A Logical Formulation for Negotiation Among Dishonest Agents*

Chiaki Sakama
Dept. Computer and Communication Sciences
Wakayama University
Wakayama 640-8510, Japan
sakama@sys.wakayama-u.ac.jp

Tran Cao Son and Enrico Pontelli
Department of Computer Science
New Mexico State University
Las Cruces, NM 88003, USA
{tson, epontell}@cs.nmsu.edu

Abstract

The paper introduces a logical framework for negotiation among *dishonest* agents. The framework relies on the use of *abductive logic programming* as a knowledge representation language for agents to deal with incomplete information and preferences. The paper shows how intentionally false or inaccurate information of agents could be encoded in the agents' knowledge bases. Such disinformation can be effectively used in the process of negotiation to have desired outcomes by agents. The negotiation processes are formulated under the answer set semantics of abductive logic programming and enable the exploration of various strategies that agents can employ in their negotiation.

1 Introduction

Negotiation has been an important research topic in multi-agent systems and artificial intelligence, and several formalisms have been developed to model negotiations among agents (e.g., [Sadri *et al.*, 2002; Kakas and Moraitis, 2006; Amgoud *et al.*, 2007; Son and Sakama, 2009]). In real-life negotiation, it is a common practice for one to misstate their bargaining positions to gain his/her advantage over the other [Carson, 2010]. Such a bargaining tactic is effective when complete information is unavailable wrt the other party. Suppose a dialogue between a buyer b and a seller s.

b_1 : *"I like a digital camera by the maker C. I want to get one that has good quality at a discount price."*

s_1 : *"The product A is made by C and has good quality. We provide a discounted price to students."* (In reality, the seller does not know the quality of A.)

b_2 : *"I am not a student."*

s_2 : *"The product B by the maker D is on bargain sale. It has good quality and is provided at a discount price for every customer paying in cash."* (In reality, the seller knows that B is not of good quality.)

b_3 : *"I do not want products by the maker D at the price."*

s_3 : *"If you join our mailing list, we can provide the product at the lowest price."*

b_4 : *"I'd like to join the list and buy it at the price."* (In reality, the buyer does not want to join the list.)

*Partially supported by NSF grant IIS-0812267.

In this negotiation, the seller has the goal of selling a product while the buyer wants to buy a product. Although fairly simple, the negotiation highlights several difficulties that need to be addressed in any framework for formalizing negotiation: (**i**) *incomplete information*: the seller does not know whether the buyer is a student or not at the beginning and comes to learn that the buyer is not a student only during the negotiation; (**ii**) *dishonesty*: the seller intentionally misstates the quality of the product A or B to achieve his goal, while the buyer intentionally agrees to join the mailing list only to get the deal which, he thought, is a good one; and (**iii**) *preference and goal change*: the buyer prefers a product made by the maker C but ends up buying a product made by the maker D.

There are some studies that can manage incomplete information of agents or preferences, however, few studies handle dishonesty that arises in negotiation. The goal of this paper is providing an abstraction of real-life negotiation where people may behave dishonestly. To represent incomplete information and preferences, we use *abductive logic programming* (ALP) [Kakas *et al.*, 1998] and introduce the notion of *abductive programs with disinformation* (*ALD-program*). ALD-programs can represent disinformation and realize dishonest reasoning by agents. We formulate negotiation using *negotiation knowledge bases* represented by ALD-programs and explore various strategies used in negotiation. The proposed framework realizes the above mentioned issues in a single framework that makes our work significantly different from previously developed models of negotiation. Moreover, the use of logic programming enables to realize a platform for negotiation systems on top of the existing answer set solvers.

In the rest of this paper, we start by reviewing preferential reasoning in ALP and introduce the framework of ALD-programs in Section 2. The notions of negotiation knowledge bases and proposals are introduced in Section 3. Negotiations and strategies are formulated in Section 4. We relate our work to others and discuss future work in Section 5. Due to lack of space, we omit the proofs of propositions in this paper.

2 Dishonest Reasoning by Abductive Logic Programming

2.1 Abductive Programs with Preferences

A *(logic) program* consists of *rules* of the form:

$$\ell_1; \cdots ; \ell_l \leftarrow \ell_{l+1}, \ldots, \ell_m, \, not\, \ell_{m+1}, \ldots, not\, \ell_n$$

where each ℓ_i ($n \geq m \geq l \geq 0$) is a positive/negative literal of a propositional language.[1] The symbol ; represents disjunction and *not* is *negation as failure*. The left-hand side of the rule is the *head*, and the right-hand side is the *body*. Given a rule r of the above form, $head(r) = \{\ell_1, \ldots, \ell_l\}$. A rule with the empty head is a *constraint*, while a rule with the empty body is a *fact*. A fact $\ell \leftarrow$ is identified with a literal ℓ. The semantics of a program is given by the *answer set semantics* [Gelfond and Lifschitz, 1991].[2]

An *abductive program* is a pair $\langle P, \mathcal{A} \rangle$ where P and $\mathcal{A}$ are (propositional) programs. Every element in $\mathcal{A}$ is called an *abducible*. An abducible $a \in \mathcal{A}$ is also called an *abducible rule* (resp. *abducible fact*) if a is a rule (resp. a fact). Without loss of generality, we assume that literals in the heads of rules of $\mathcal{A}$ do not occur in the heads of rules of P [Kakas *et al.*, 1998]. Abducibles are hypothetical rules which are used to account for an observation together with the background knowledge P. A set S of literals is a *belief set* of $\langle P, \mathcal{A} \rangle$ if S is a consistent answer set of $P \cup E$ for some $E \subseteq \mathcal{A}$. An abductive program $\langle P, \mathcal{A} \rangle$ is *consistent* if it has a belief set; otherwise, it is *inconsistent*. In what follows, we associate a *name* n_r to each rule r and freely use the name to represent the rule.[3]

