Description Logics for Multi-issue Bilateral Negotiation with Incomplete Information*

Azzurra Ragone, Tommaso Di Noia, Eugenio Di Sciascio
SisInfLab—Politecnico di Bari
via Re David 200, 70125 Bari, Italy

Francesco M. Donini
DISCOM—Università della Tuscia
via San Carlo 32, 01100 Viterbo, Italy

Abstract

We propose a framework for multi-issue bilateral negotiation, where issues are expressed and related to each other via Description Logics. Agents' goals are expressed through (complex) concepts, and the worth of goals as weights over concepts. We adopt a very general setting with incomplete information by letting agents keep both goals and worths of goals as private information. We introduce a negotiation protocol for such a setting, and discuss different possible strategies that agents can adopt during the negotiation process. We show that such a protocol converges, if the Description Logic used enjoys the finite implicants property.

Introduction

Several recent research works (Bouveret *et al.* 2005; Zhang & Zhang 2006; Ragone *et al.* 2006) propose modeling preferences (goals) in multi-issue negotiation with the aid of logic. Usually such works adopt Propositional Logic that, when equipped with a theory, allows modeling bundles of issues and implications between them.

In this paper we propose Description Logics (DLs) as languages to model agents' goals in Multi-issue Bilateral Negotiation. We have two main reasons for choosing DLs: (1) they are the basis of Semantic Web Languages as OWL[1]; (2) they can be much more expressive than Propositional Logic, yet they have decidable inference problems—satisfiability and subsumption—available as services in optimized reasoners. Satisfiability is useful to catch inconsistency between agent's goals w.r.t. the ontology $\mathcal{T}$, *i.e.*, inconsistent goals cannot be in the same agreement, (*e.g.*, agents cannot agree on A and B at the same time if in $\mathcal{T}$ A is defined as disjoint from B). Through subsumption one can discover if an agent's goal is implied by a goal of its opponent, even if this fact does not immediately appear at the syntactic level.

We remark that our framework could be applied to *every* logic whose logical implication and satisfiability are decidable. We set the stage of our approach in a generic scenario with incomplete information, and define a logic-based protocol. We prove that the protocol converges for DLs restricted by the finite implicant property.

The protocol leaves open the possibility to the agents to pursue different strategies. Among others, we study two possible strategies that agents can adopt with reference to the rules set up by the protocol and analyze relevant properties of such strategies.

As the negotiation scenario is one with fully *incomplete information*, we assume agents do not reveal their goals either to the opponent or to a mediator, but they keep as private information both goals and their worths.

Actually, the difficulty to model scenarios with incomplete information is due to the fact that the agent cannot be sure how the opponent will evaluate its offers, and therefore it may be unable to negotiate to the best of its capacity (Kraus 2001). Usually, to overcome such drawback, a preliminary step is added to strategic negotiations where agents reveal some (or all) private information. Obviously, the revelation mechanism has to ensure agents truthfully report their private information and punish liars (Rosenschein & Zlotkin 1994). Yet it is not always possible to design truthful revelation mechanisms, since they depend on the particular scenario taken into account (see (Kraus 2001, p.64) for an extensive discussion). When negotiation involves organizations, *e.g.*, companies, revealing information may conflict with company's interests and assessing the truthfulness of the agents' declarations can can be very hard or impossible.

The protocol we propose here is able to deal with such incomplete information without forcing agents to reveal either their goals or utility functions, so it suits all scenarios where agents are not willing to reveal private information or when it is hard to design a truthful revelation mechanism.

The remaining of this paper is structured as follows: first we outline the notation we adopt and define an agent in our setting together with the information characterizing its type (goals, utility functions, thresholds). Next we present the negotiation mechanism, illustrating the logic-based negotiation protocol and the convergence property of the protocol itself. Then we introduce and discuss two possible strategies compliant with the protocol, highlighting their properties. Discussion and conclusion close the paper.

*This research has been partially supported by Apulia Region Strategic Project PS_092 DIPIS (DIstributed Production as Innovative System) and by EU FP-6 IST STREP TOWL co. 026896.

[1] www.w3.org/TR/owl-features/

Notation

Hereafter, we assume the reader be familiar with Description Logics (DLs) syntax and semantics (Baader *et al.* 2003). We use symbols $A, B, C, D, \ldots, \top, \bot$ to denote concepts, and symbols $\mathcal{T}$, $\mathcal{G}$, $\mathcal{R}$, to denote sets (of inclusions, of concepts, etc.). Note that we refer both to concept names and to DL formulas with the term "concept". The symbol $\top$ denotes the universal concept, while $\bot$ the empty (contradictory) one. Here we use $\mathcal{T}$ (for Terminology) to indicate a DL ontology, *i.e.*, a set of axioms of the form $C \sqsubseteq D$ (inclusion) and $C \equiv D$ (definition) with C and D being concepts. We say C is subsumed by D w.r.t. $\mathcal{T}$ when $\mathcal{T} \models C \sqsubseteq D$, or equivalently $C \sqsubseteq_{\mathcal{T}} D$; C is not satisfiable w.r.t. the ontology $\mathcal{T}$ when it is subsumed by the most specific concept $\mathcal{T} \models C \sqsubseteq \bot$, or equivalently $C \sqsubseteq_{\mathcal{T}} \bot$; C is not subsumed by D w.r.t. $\mathcal{T}$ when $\mathcal{T} \not\models C \sqsubseteq D$, or equivalently $C \not\sqsubseteq_{\mathcal{T}} D$.

