \mathsf{At}$ and $r \in \mathsf{Pr}$):

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \psi \mid \langle\alpha\rangle\varphi$$

$$\alpha ::= r \mid \alpha \cup \beta \mid \alpha \cap \beta \mid \alpha;\beta \mid \alpha^* \mid \overline{\alpha} \mid \alpha^{-1}$$

Other connectives and operators are defined as usual. For example, $\varphi \wedge \psi = \neg(\neg\varphi \vee \neg\psi)$ and $[\alpha]\varphi = \neg\langle\alpha\rangle\neg\varphi$. Note that for simplicity we do not include the test-operator. Let $\mathcal{L}_{\mathsf{PDL}}$ be the set of all such well-formed formulas. Given an arbitrary program α, we define relations R_α as usual [8]. Formulas are interpreted in Kripke structures $\mathcal{M} = (W, \{R_r\}_{r\in\mathsf{Pr}}, V)$ where each $R_r \subseteq W \times W$ and $V : \mathsf{At} \to 2^W$. Truth in a model is defined as usual (see Section 2.3 and [8]). A model $\mathcal{M}$ is called a PDL model provided $\mathcal{M}$ and the relations R_α for any program α are defined as above. By PDL we mean the set of all formulas which are valid in any PDL model. We now survey the main results relevant for our discussion in this paper.

Harel [8] showed that assuming that all atomic programs are deterministic, PDL with intersection is highly undecidable. However, the result is more positive if we allow for arbitrary (non-deterministic) atomic programs. Balbiani and Vakarelov [16] showed that PDL with intersection is axiomatizable with the use of an infinitary proof rule. Passy and Tinchev [15] prove a similar result using nominals. Early on it was shown by Fischer and Ladner [17] that the satisfiability problem for $\mathcal{L}_{\mathsf{PDL}}$ with respect to the class of all PDL models is decidable. Pratt [18] went on to show that it is EXPTIME-complete. It was observed by Harel [8] that the validity problem with complementation is undecidable. However, recently it was shown that allowing complementation of *atomic* programs only allows us to retain decidability.

Theorem 1 (Lutz & Walther [19]). *The satisfiability problem for $\mathcal{L}_{\mathsf{PDL}}$ with complement applied only to atomic programs is decidable.*

The satisfiability problem for $\mathcal{L}_{\mathsf{PDL}}$ (with or without complement) interpreted over PDL models in which the atomic programs are deterministic is Σ^1_1-complete. If the restriction to deterministic atomic programs is dropped then the situation becomes much more manageable.

Theorem 2 (Danecki [20]; Lange & Lutz [21]). *The satisfiability problem for $\mathcal{L}_{\mathsf{PDL}}$ with intersection (but without complement) is 2-EXPSPACE-complete.*

Finally, in a recent paper Lange [13] points out that model checking $\mathcal{L}_{\mathsf{PDL}}$ formulas remains in PTIME,

Theorem 3 (Lange [13]). *The model checking problem for $\mathcal{L}_{\mathsf{PDL}}$ with respect to PDL models is in PTIME.*

Returning to the logics presented in this paper, it is not hard to see that the language $\mathcal{L}_{\mathsf{Pareto}}$ is a fragment of $\mathcal{L}_{\mathsf{PDL}}$. The idea is to interpret each preference relation R_i as an atomic program. Then the operators $\langle\text{PAR}\rangle$ and $\langle R_i^s\rangle$ become definable in $\mathcal{L}_{\mathsf{PDL}}$. Of course, this interpretation uses the converse, complement and intersection operators. Thus as remarked above, in the presence of the complement operator, the validity problem for $\mathcal{L}_{\mathsf{PDL}}$ is undecidable. However, we are working in a fragment in which the complement operator is only applied to atomic and the converse of atomic programs. The logic $\mathcal{L}_{\langle\mathcal{A},\mathcal{R}\rangle}$ is decidable due to the chosen semantics which fixes the set of possible worlds (cf. Proposition 1).