When multiple sets of abducible rules can be used to generate belief sets, a *preference* relation among abducibles is introduced [Sakama and Inoue, 2000]. Given two abducibles $n_1, n_2 \in \mathcal{A}$, P can include atoms of the form $n_1 < n_2$ meaning that n_2 is *preferred* to n_1. The relation $<$ is a strict partial order that is transitive and asymmetric. The semantics of abductive programs with such preference relations is defined as follows. First, the relation $<$ is extended to define preference among sets of abducible rules: for $Q_1, Q_2 \subseteq \mathcal{A}$, Q_1 is *preferred* to Q_2 if either **(i)** $Q_1 \subseteq Q_2$ or **(ii)** there is $n_1 \in Q_1 \setminus Q_2$ such that $n_2 < n_1$ for some $n_2 \in Q_2 \setminus Q_1$ and $n_1 \not< n_3$ for any $n_3 \in Q_2 \setminus Q_1$. In turn, this provides a means to compare belief sets of an abductive program $\langle P, \mathcal{A} \rangle$; if S_1 (resp. S_2) is a belief set obtained from $P \cup Q_1$ (resp. $P \cup Q_2$), then S_1 is *preferred* to S_2 (written $S_2 \ll S_1$) if Q_1 is preferred to Q_2. A belief set S of $\langle P, \mathcal{A} \rangle$ is *most preferred* if there is no belief set S' of $\langle P, \mathcal{A} \rangle$ such that $S \ll S'$.

Example 1 Consider the abductive program $\langle P, \mathcal{A} \rangle$ where $P = \{ \leftarrow not\, p, not\, q, \quad r \leftarrow, \quad n_2 < n_1 \}$, $\mathcal{A} = \{ n_1 : p \leftarrow r, \quad n_2 : q \leftarrow not\, p \}$. Then, $\langle P, \mathcal{A} \rangle$ has two belief sets $S_1 = \{p, r, n_2 < n_1\}$ and $S_2 = \{q, r, n_2 < n_1\}$, obtained by adding $\{n_1\}$ and $\{n_2\}$ to P, respectively. S_1 is preferred to S_2 and S_1 is the most preferred belief set of $\langle P, \mathcal{A} \rangle$.

2.2 Abductive Programs with Disinformation

Dishonest agents are those who use intentionally false or inaccurate information. In this paper, we consider the following two cases. First, an agent a, who believes a proposition $\neg p$, informs another agent b that p is true. Second, an agent a, who believes neither p nor $\neg p$, informs another agent b that p (or $\neg p$) is true. The first one is called a *lie* [Mahon, 2008], while the second one is called *bullshit* (shortly, *BS*) [Franfurt, 2005]. In both cases, information p brought to another agent b is false or inaccurate (in contrast to the reality as believed by the agent a). We call such p *disinformation*. In abductive programs, disinformation is defined as follows.

Definition 1 (Disinformation) Let $\langle P, \mathcal{A} \rangle$ be an abductive program, and L and B be two sets of literals s.t.

- $\forall l \in L$, $\neg l$ belongs to every belief set of $\langle P, \mathcal{A} \rangle$;
- $\forall l \in B$, neither l nor $\neg l$ belongs to any belief set of $\langle P, \mathcal{A} \rangle$.

Then, $\mathcal{D} = (L, B)$ is called *disinformation* wrt $\langle P, \mathcal{A} \rangle$.

Literals in L represent lies, as their opposite facts are included in every belief set of $\langle P, \mathcal{A} \rangle$. Literals in B represent BS, as none of them (or their negations) are present in any belief set. By the definition, $L \cap B = \emptyset$. We next introduce a framework for realizing dishonest reasoning using abductive programs.

Definition 2 (ALD-program) Let $\langle P, \mathcal{A} \rangle$ be an abductive program and $\mathcal{D} = (L, B)$ disinformation wrt $\langle P, \mathcal{A} \rangle$. Let

$$\begin{aligned} I &= \{ r \mid r \in P \text{ and } head(r) \cap L^{\neg} \neq \emptyset \}, \\ \Phi &= \{ n_i < n_j \mid n_i \in \mathcal{A} \text{ and } n_j \in I \} \\ &\quad \cup \{ n_h < n_k \mid n_h \in (L \cup B) \text{ and } n_k \in \mathcal{A} \cup I \} \\ &\quad \cup \{ n_s < n_t \mid n_s \in L \text{ and } n_t \in B \} \end{aligned}$$

where $L^{\neg} = \{ \neg \ell \mid \ell \in L \}$.[4] An *abductive program with disinformation (ALD-program)* (wrt $\langle P, \mathcal{A} \rangle$ and $\mathcal{D}$) is defined as $\langle P_{\mathcal{D}}, \mathcal{A}_{\mathcal{D}} \rangle = \langle (P \setminus I) \cup \Phi, \mathcal{A} \cup I \cup L \cup B \rangle$.