In the following, the symbol *a* denotes a generic agent, and the symbol *o* denotes *a*'s opponent.

Negotiating over related issues

Extending the well-known preference mechanism of Weighted Propositional Formulas (Pinkas 1991; Lafage & Lang 2000; Chevaleyre, Endriss, & Lang 2006), we assign normalized weights to concepts $C_1, \ldots, C_n$ in a Description Logic, and define the utility of a proposal C as the sum of the weights of the concepts subsumed by C w.r.t. a ontology $\mathcal{T}$. The agent's type embodies any private information which is not common knowledge to all agents, but is relevant to the agent's decision making (Fudenberg & Tirole 1991). In our setting the type embodies goals, their worth, an utility threshold and unwanted issues. More formally, we define agent *a* and its opponent *o* as in the following.

Definition 1 *[Agent] Let $\mathcal{T}$ be an ontology in a DL $\mathcal{L}$. Agent a is defined as a tuple $(\mathcal{G}_a, u_a(\cdot), t_a, \mathcal{R}_a)$, where:*

- *$\mathcal{G}_a$ is a set of concepts $\{C_1, \ldots, C_n\}$, each one satisfiable in $\mathcal{T}$, representing a's* goals. *To each goal C_i is associated a* weight *$w_i^a \in \Re^+$ such that $\sum_{i=1}^n w_i^a = 1$.*
- *$u_a(\cdot) : \mathcal{L} \to \Re^+ \cup \{0\}$ is a function that assigns a worth (*utility*) to every concept $C \in \mathcal{L}$ as follows:*

$$u_a(C) = \sum \{w_i^a \mid i = 1, \ldots, n, \text{ and } C \sqsubseteq_{\mathcal{T}} C_i\}$$

- *t_a is a* threshold *for a's utility. If an agreement is reached such that the corresponding a's utility is lower than t_a, then a rejects the agreement (conflict deal).*
- *$\mathcal{R}_a$ is the set of concepts that a explicitly* rejects *during the negotiation.*

In an analogous way, the opponent o is defined as $(\mathcal{G}_o, u_o(\cdot), t_o, \mathcal{R}_o)$.

Note that the weight of a goal may not be its worth: *e.g.*, let $\mathcal{T}$ embody knowledge about guarantees, and let C_1, C_2 represent in $\mathcal{T}$ a guarantee of at least one year and at least two years, respectively, both with weight $w_1^a = w_2^a = 0.1$; then $u_a(C_2) = 0.2 \neq w_2^a$ since $C_2 \sqsubseteq_{\mathcal{T}} C_1$.

Here we are not interested in how to compute weights w_i^a (or w_j^o); without loss of generality, we might assume the use of either direct assignment methods (Ordering, Simple Assessing or Ratio Comparison) or pairwise comparison methods (AHP and Geometric Mean) (Pomerol & Barba-Romero 2000). $\mathcal{R}_a$, $\mathcal{R}_o$ are sets of concepts declared unacceptable by *a* and *o*, respectively, *i.e.*, if $N \in \mathcal{R}_a$ then *a* declared that it will reject any agreement implying N. Both agents can start negotiation with an empty or a non-empty $\mathcal{R}$, depending on their own strategy. We build on the definition of Multi-issue Bilateral Negotiation (MBN) problem as proposed in (Ragone *et al.* 2006) and extend it as follows:

Definition 2 (MBN) *Given two agents a and o, a **M**ulti-issue **B**ilateral **N**egotiation (MBN) problem is finding a concept A (for **A**greement) in $\mathcal{L}$, consistent w.r.t. $\mathcal{T}$ and such that:*

1. *$u_a(A) \geq t_a$ and $u_o(A) \geq t_o$;*
2. *for each concept $C_i \in \mathcal{G}_a$ either $A \sqsubseteq_{\mathcal{T}} C_i$ or $A \sqsubseteq_{\mathcal{T}} \neg C_i$ and for each concept $C_j \in \mathcal{G}_o$ either $A \sqsubseteq_{\mathcal{T}} C_j$ or $A \sqsubseteq_{\mathcal{T}} \neg C_j$;*
3. *for each concept $N_i \in \mathcal{R}_a$ and $N_j \in \mathcal{R}_o$ both $A \not\sqsubseteq_{\mathcal{T}} N_i$ and $A \not\sqsubseteq_{\mathcal{T}} N_j$ hold.*

In the above definition, Condition 1 states that agents accept an agreement A iff their utility is greater or equal to their threshold's utility. With the second condition we impose all the goals have to be decided in the final agreement A, *i.e.*, the agent is able to compute if each goal is attained or not. Finally, the last condition says that an agent will never accept any agreement implying any unwanted issue.