Intuitively, I is a set of rules whose heads contain a literal that conflicts with lies in L. The set Φ specifies preferences over abducibles in three ways: **(i)** every rule in I is preferred to abducibles in $\mathcal{A}$, **(ii)** every rule in $\mathcal{A} \cup I$ is preferred to disinformation $L \cup B$, and **(iii)** BS is preferred to lies. **(i)** represents that every rule from the background knowledge P is preferred to hypotheses in $\mathcal{A}$. **(ii)** stands for ethics of rational agents; agents try to be honest as much as possible. Comparing lies and BS, lies are considered more sinful than BS, since lies are wrong beliefs while BS are ungrounded beliefs. **(iii)** represents this preference: an agent tries to keep lies as small as possible. In $P_{\mathcal{D}}$, I is removed from P because a program could become inconsistent when L is introduced to P, and the preference relations Φ are added to P. In $\mathcal{A}_{\mathcal{D}}$, on the other hand, new abducibles are set as $\mathcal{A}$ plus I and disinformation $L \cup B$. Observe that an ALD-program $\langle P_{\mathcal{D}}, \mathcal{A}_{\mathcal{D}} \rangle$ with $\mathcal{D} = (\emptyset, \emptyset)$ reduces to the original abductive program $\langle P, \mathcal{A} \rangle$. Because $\langle P_{\mathcal{D}}, \mathcal{A}_{\mathcal{D}} \rangle$ is also an abductive program, its belief sets and their preference relations are defined as before.

Proposition 1 *Let $\langle P, \mathcal{A} \rangle$ be an abductive program and $\mathcal{D}$ disinformation wrt $\langle P, \mathcal{A} \rangle$. If S is a (most preferred) belief set of $\langle P, \mathcal{A} \rangle$, then $S \cup \Phi$ is a (most preferred) belief set of the ALD-program $\langle P_{\mathcal{D}}, \mathcal{A}_{\mathcal{D}} \rangle$.*

If $\langle P, \mathcal{A} \rangle$ is consistent, then $\langle P_{\mathcal{D}}, \mathcal{A}_{\mathcal{D}} \rangle$ is also consistent. On the other hand, an inconsistent abductive program could recover consistency using disinformation.

Example 2 Consider the abductive program $\langle P, \mathcal{A} \rangle$ where $P = \{ \leftarrow not\, q, \quad q \leftarrow p, r, \quad \neg p \leftarrow \}$, $\mathcal{A} = \{r\}$, and disinformation $\mathcal{D} = (\{p\}, \{q\})$ wrt $\langle P, \mathcal{A} \rangle$. The ALD-program $\langle P_{\mathcal{D}}, \mathcal{A}_{\mathcal{D}} \rangle$ then becomes $P_{\mathcal{D}} = \{ \leftarrow not\, q, \quad q \leftarrow p, r, \quad n_2 < n_1, \quad n_3 < n_1, \quad n_4 < n_1, \quad n_3 < n_2, \quad n_4 < n_2, \quad n_3 < n_4 \}$ and $\mathcal{A}_{\mathcal{D}} = \{ n_1 : \neg p, \quad n_2 : r, \quad n_3 : p, \quad n_4 : q \}$. Here,

[1] A rule with variables is viewed as the set of its ground instances.

[2] For space limitations, we omit the definition of answer sets.

[3] We omit the rule names when not needed in the discussion.

[4] We assume $\neg\neg a = a$ to represent the atom a.

$\langle P, \mathcal{A}\rangle$ is inconsistent, while $\langle P_{\mathcal{D}}, \mathcal{A}_{\mathcal{D}}\rangle$ has the most preferred belief set $\{q\} \cup \Phi$. Here, bullshit q is used in creating the belief set. Note that $\{p, q, r\} \cup \Phi$ is also a belief set but it is not a preferred one as it uses a lie p.

3 Knowledge Bases and Proposals

3.1 Negotiation Knowledge Bases

A knowledge base for negotiation is expected to serve as a means for an agent to create his/her proposals/responses in negotiation, and to decide whether he/she should accept/reject a proposal. To this end, it must encode an agent's beliefs, rules for negotiation with their preferences, possible assumptions about the other agent, and possible information that he/she could behave dishonestly. To model this, we introduce negotiation knowledge bases using ALD-programs.

Definition 3 (Negotiation KB) Given an abductive program $\langle P, \mathcal{A}\rangle$ and disinformation $\mathcal{D} = (L, B)$ wrt $\langle P, \mathcal{A}\rangle$, a *negotiation knowledge base* (*NKB*) is defined as a tuple $K = (\Pi, H, N^{\prec})$ such that:

- $\Pi = \langle P_{\mathcal{D}}, \mathcal{A}_{\mathcal{D}}\rangle$ is an ALD-program wrt $\langle P, \mathcal{A}\rangle$ and $\mathcal{D}$.
- H is a set of literals (called *assumptions*) s.t. $H \subseteq \mathcal{A}$.
- $N^{\prec}$ is a set of literals (called *negotiation conditions*) associated with a strict partial order $\prec$ on its elements.

An NKB K is *consistent* if Π is consistent.

Π represents an agent's domain-specific knowledge, goals and preferences with disinformation. H represents assumptions about the other agent that is unknown at the beginning of a negotiation. $N^{\prec}$ specifies desired outcomes in a negotiation. $p \prec q$ means that q is preferred to p.[5] For simplicity, we often write $N^{\prec} = \{p \prec q \prec r\}$ if $p \prec q$ and $q \prec r$ hold over the set $N^{\prec} = \{p, q, r\}$. Since an ALD-program serves as a means for an agent to generate arguments in negotiation, Π is assumed to be consistent.

We now present two NKBs, one for a seller and one for a buyer, describing the agents addressed in the introduction.

Example 3 Suppose a seller agent s, who has the abductive program $\langle P_s, \mathcal{A}_s\rangle$ and disinformation $\mathcal{D}_s$ such that

$$P_s = \{\, sale \leftarrow prod_A, \mathbf{price_1},\;\; sale \leftarrow prod_B, \mathbf{price_2},\;\; \leftarrow not\; sale,\;\; mak_C \leftarrow prod_A,\;\; mak_D \leftarrow prod_B,\;\; \neg qual_B \leftarrow prod_B,\;\; bargain \leftarrow prod_B,\;\; prod_A \leftarrow,\;\; prod_B \leftarrow,\;\; n_i < n_1,\;\; n_4 < n_j,\;\; \leftarrow high, low,\;\; \leftarrow high, lowest,\;\; \leftarrow low, lowest \,\}$$

where $\mathbf{price_1} \in \{high, low\}$, $\mathbf{price_2} \in \{low, lowest, high\}$, $i \in \{2, 3, 4\}$, and $j \in \{2, 3\}$.

$$\mathcal{A}_s = \{\, n_1 : high,\;\; n_2 : low \leftarrow student,\;\; n_3 : low \leftarrow bargain, cash,\;\; n_4 : lowest \leftarrow mailing, cash,\;\; n_5 : student,\;\; n_6 : cash,\;\; n_7 : mailing \,\}.$$

$$\mathcal{D}_s = (L_s, B_s) = (\{qual_B\}, \{qual_A\}).$$

Here, $prod$, $qual$ and mak mean product, quality, and maker, respectively. P_s states features about products and sales conditions, together with preferences among abducibles. $\mathcal{A}_s$ specifies different pricing scenarios. (n_1) Any customer buying a product with $high$ is accepted. (n_2) Students are entitled to low. (n_3) low is also applied to bargain products and for every customer paying in cash. (n_4) A special discount $lowest$ is applied to a customer who subscribes to the shop's mailing list and purchases the product in cash. The seller intends to claim that both products A and B are of good quality, if needed. This is represented by $\mathcal{D}_s$ in which the seller would lie about $qual_B$ and bullshit about $qual_A$.

The seller agent then constructs his/her NKB $K_s = (\Pi_s, H_s, N_s^{\prec})$ as

- $\Pi_s = \langle P^s_{\mathcal{D}}, \mathcal{A}^s_{\mathcal{D}}\rangle$ with $P^s_{\mathcal{D}} = (P_s \setminus \{n_8\}) \cup \{n_i < n_8 \mid i = 1, \ldots, 7\} \cup \{n_h < n_k \mid k = 1, \ldots, 8,\; h = 9, 10\} \cup \{n_9 < n_{10}\}$ where $n_8 : \neg qual_B \leftarrow prod_B$, $n_9 : qual_B$, $n_{10} : qual_A$ and $\mathcal{A}^s_{\mathcal{D}} = \mathcal{A}_s \cup \{n_8\} \cup L_s \cup B_s$.
- $H_s = \{\, student,\; cash,\; mailing \,\}$.
- $N_s^{\prec} = \{\, lowest \prec low \prec high \,\}$.

$N_s^{\prec}$ indicates that $high$ is preferred to low and $lowest$. It is easy to see that Π_s is consistent so K_s is consistent.

Example 4 Suppose a buyer agent b, who has the abductive program $\langle P_b, \mathcal{A}_b\rangle$ and disinformation $\mathcal{D}_b$ such that

$$P_b = \{\, purchase \leftarrow prod_X, qual_X, \mathbf{price_3},\;\; \leftarrow high,\;\; \leftarrow not\, purchase,\;\; \neg student \leftarrow,\;\; cash \leftarrow,\;\; \leftarrow low, lowest,\;\; n_2 < n_1,\;\; n_3 < n_2,\;\; n_3 < n_1 \,\}$$

where $X \in \{A, B\}$ and $\mathbf{price_3} \in \{\, low,\; lowest \,\}$.

$$\mathcal{A}_b = \{\, n_1 : lowest \leftarrow mak_C,\;\; n_2 : lowest \leftarrow mak_D,\;\; n_3 : low \leftarrow mak_C,\;\; n_4 : qual_A,\;\; n_5 : qual_B,\;\; n_6 : mak_C,\;\; n_7 : mak_D,\;\; n_8 : prod_A,\;\; n_9 : prod_B \,\}.$$

$$\mathcal{D}_b = (L_b, B_b) = (\emptyset, \{mailing\}).$$

The buyer does not care about the mailing list of the seller but could pretend to join it if it works to his/her advantage. Using $\langle P_b, \mathcal{A}_b\rangle$ and $\mathcal{D}_b$, the buyer agent constructs his/her NKB $K_b = (\Pi_b, H_b, N_b^{\prec})$ as

- $\Pi_b = \langle P^b_{\mathcal{D}}, \mathcal{A}^b_{\mathcal{D}}\rangle$ with $P^b_{\mathcal{D}} = P_b \cup \{n_{10} < n_k \mid k = 1, \ldots, 9\}$ where $n_{10} : mailing$ and $\mathcal{A}^b_{\mathcal{D}} = \mathcal{A}_b \cup L_b \cup B_b$.
- $H_b = \{\, n_4, n_5, n_6, n_7, n_8, n_9 \,\}$.
- $N_b^{\prec} = \{\, low \prec lowest \,\}$.

It is easy to check that K_b is also consistent.