A Cumulative Protocol

Following the idea of Parsons, Sierra, & Jennings (1998), here we introduce a DL-protocol for MBN based on proposals and counter-proposals. Basically, in this protocol agent *a* makes a proposal C to agent *o* trying to satisfy its goals in $\mathcal{G}_a$ and avoiding elements in $\mathcal{R}_a$. On the other hand, *o* may decide to accept, reject or refine *a*'s proposal taking into account its own goals $\mathcal{G}_o$ and unwanted issues modeled in $\mathcal{R}_o$. If *o* decides to refine C proposing C', then it is up to *a* the decision to accept/reject/refine C' and so on, until an agreement is reached or an agent opts out. It is easy to see that whenever during the negotiation either *a* or *o* decide to accept, they agree on some issues they are negotiating on. Hence, at each step of the protocol there is a part of the agreement they already set.

In our framework, given a DL $\mathcal{L}$, a *proposal* is a concept $C \in \mathcal{L}$. Similarly, A_{sf} is a concept in $\mathcal{L}$ representing aspects the agents **A**greed **S**o **F**ar in the negotiation. When negotiation starts, $A_{sf} = \top$.

Reasonably, an agent makes a proposal because it expects increasing its own utility with respect to the agreement so far A_{sf}.

Definition 3 *[Utility of a proposal] Given a proposal C made by agent a, the* utility of C for a with respect to A_{sf} *is $u_a(C \sqcap A_{sf})$.*

i.e., $u_a(C \sqcap A_{sf})$ is the sum of all utilities of goals attained by *a* if C is accepted.

Hereafter, when no ambiguity arises, we write $u_a(C)$ instead of $u_a(C \sqcap A_{sf})$.

Definition 4 (Motivated proposal) *Let $\mathcal{T}$ be a reference ontology for both a and o, and let A_{sf}, $\mathcal{R}_o$ as in the previous section. A concept C is a* motivated proposal *for Agent a, w.r.t. $\mathcal{T}$, A_{sf}, $u_a(\cdot)$, if all of the conditions below hold:*

1. *$A_{sf} \sqcap C$ is satisfiable w.r.t. $\mathcal{T}$*
2. *$A_{sf} \sqcap C \not\sqsubseteq_{\mathcal{T}} N$ for every $N \in \mathcal{R}_o$*
3. *$u_a(C) > u_a(\top)$*

First condition states that new proposals cannot be in contrast with what has been agreed so far. Second condition requires that *a* must take into account what *o* already declared unacceptable. Third condition requires that C should actually increase *a*'s utility w.r.t. what *a* already gained with A_{sf}. Observe that $\top$ represents a shallow proposal (it implies no goal), and that $\top$ is never a motivated proposal, because it fails on Condition 3.

In what follows, we require that both agents make only motivated proposals. In doing so, we assume agents are self-interested and willing to conclude the negotiation.

Theorem 1 *Condition 3 in Definition 4 is equivalent to the following one: there exists a concept D such that*

1. *$u_a(D) > 0$, and*
2. *$A_{sf} \not\sqsubseteq_{\mathcal{T}} D$, and*
3. *$A_{sf} \sqcap C \sqsubseteq_{\mathcal{T}} D$*

Proof. If. If such a D exists, then proposing C should add at least $u_a(D)$ to global utility of *a*, implying Condition 3. *Only if.* Applying Defs. 1 and 3, Condition 3 can be rewritten as $\sum\{w_i^a \mid i = 1, \ldots, n, \text{ and } C \sqcap A_{sf} \sqsubseteq_{\mathcal{T}} C_i\} > \sum\{w_i^a \mid i = 1, \ldots, n, \text{ and } A_{sf} \sqsubseteq_{\mathcal{T}} C_i\}$. This implies that there exists an $h \in \{1, \ldots, n\}$ such that $C_h \in \mathcal{G}_a$, and both $A_{sf} \not\sqsubseteq_{\mathcal{T}} C_h$ and $A_{sf} \sqcap C \sqsubseteq_{\mathcal{T}} C_h$. Then the claim holds with $D = C_h$, since $u_a(C_h) \geq w_h^a > 0$. □

Negotiation starts with $A_{sf} = \top$ (nothing agreed so far), and with one agent making a (motivated) proposal C. We leave open in the protocol which agent should start. Then, the agent that *receives* the proposal C—say, agent *a*—has the following options:

1. **reject**: let $\mathcal{R}_a := \mathcal{R}_a \cup \{C\}$;
2. **accept**: let $A_{sf} := A_{sf} \sqcap C$;
3. **refine**: propose $C \sqcap D$ such that $C \sqcap D$ is a motivated proposal for *a* w.r.t. $\mathcal{T}, A_{sf} \sqcap C, u_a(\cdot)$.