3.2 Proposals and Acceptability

In building a proposal, an agent has a goal and can make assumptions about the receiver of the proposal. The agent may also decide to reveal information about his/her state-of-belief, rendering some conditions on the feasibility of the proposal. Given a set S of literals, let $Goal(S) = \{\leftarrow not\; \ell \mid \ell \in S\}$.

Definition 4 (Proposal) Let $K = (\Pi, H, N^{\prec})$ be an NKB of an agent with $\Pi = \langle P_{\mathcal{D}}, \mathcal{A}_{\mathcal{D}}\rangle$. For a set of literals $G \subseteq N^{\prec}$, a tuple $\gamma = (G, S, R)$ is a *proposal* wrt K if $\langle P_{\mathcal{D}} \cup Goal(G), \mathcal{A}_{\mathcal{D}}\rangle$ has a belief set M such that $S = M \cap H$ and $R \subseteq M \setminus H$. We refer to G, S, R, and M as the *goal*, *assumptions*, *conditions*, and *support* of γ, respectively. The proposal is *honest* if $M \cap (L \cup B) = \emptyset$; it is *deceptive* if $M \cap L \neq \emptyset$; and it is *unreliable* if $M \cap B \neq \emptyset$.

Intuitively, a proposal $\gamma = (G, S, R)$ states that the goal of an agent is to negotiate for the objective G. The reason to put forward γ is that the agent has a support M for it. The agent indicates assumptions S that he/she has made about the receiver of γ. In addition, the agent also reveals additional information R supporting the goal G, which informs the receiver that the information in R should not be violated.

[5] We distinguish $\prec$ from $<$ that is defined over abducibles.

Example 5 In Example 3, $(\{high\}, \emptyset, \{prod_A\})$ and $(\{low\}, \{student\}, \{prod_B\})$ are two honest proposals by the seller. The former states that he/she can sell $prod_A$ for the price $high$. The latter states that he/she can sell $prod_B$ for the price low if the customer is a student. A deceptive proposal by the seller is $(\{low\}, \{student\}, \{prod_B, qual_B\})$, indicating that he/she can sell $prod_B$ having good quality for low provided that the buyer is a student. The proposal is deceptive because $qual_B$ is a lie. Similarly, $(\{high\}, \emptyset, \{prod_A, qual_A\})$ is an unreliable proposal wrt K_s.

Next we consider the acceptability of proposals.

Definition 5 (Acceptability) Let $K_a = (\Pi, H, N^{\prec})$ be an NKB of an agent a with $\Pi = \langle P_{\mathcal{D}}, \mathcal{A}_{\mathcal{D}} \rangle$, and $\gamma_b = (G, S, R)$ a proposal by another agent b. Then,

- γ_b is *acceptable* wrt K_a if $Q = \langle P_{\mathcal{D}} \cup Goal(G), \mathcal{A}_{\mathcal{D}} \rangle$ has a belief set M such that $S \subseteq M$ and $M \cap H \subseteq R$. We say that γ_b is acceptable *without disinformation* if $M \cap (L \cup B) = \emptyset$; γ_b is acceptable *with disinformation*, otherwise.
- γ_b is *rejectable* wrt K_a if Q is inconsistent.
- γ_b is *negotiable* wrt K_a, otherwise.

Intuitively, Q encodes the possibility of satisfying b's goal G in the ALD-program Π of the agent a. If Q is inconsistent, there is no way to accept γ_b. γ_b is acceptable if Q has a belief set M such that: **(i)** M is compatible with the assumptions S about the agent a ($S \subseteq M$); and **(ii)** if there are assumptions $M \cap H$ made by the agent a about the proposer b, then these must be compatible with the information R revealed by the proposer ($M \cap H \subseteq R$). The first condition is needed, since a negotiated goal is acceptable to both parties only if their supports agree. The second condition implies that a proposal is based on the same set of shared assumptions. A proposal is negotiable if it is neither acceptable nor rejectable. Note that when an agent considers a proposal acceptable or negotiable, he/she may use disinformation included in his/her NKB.

Example 6 For K_s and K_b from Examples 3 and 4,

- $(\{high\}, \{prod_A, qual_A\}, \emptyset)$ is acceptable with disinformation wrt K_s, since $\langle P^s_{\mathcal{D}} \cup Goal(\{high\}), \mathcal{A}^s_{\mathcal{D}} \rangle$ has a belief set M containing $high, prod_A$, and $qual_A$. Any belief set that allows the seller to accept this proposal contains disinformation $qual_A$.
- $(\{low\}, \{prod_B, mak_D, qual_B\}, \emptyset)$ is a negotiable proposal wrt K_s, since $\langle P^s_{\mathcal{D}} \cup Goal(\{low\}), \mathcal{A}^s_{\mathcal{D}} \rangle$ has a belief set containing its assumptions but requires at least one of the sets $\{student\}$ or $\{cash\}$.
- $(\{high\}, \emptyset, \{prod_A, mak_C, qual_A\})$ is a rejectable proposal wrt K_b because $\langle P^b_{\mathcal{D}} \cup Goal(\{high\}), \mathcal{A}^b_{\mathcal{D}} \rangle$ has no belief set containing $high$.

Let Γ^A_K, Γ^N_K, and Γ^R_K be the sets of proposals that are acceptable, negotiable, and rejectable wrt an NKB K.

Proposition 2 *Γ^A_K, Γ^N_K, and Γ^R_K are pairwise disjoint. Furthermore, $\gamma \in \Gamma^A_K \cup \Gamma^N_K \cup \Gamma^R_K$ for any proposal γ.*

4 Negotiation Using NKBs

4.1 Negotiation

We will now present a model of negotiation between two agents a and b who respectively use NKBs K_a and K_b that share the same language. Each agent does not assume anything related to his/her condition, so the set H of assumptions in K_a is assumed to be disjoint from the one in K_b. In negotiation, an agent a puts forward a proposal $\gamma_a = (G, S, R)$, and the opponent b will respond with a proposal with the same structure. At the moment, a response could be an arbitrary proposal, an acceptance or a rejection of the current proposal. (We will address more sophisticated responses in Def. 13.)

Definition 6 (Response) Let $K_a = (\Pi, H, N^{\prec})$ be an NKB of an agent a, and $\gamma_b = (G, S, R)$ a proposal by b wrt its NKB K_b. A *response* to γ_b by a is either **(i)** a proposal $\gamma_a = (G', S', R')$; or **(ii)** $(\top, \emptyset, \emptyset)$, denoting *acceptance* of the proposal if γ_b is acceptable wrt K_a; or **(iii)** $(\bot, \emptyset, \emptyset)$, denoting *rejection* of the proposal. The set of all responses (by a wrt K_a) to a proposal γ_b is denoted by $Res(K_a, \gamma_b)$.

A negotiation is a series of responses between two agents, who, in alternation, take into consideration the other agent's response and put forward a new response; this can be either accept, reject, or a new proposal that may involve explanations of why the latest proposal (of the other agent) was not acceptable. A possibly infinite sequence of responses $\omega_1, \ldots, \omega_i, \ldots$ is denoted by $\langle \omega_i \rangle_{i>0}$.

Definition 7 (Negotiation) A *negotiation* between two agents a and b, starting with a, is a possibly infinite sequence of responses $\langle \omega_i \rangle_{i>0}$ where $\omega_i = (G_i, S_i, R_i)$ and **(i)** ω_{2k+1} is a proposal wrt K_a ($k \geq 0$); **(ii)** ω_{2k} is a proposal wrt K_b ($k \geq 1$); **(iii)** ω_{i+1} is a response to ω_i for every $i \geq 1$. A negotiation *ends* at i if $\omega_i = (\top, \emptyset, \emptyset)$ or $\omega_i = (\bot, \emptyset, \emptyset)$. When $G_i \neq G_{i+2}$, we say that a *goal change* has occurred for the agent who responded ω_i.

Definition 8 (Un/Successful Negotiation) A negotiation is *successful* (resp. *unsuccessful*) if it is finite and ends with $\omega_i = (\top, \emptyset, \emptyset)$ (resp. $\omega_i = (\bot, \emptyset, \emptyset)$). We call ω_{i-1} the *accepted* (resp. *rejected*) proposal of the negotiation.

Example 7 The negotiation dialogue between the seller s and the buyer b in the introduction is realized using NKBs of Examples 3 and 4 as follows.

b_1 : $(\{low\}, \{prod_A, qual_A, mak_C\}, \emptyset)$
s_1 : $(\{low\}, \{student\}, \{prod_A, qual_A, mak_C\})$
b_2 : $(\{low\}, \{prod_A, qual_A, mak_C\}, \{\neg student\})$
s_2 : $(\{low\}, \{cash\}, \{prod_B, mak_D, qual_B\})$
b_3 : $(\{lowest\}, \{prod_B, mak_D, qual_B\}, \{cash\})$
s_3 : $(\{lowest\}, \{cash, mailing\}, \{prod_B, mak_D, qual_B\})$
b_4 : $(\top, \emptyset, \emptyset)$.

The seller bullshits in s_1 and lies in s_2. The buyer bullshits in b_4. A goal change has occurred at b_3 (for the buyer) and s_3 (for the seller).

A negotiation represents one possible way for two agents to reach an agreement (or disagreement). In the course of reaching an agreement, two agents might have different alternatives. All possible negotiations between two agents are represented by a *negotiation tree*. In what follows, the *level* of a node in a tree means the number of links lying on the path connecting the root to this node.

Definition 9 (Negotiation Tree) A *negotiation tree* between two agents a and b, starting with a, is a labeled tree $T_{a,b}$

where **(i)** each child of the root node has the label of the form (K_a, γ_a, K_b) where γ_a is a proposal wrt K_a; **(ii)** if $\eta_i = (K, \omega, K')$ is a node at the level $i \geq 1$, then every child of η_i has the label of the form (K', ω', K) where $\omega' \in Res(K', \omega)$; and **(iii)** nodes labeled by $(K, (\top, \emptyset, \emptyset), K')$ or $(K, (\bot, \emptyset, \emptyset), K')$ have no children.

Each path from a node η_1 of $T_{a,b}$ to a leaf is a negotiation between a and b. There is a possibility of reaching an agreement if $T_{a,b}$ has a finite length of path having the leaf labeled by $(K, (\top, \emptyset, \emptyset), K')$; otherwise, no agreement is reached.

4.2 Negotiation Strategies

The previous section provides basic definitions for modeling negotiation. In practice, agents commonly employ their own strategies in a negotiation. We next formalize this notion.

Definition 10 (Strategy) Given an agent a with the NKB K_a, a *negotiation strategy* for a is a function F that maps a proposal γ_b by another agent b and a negotiation $\langle \omega_i \rangle$, a finite sequence $\omega_1, \ldots, \omega_i$, to a new proposal that satisfies **(i)** $F(\gamma_b, \langle \omega_i \rangle) \in Res(K_a, \gamma_b)$ and **(ii)** the sequence $\langle \omega_i \rangle, F(\gamma_b, \langle \omega_i \rangle)$ is a negotiation.