For Options 1–2, it is left open who will propose next; for Option 3, *o* should receive $C \sqcap D$ using the options above.

A ***round*** is either a proposal or one of the three actions above (reject, accept, refine). Negotiation ends when no agent can make a motivated proposal, which means no agent can increase its utility by making a consistent addition to A_{sf}. At this point, each agent *a* is supposed to accept A_{sf} as the specification of a deal if $u_a(C) > t_a$ (the threshold utility for *a*), and reject otherwise; if both agents accept, the deal is reached, otherwise the conflict deal is the outcome.

Observe that the satisfiability condition in Definition 4 forces A_{sf} to be always satisfiable w.r.t. $\mathcal{T}$.

Convergence

When using the Cumulative Protocol, an agent *a* could propose directly its goals $C_1, \ldots, C_n$, one at a time, or as many as possible at once, or some combination of the goals, depending on *a*'s strategy. However, proposing a goal *reveals* it to the opponent, who can exploit such a knowledge. Instead, the Cumulative protocol allows an agent to propose a concept D which only *implies* one or more goals C. The convergence of the protocol depends on how many motivated proposals can be made, and rejected, before negotiation ends.

We start by extending to DLs a standard terminology in Propositional and First-Order Logic. Given a concept C, and a TBox $\mathcal{T}$, we say that D is an *implicant* of C w.r.t. $\mathcal{T}$ if $D \sqsubseteq_{\mathcal{T}} C$. Two implicants D_1, D_2 of C are *independent* if neither $D_1 \sqsubseteq_{\mathcal{T}} D_2$ nor $D_2 \sqsubseteq_{\mathcal{T}} D_1$. A description logic $\mathcal{L}$ has the *finite implicants* property if for every concept $C \in \mathcal{L}$, the set of concepts $\{D \in \mathcal{L} \mid D \sqsubseteq C\}$ is finite, and $\mathcal{L}$ has the *finite independent implicants* if every set of mutually independent implicants is finite. Observe that the latter notion rules out syntactic variants of semantically equivalent formulas.

While Propositional Logic has the finite independent implicants property, DLs in general do not. Consider the very basic DL $\mathcal{FL}_0$ (Nebel 1990), whose connectives are just $\sqcap$ and $\forall R.C$. Given a concept $C \in \mathcal{FL}_0$, and a concept name A not appearing in C, a set of independent implicants of C is $D_0 = C \sqcap A$, $D_1 = C \sqcap \forall R.A$, $D_2 = C \sqcap \forall R.\forall R.A, \ldots$ Therefore, no DL has the finite implicants property. However, taking into account our application to negotiation, several meaningful syntactic restrictions would yield the finite implicants property: for instance, a bound on the quantification nesting of concepts; or a restriction that allows an agent to propose only concept names (defined in the TBox); or to propose only concepts of the syntactic closure of $\mathcal{T}$ and the goals of both agents.

Theorem 2 *Let $\mathcal{L}$ be a description logic with the finite implicants property; let $u_a(\cdot), u_o(\cdot)$ two utility functions over n_a and n_o concepts in $\mathcal{L}$, respectively; let $\mathcal{T}$ be a TBox. Every run of the Cumulative Protocol terminates in a finite number of rounds.*

Proof. First we prove that there cannot be an infinite sequence of refinements. In fact, every refinement $C \sqcap D$ must be a motivated proposal, hence from Theorem 1 it must "use" a goal not previously attained. Hence, after at most $n_a + n_o$ refinements, either an "accept" or a "reject" must be issued by the agent who cannot make further refinements.

Then, we prove that at each rejection or acceptance, a finite set of concepts is reduced by one. Regarding acceptance, since the accepted concept is a motivated proposal, from Theorem 1 there is at least one goal D which is implied by A_{sf} after acceptance. Hence, D cannot be the objective of any further motivated proposal. Regarding rejection, we use the finite implicants property of $\mathcal{L}$: every goal has a finite number of implicants. When a motivated proposal C made by *a* is rejected by *o*, at least one implicant—C itself—of a goal cannot be proposed anymore by *a*, since now C is in $\mathcal{R}_o$, and vice versa for implicants of *o*'s goals. Therefore, there cannot be an infinite number of acceptances and rejections. □

The above theorem shows that the representation language is not neutral w.r.t. the negotiation mechanism. Some languages may allow agents to go astray, while others, limiting the number of possible proposals, force termination sooner of later.

Strategies

We now analyze possible strategies that the agents may adopt when negotiating with a Cumulative protocol. Given an ontology $\mathcal{T}$, an ongoing agreement A_{sf}, and previous rejections $\mathcal{R}_a$, $\mathcal{R}_o$, a goal C of a is *attainable* w.r.t. A_{sf}, $\mathcal{T}$, $\mathcal{R}_o$ if both $A_{sf} \not\sqsubseteq_{\mathcal{T}} C$, and for every $N \in \mathcal{R}_o$, $A_{sf} \sqcap C \not\sqsubseteq_{\mathcal{T}} N$. That is, C does not imply any concept N already rejected by o. In other words, at a given stage A_{sf}, $\mathcal{R}_a$, $\mathcal{R}_o$ of the negotiation, a can expect to gain the utility of each attainable goal—although maybe not all together. In what follows, we refer to the set of attainable goals of an agent.

IWIN

The first strategy we consider is IWIN: when making a proposal or a refinement, an agent maximizes immediately its utility—of course, always relative to satisfiability with A_{sf}.

Definition 5 (SUP) *Given an ontology $\mathcal{T}$, an ongoing agreement A_{sf}, and previous rejections $\mathcal{R}_a$, $\mathcal{R}_o$, let $\{C_1, \ldots, C_k\}$ be the set of a's attainable goals. Then C is a* saturated utility *motivated proposal (SUP) w.r.t. $\mathcal{T}$, A_{sf}, $u_a(\cdot)$, if all conditions below hold:*

1. *C is a* motivated proposal *for a, w.r.t. $\mathcal{T}$, A_{sf}, $u_a(\cdot)$,*
2. *either $A_{sf} \sqcap C \sqsubseteq_{\mathcal{T}} C_i$ or $A_{sf} \sqcap C \sqsubseteq_{\mathcal{T}} \neg C_i$, $\forall i \in \{1, \ldots, k\}$*
3. *$u_a(C)$ is maximal over all possible motivated proposals.*

Hence a SUP is a proposal in which a decides every goal of hers not already decided, getting the worth of the goal if possible. Observe that, in general, the concept $C_1 \sqcap \cdots \sqcap C_k$ may not be a SUP, since it may be unsatisfiable by itself or w.r.t. A_{sf} and $\mathcal{T}$. For instance, if a sets increasing utilities over several mutually disjoint alternatives (*e.g.*, colors, methods of payment, ways of shipping, etc.), then a cannot attain all such goals at once.

When *receiving* a proposal C, IWIN suggests that a should:

1. **reject** if a has a proposal C' starting from A_{sf}, such that $u_a(C') > u_a(C)$, and $A_{sf} \sqcap C \sqcap C'$ is unsatisfiable w.r.t. $\mathcal{T}$; otherwise
2. **refine** proposing $C \sqcap D$, if there is a concept D which is a SUP w.r.t. $\mathcal{T}$, $A_{sf} \sqcap C$, $u_a(\cdot)$; otherwise
3. **accept** if there is neither a proposal nor a refinement better than C.

If both agents adopt this strategy, refinements end in one round: a makes a SUP C (w.r.t. $\mathcal{T}$, A_{sf}, $u_a(\cdot)$), o may refine it adding a SUP D (now w.r.t. $\mathcal{T}$, $A_{sf} \sqcap C$, $u_o(\cdot)$), and then a can only accept or reject it, since no motivated addition could be further made on $C \sqcap D$ (every attainable goal is implied by C). If a accepts, negotiation ends; if a rejects, a new SUP can be made by either agent.

This strategy is naive (it reveals all of one's true goals) and computationally demanding: it is (at least) NP-hard to find a SUP, since *e.g.*, the problem MAX-SAT (given a set of clauses, find an assignment satisfying the maximum number of them) can be immediately reduced to finding a SUP. However, maximization of utility in SUP yields Pareto-efficiency w.r.t. the agreements allowed by A_{sf}, as formalized by the following theorem.

Theorem 3 *If both agents adopt IWIN, then every accepted proposal is Pareto-efficient w.r.t. A_{sf}.*

Proof. Suppose C is a proposal that is accepted by a. Since a adopts a saturated utility strategy, if a accepts C then it must be so because a cannot make any refinement on C—that is, $A_{sf} \sqcap C$ already decides (positively or negatively) every goal of a. We distinguish two cases: (i) C was a proposal of o, (ii) C was a refinement of o on a proposal of a.

Suppose C was a SUP made by o. Then, if C is not Pareto-efficient there should be a concept C' such that either $u_a(C) \leq u_a(C')$ or $u_o(C) \leq u_o(C')$, with at least one strict inequality. But if the latter inequality is strict, then C would not be a SUP (o should have offered C' instead). If the former inequality is strict, then a would reject C and propose C' or better in the next round. Contradiction.