Agents are interested in different types of strategies. For example, strategies that guarantee the termination of a negotiation, strategies that do not use disinformation, strategies that guarantee the success of a negotiation, etc. We will discuss some of these next. For two negotiations $\langle \omega_i \rangle$ and $\langle \omega_j \rangle$, we write $\langle \omega_i \rangle \lhd \langle \omega_j \rangle$ if $\langle \omega_i \rangle$ is a proper prefix of $\langle \omega_j \rangle$.

Definition 11 (Observant Strategy) A strategy F is *observant* if $F(\gamma, \langle \omega_i \rangle) \neq F(\gamma, \langle \omega_j \rangle)$ for every pair of negotiations $\langle \omega_i \rangle$ and $\langle \omega_j \rangle$ such that $\langle \omega_i \rangle \lhd \langle \omega_j \rangle$.

The observant strategy says an agent does not repeat the same response to the same proposal in a negotiation. If at least one agent's strategy is observant, then the negotiation will terminate. This is because an agent does not have infinitely many responses and thus the negotiation will either fail/succeed.

Proposition 3 *Let a and b be two agents with strategies F_a and F_b, respectively. If either F_a or F_b is observant, then every negotiation between a and b terminates.*

In practice, an agent might prefer to be honest before he/she uses disinformation in achieving his/her goals. We address some strategies that differentiate between honesty and dishonesty. In particular, strategies, which guarantee that an agent lies or BS only if he/she has no alternative, can be built using the preference relation between belief sets, since such relation favors belief sets without disinformation.

Definition 12 (Deliberate Strategy) Given a proposal γ_b, let $\Sigma_a(\gamma_b) = \{ M \mid$ a belief set M (of an agent a) supports some $\omega \in Res(K_a, \gamma_b)\}$. Then, a strategy F is *deliberate* if **(i)** $F(\gamma_b, \langle \omega_i \rangle)$ is supported by a most preferred belief set in $\Sigma_a(\gamma_b)$ whenever $\Sigma_a(\gamma_b) \neq \emptyset$; **(ii)** $F(\gamma_b, \langle \omega_i \rangle) = (\bot, \emptyset, \emptyset)$, otherwise. A deliberate strategy, which is also observant, is called a *best-practice* strategy.

A deliberate strategy does not guarantee termination of a negotiation. However, a best-practice strategy does. Furthermore, we can observe that an agent with a best-practice strategy may accept a less preferred outcome (under $\prec$) of a negotiation even though he/she might obtain a more preferred outcome had he/she used disinformation. Similarly, he/she may sometimes reject a proposal even though this might be negotiable and further negotiation might yield a preferred outcome, had he/she lied or bullshitted. This can be seen in Example 7: a deliberate seller will respond to b_1 with $s'_1 = (\{low\}, \{student\}, \{prod_A, mak_C\})$ rather than s_1, since he/she has a belief set without disinformation ($qual_A$) that supports s'_1. The negotiation does not reach an agreement by b_4 if s_3 does not include a lie $qual_B$.

One disadvantage of the observant strategy is that it requires an agent to memorize the full history of the negotiation. We next consider a possible way to avoid this. In negotiation, an agent, who would be willing to accept rather than reach no agreement at all, needs to consider **(i)** the assumptions that have been made by the opponent, and **(ii)** the information that the opponent reveals about him/herself. As for **(i)**, the response should identify and make explicit those assumptions that are wrong about him/herself, as far as he/she will not lie/BS on those assumptions—e.g., if the seller assumes that the buyer is a student but the buyer is not, then the buyer should identify this and inform the seller, as far as if he/she does not disguise him/herself as a student. As for **(ii)**, the response needs to conform to this information—e.g., if the seller says that he/she does not have the product A, then the the buyer should not assume that $prod_A$ is available, even though $prod_A$ is a viable assumption in his/her KB. These considerations lead to more sophisticated responses.

Definition 13 (Conscious Response) Let $K_a = (\Pi, H, N^{\prec})$ be an NKB of an agent a, and $\gamma_b = (G, S, R)$ a proposal by another agent b. A *conscious response* to γ_b by a is

(i) $\gamma_a = (G', S', R')$ wrt K_a with a support M such that $G \preceq G'$, $R \cap H \subseteq S'$, and $S^{\neg} \cap M \subseteq R'$, where $S^{\neg} = \{\neg \ell \mid \ell \in S\}$, if γ_b is not rejectable wrt K_a; **or**

(ii) $\gamma_a = (G', S', R')$ wrt K_a with a support M such that $G \not\preceq G'$ and $S^{\neg} \cap M \subseteq R'$, if γ_b is rejectable wrt K_a; **or**

(iii) $(\top, \emptyset, \emptyset)$, denoting *acceptance of the proposal*, if γ_b is acceptable wrt K_a; **or**

(iv) $(\bot, \emptyset, \emptyset)$, denoting *rejection of the proposal*.

If the proposal γ_b is acceptable to an agent a, then he/she could accept it (case (**iii**)) or attempt to negotiate for some better options (case (**i**)). If γ_b is negotiable, he/she could continue and attempt to get a better option (case (**i**)). If γ_b is rejectable, he/she could try to negotiate for something that is not as good as the current goal (case (**ii**)). In each case, the agent can stop with rejection (case (**iv**)). An agent a should generate a new proposal γ_a whose goal G' depends on the goal G of the given proposal γ_b, whose assumptions S' cover the conditions R stated in γ_b ($R \cap H \subseteq S'$), and whose conditions R' identify all incorrect assumptions made in γ_b ($S^{\neg} \cap M \subseteq R'$). An agent, who considers preferable proposals, would require that the support for the new proposal must be preferred to any support for accepting γ_b.