Suppose now C is a refinement of o on a previous SUP C_a made by a. If o had a better refinement or proposal C' then it would have not refined C_a as C, hence there does not exist a C' such that $u_o(C) \leq u_o(C')$. On the other hand, if a had a better proposal C', then it would have proposed C' instead of C_a. Again, such a C' does not exist. □

Observe that we had to state Pareto-efficiency only relative to agreements allowed by A_{sf}, that is, conditions which both agents agreed so far. The strategy allows an agent a to reject a convenient proposal C, and in the rest of the negotiation C is added to $\mathcal{R}_a$. In this way, later on agents may agree on a proposal D, which was dominated by C in the initial setting, but which becomes Pareto-efficient after C has been excluded by $\mathcal{R}_a$.

Of course, this strategy yields highly *unfair* distributions of utility in proposals: namely, the agent making a SUP gets more than the opponent who can make only refinements. Therefore, a SUP will likely be rejected by a self-interested opponent, without any refinement at all. Hence, a negotiation where both agents immediately saturate their utility is likely to be a very long sequence of SUP-s and rejections.

Theorem 4 (IWIN vs. IWIN) *If both agents adopt IWIN, the number of steps needed to end negotiation can be, in the worst case, exponential in n_a, n_o, and the size of $\mathcal{T}$, even when agents are allowed to propose only conjunctions of their goals.*

Proof. Let o have the goals $C_1^o, \ldots, C_n^o$, and a the goals $C_1^a, \ldots, C_n^a$ (same number n of goals), with n even, all goals with weight $\frac{1}{n}$ for each agent: $w_i^a = w_i^o = \frac{1}{n}$ for every $i = 1, \ldots, n$. Moreover, let $\mathcal{T} = \{C_i^o \sqcap C_i^a \sqsubseteq \bot \mid i =$

$1,\ldots,n\}$. This is equivalent to say that agents fully compete in a zero-sum game. Then, no matter who starts, both agents will propose their combinations of $n, n-1, n-2, \ldots$ goals, rejecting the opponent's proposals until they reach an agreement on sharing $\frac{n}{2}$ goals each. To reach such an agreement, they need a number of proposals which is the sum of binomial coefficients from $\binom{n}{n}$ down to $\binom{n}{n/2}$, which is $O(2^{cn})$. □

Observe that the above theorem holds independently of the thresholds. Interestingly, in the case of the above proof if one drops axioms in the ontology—*i.e.*, $\mathcal{T} = \emptyset$—one obtains a total cooperation game, in which there is an agreement $C_1^a \sqcap \cdots \sqcap C_n^a \sqcap C_1^o \sqcap \cdots \sqcap C_n^o$ in which each agent can get all of its utility. If both agents adopt IWIN, they reach it in two rounds—a SUP and a refinement. Hence, the exponential behavior of IWIN seems to depend on the number of incompatible goals between the two agents.

Compensating Conceder (CoCo)

Real life negotiations are often cooperative rather than non-cooperative and maximizing utility is not the only criterion to pursue, but also fairness, establishing a reputation or trust relation and maintaining good relationships would also play a major role. Hence a different strategy with the above characteristics, that may converge more rapidly to a fair agreement, is the following one: when making a proposal, an agent a proposes C such that C implies one of the most valuable a's goals, leaving other goals to be decided in further negotiation rounds. More formally,

Definition 6 (BFMP) *Let $\{C_1,\ldots,C_k\}$ the set of a's attainable goals w.r.t. A_{sf}, $\mathcal{T}$, $\mathcal{R}_o$. Then C is a* best first *motivated proposal (BFMP) w.r.t. $\mathcal{T}$, A_{sf}, $u_a(\cdot)$, if there exists one $i \in \{1,\ldots,k\}$ such that $A_{sf} \sqcap C \sqsubseteq_{\mathcal{T}} C_i$, and moreover, $u_a(C_i)$ is maximal among $\{u_a(C_1),\ldots,u_a(C_k)\}$.*

When issuing a proposal, an agent adopting CoCo should propose a BFMP. Of course, such an agent should also protect itself from more aggressive strategies such as IWIN. Hence, when *receiving* a proposal C, CoCo suggests that a should first evaluate how much of a's own utility is definitely lost—and possibly, how much is gained—if a accepts C, as follows.

Definition 7 (Evaluation) *Let C be a proposal by o and let $\{C_1,\ldots,C_k\}$ the set of a's attainable goals w.r.t. A_{sf}, $\mathcal{T}$ and $\mathcal{R}_o$. Then, a's* evaluation *$E(C, A_{sf}, \mathcal{T})$ about C is $\sum\{w_i^a \mid A_{sf} \sqcap C \sqsubseteq_{\mathcal{T}} C_i\} - \sum\{w_i^a \mid A_{sf} \sqcap C \sqsubseteq_{\mathcal{T}} \neg C_i\}$*

Intuitively, the evaluation of C made by a is the sum of utilities that a gains minus the sum of utilities that a loses for sure if a accepts C. The CoCo strategy then suggests to accept C if its evaluation is positive, otherwise refine C as $C \sqcap D$ in a way that the loss induced by accepting C is compensated by D—if possible—and reject otherwise. Some care must be taken for the last proposal (the one which "burns" the last goal), otherwise since the last goal cannot have a compensation CoCo would always reject in the end. More precisely, CoCo suggests to:

1. **accept** if either $E(C, A_{sf}, \mathcal{T}) \geq 0$, or if the set of a's attainable goals w.r.t. C, A_{sf}, $\mathcal{T}$ is empty and $u_a(C) \geq t_a$ (a's threshold); otherwise
2. **refine** proposing $C \sqcap D$ if $E(C, A_{sf}, \mathcal{T}) < 0$ and there exists a concept D such that
   (a) either $E(C \sqcap D, A_{sf}, \mathcal{T}) \geq 0$, and D is a BFMP w.r.t. $\mathcal{T}$, $A_{sf} \sqcap C$, $u_a(\cdot)$,
   (b) or there are no attainable goals w.r.t. $A_{sf} \sqcap C \sqcap D$, and $u_a(A_{sf} \sqcap C \sqcap D, A_{sf}, \mathcal{T}) \geq t_a$;
3. **reject** otherwise.

Observe that Case 2 implies that the set of a's attainable goals w.r.t. $A_{sf} \sqcap C$, $\mathcal{T}$, $\mathcal{R}_o$, is *not* empty, otherwise there cannot be any concept D compensating the evaluation of C. In particular, in Case 2(a) a has enough still attainable goals to completely compensate C, while in Case 2(b) although a cannot completely compensate, it can make a refinement whose utility exceeds a's threshold. Instead, when C leaves a no more attainable goals, either Case 1 or Case 3 applies.

When both agents adopt CoCo, negotiations tend to end very quickly on a fair agreement, if one exists. For instance, in the case used in the proof of Theorem 4, negotiation ends in $2n$ rounds on an agreement which evenly divides goals between o and a. Clearly, since CoCo uses a greedy approach, the final agreement might not be Pareto-efficient. Yet the decision about which is the next BFMP to propose is a look-up on attainable goals, and so is the refinement. Hence CoCo has a low computational effort compared to IWIN.

Comparing IWIN and Coco

When one agent—say, a—adopts CoCo and the opponent adopts IWIN, a will reject all of o's proposals and refinements until o makes a SUP that either directly exceeds a's threshold, or can be refined to an acceptable proposal. Here is evident that the threshold t_a is not just encoding overheads and a's general attitude towards concluding the negotiation; t_a encodes also CoCo's defense against aggressive opponents.

Theorem 5 (CoCo vs. IWIN) *If agent a adopts CoCo, and agent o adopts IWIN, the number of steps needed to end negotiation can be, in the worst case, exponential in n_a, n_o, and the size of $\mathcal{T}$, even when agents are allowed to propose only conjunctions of their goals.*

Proof. As in the proof of Theorem 4, let $\mathcal{G}_o = C_1^o,\ldots,C_n^o$, and the $\mathcal{G}_a = C_1^a,\ldots,C_n^a$, with n even, all goals worth $\frac{1}{n}$ for each agent. Again, let $\mathcal{T} = \{C_i^o \sqcap C_i^a \sqsubseteq \perp \mid i = 1,\ldots,n\}$ setting up a zero-sum game. Suppose that o always proposes first (a SUP): then, a rejects all combinations of $n, n-1, n-2, \ldots$ goals until o proposes a combination of $n-k$ goals such that $\frac{k}{n} \geq t_a$, that is, $k = \lceil n \cdot t_a \rceil$. Then a accepts such a deal. To reach it, however, it is necessary a number of rounds that is the sum of $\binom{n}{n}$ down to $\binom{n}{n-k} = \binom{n}{\lceil n(1-t_a) \rceil}$, which is again an exponential in n. If instead a proposes first, then o always refines a's proposals with a saturation of remaining goals, obtaining a's rejections until the same bound above is reached. □

When time is brought into our framework as a private bound set by each agent on the number of rounds, agents adopting IWIN will tend to reach no agreement at all in competitive cases, both against other IWIN-agents, and CoCo-agents. In contrast, when both agents adopt CoCo, they tend more often (if their threshold is not too high) to reach agreements, although non-optimal ones.

Discussion and conclusion

A number of research works have focused on automated negotiation among agents and on design of high-level protocols for agents interaction; here we only give a brief overview of the most prominent ones dealing with logic-based approaches to negotiation. Several papers are based on Propositional Logic, including Bouveret *et al.* (2005) and the work by Chevaleyre, Endriss, & Lang (2006), where weighted Propositional formulas in preference modeling were considered. However, in such papers, no semantic relation among issues is taken into account. In our approach we adopt a logical theory $\mathcal{T}$, *i.e.*, an ontology, which allows us to model a set of consistent goals $\mathcal{G}_a$, to catch inconsistencies between agent's goals and to reach a satisfiable agreement. Wooldridge & Parsons (2000) define an agreement as a model for a set of formulas from both agents, but agent's preferences are not taken into account. Ragone *et al.* (2006) propose a Propositional Logic framework endowed of an ontology $\mathcal{T}$, where a one-shot protocol is exploited to reach Pareto-efficient agreements. In order to reach a Pareto-efficient agreement they need a trusted mediator, to whom agents reveal their goals, which suggests to the agents the most suitable agreement. With reference to the work of Zhang & Zhang (2006), that also adopts Propositional Logic, we note that in our approach we use a *knowledge base*, or ontology $\mathcal{T}$, of formulas which are common knowledge for both agents, whose constraints must always be enforced in the negotiation outcomes, while Zhang & Zhang consider common knowledge as just more entrenched preferences, that could be even dropped in some deals. Moreover we use *additive utilitites* over formulas: this allows an agent to make compensations between its requests and its concessions, while in (Zhang & Zhang 2006) the concession of a more entrenched formula can never be compensated by less entrenched ones, no matter how many they are. Finally we devised a *protocol* which the agents should adhere to while negotiating; in contrast Zhang & Zhang take a game-theoretic approach, presenting no protocol at all, since communication between agents is not considered. To the best of our knowledge our approach is the first one using DLs to design a logic-based negotiation mechanism, ensuring expressiveness greater than Propositional Logic. Moreover, w.r.t. to non-logic-based approaches, using an ontology $\mathcal{T}$ allows exploiting inference services that make possible to catch inconsistency and subsumption relations between goals. We have set our negotiation model in a *worst-case* scenario with no mediator and incomplete information, and illustrated a protocol ensuring the convergence under some syntactic restrictions of the DLs used. Such a protocol allows agents using different strategies; here we have presented two possible ones and shown that if both agents adopt an IWIN strategy the negotiation can end in a number of steps which is exponential in the size of agents' definition. Adopting a CoCo strategy—a greedy approach—they may not always reach a Pareto-efficient agreement, but the negotiation tends to end in a fair agreement more quickly and with lower computational cost than the IWIN strategy. Also in the case when an agent adopts an IWIN strategy and the opponent a CoCo strategy the number of rounds to reach an agreement can be exponential. So it would seem that for agents a good way to reach a fair agreement in a reasonable amount of time is to adopt a CoCo strategy. Future work will include the definition of alternative strategies in more restricted frameworks, including the presence of a mediator, the implementation of a prototype test-bed to numerically evaluate best strategies to adopt w.r.t. the negotiation mechanism.

References

Baader, F.; Calvanese, D.; Mc Guinness, D.; Nardi, D.; and Patel-Schneider, P., eds. 2003. *The Description Logic Handbook*. Cambridge University Press.

Bouveret, S.; Lemaitre, M.; Fargier, H.; and Lang, J. 2005. Allocation of indivisible goods: a general model and some complexity results. In *Proc. of AAMAS'05*, 1309–1310.

Chevaleyre, Y.; Endriss, U.; and Lang, J. 2006. Expressive power of weighted propositional formulas for cardinal preference modeling. In *Proc. of KR 2006*, 145–152.

Fudenberg, D., and Tirole, J. 1991. *Game Theory*. The MIT Press.

Kraus, S. 2001. *Strategic Negotiation in Multiagent Environments*. The MIT Press.

Lafage, C., and Lang, J. 2000. Logical representation of preferences for group decision making. In *Proc. of KR 2000*, 457–468.

Nebel, B. 1990. Terminological reasoning is inherently intractable. *Artificial Intelligence* 43:235–249.

Parsons, S.; Sierra, C.; and Jennings, N. 1998. Agents that reason and negotiate by arguing. *J. of Log. and Comp.* 8(3):261–292.

Pinkas, G. 1991. Propositional non-monotonic reasoning and inconsistency in symmetric neural networks. In *Proc. of IJCAI'91*, 525–531.

Pomerol, J., and Barba-Romero, S. 2000. *Multicriterion Decision Making in Management*. Kluwer Series in Operation Research. Kluwer Academic.

Ragone, A.; Di Noia, T.; Di Sciascio, E.; and Donini, F. 2006. A logic-based framework to compute pareto agreements in one-shot bilateral negotiation. In *Proc. of ECAI'06*, 230–234.

Rosenschein, J. S., and Zlotkin, G. 1994. *Rules of Encounter*. MIT Press,.

Wooldridge, M., and Parsons, S. 2000. Languages for negotiation. In *Proc. of ECAI 2000*, 393–400.

Zhang, D., and Zhang, Y. 2006. A computational model of logic-based negotiation. In *Proc. of AAAI 2006*, 728–733.