Example 8 The proposal $\gamma_s = (\{low\}, \{cash\}, \{prod_A, qual_A, mak_C\})$ by the seller ("I can sell you the $prod_A$, made by mak_C, and has good quality for low if you pay in cash") is acceptable by the buyer wrt K_b. However, the buyer could respond with the proposal $\gamma_b = (\{lowest\}, \{prod_A, qual_A, mak_C\}, \{cash\})$ ("Can I get the $lowest$?").

Intelligent agents also update their negotiation KBs using incoming information during negotiation.

Definition 14 (Adaptive Agent) An agent a with the NKB $K_a = (\Pi, H, N^{\prec})$ and a strategy F is said to be *adaptive* if for every proposal $\gamma_b = (G, S, R)$ by another agent b,

- $F(\gamma_b, \langle \omega_i \rangle) = (G', S', R')$ is a conscious response to γ_b; and if γ_b is acceptable wrt K_a, then $G \prec G'$ or $(G', S', R') = (\top, \emptyset, \emptyset)$.
- a changes his/her NKB to $K'_a = (\Pi', H, N^{\prec})$ where $\Pi' = \langle P_{\mathcal{D}} \cup (R \cap (H \cup H^{\neg})), \mathcal{A}_{\mathcal{D}} \rangle$ after his/her response to γ_b.

Intuitively, an agent is adaptive if it imports information received during a negotiation into his/her NKB and keeps this information for the next round of the negotiation. Furthermore, an adaptive agent prefers to accept a proposal if a better outcome cannot be achieved. It is easy to see that if both agents are adaptive then a negotiation will terminate.

5 Discussion

Logic programming has been used for formulating negotiation by many researchers. Chen *et al.* [2007] use answer sets of logic programs as a means for negotiation between agents. Their goal is to coordinate answer sets of two programs, and it has no mechanism for developing proposals for a particular goal. Sadri *et al.* [2002] realize negotiation using abductive programs. In their framework, a program specifies a negotiation plan to achieve a goal, and the behavior of agents is operationally specified by an observe-think-act cycle. In our framework, an agent can generate proposals using abductive assumptions, and the behavior of agents is flexibly changed by strategies. Son *et al.* [2009] use logic programs with consistency restoring rules (CR-Prolog) to formulate negotiation. Unlike abductive programs, a CR-Prolog program considers the most preferred answer sets only. This restricts options for building proposals that are supported by less preferred belief sets (with disinformation). Any proposal/response produced by CR-Prolog is also built in our current framework, but the converse implication does not hold in general. Our work is in the same spirit as the approaches to argumentation-based negotiation (ABN) [Kakas and Moraitis, 2006; Amgoud *et al.*, 2007], in that it considers explanations as a part of a proposal/response. The main difference between our work and ABN lies in our use of abductive programs, a nonmonotonic logic, while ABN's logic is monotonic. Our framework does not compute explanations for accepting/rejecting a proposal in advance as in [Amgoud *et al.*, 2007], and it allows negotiators to non-monotonically modify their beliefs and change proposals using incoming information. Kakas *et al.* [2006] introduce priorities over arguments and use abduction to seek conditions to support arguments. They do not integrate abduction and preferences as done in this paper.

It is important to note that all studies mentioned above model negotiation between *honest* agents, which is not always realistic. There are few studies which provide a formal logic for negotiation between dishonest agents. Zlotkin *et al.* [1991] study negotiation in which agents may lie. The study focuses on multi-agents in a dynamic environment, where agents act and interact to achieve their individual or cooperative goals. However, it does not provide any computational method of building (dis)honest proposals. The issue has also been discussed in the context of game theory (e.g., [Ettinger and Jehiel, 2010]). The approach differs from most of the works we have discussed so far, including our own.

In this paper, we develop an abstract framework for negotiation based on logic programming. It can deal with incomplete information, preferences, disinformation, and compute proposals and conditions on a case-by-case basis using different negotiation strategies. Proposals and responses are computed by belief sets of abductive programs, so computational complexities follow from those of abductive programs. The proposed framework has been implemented on top of the ASP-Prolog platform [Nguyen *et al.*, 2011]. On the other hand, this paper does not argue the issue of detecting and dealing with disinformation made by a dishonest agent. The investigation of these issues is left for future work.

References

[Amgoud *et al.*, 2007] L. Amgoud, Y. Dimopoulos, and P. Moraitis. A unified and general framework for argumentation-based negotiation. In *Proc. 6th AAMAS*, 1018–1025, 2007.

[Carson, 2010] T. L. Carson. *Lying and deception: theory and practice*. Oxford University Press, 2010.

[Chen *et al.*, 2007] W. Chen, M. Zhang, and N. Foo. Repeated negotiation of logic programs. In *Proc. 7th NRAC*, pp. 33–39, 2007.

[Ettinger and Jehiel, 2010] D. Ettinger and P. Jehiel. A theory of deception. *American Economic Journal: Macroeconomics*, 2(1):1–20, 2010.

[Franfurt, 2005] H. G. Frankfurt. *On Bullshit*. Princeton University Press, 2005.

[Gelfond and Lifschitz, 1991] M. Gelfond and V. Lifschitz. Classical negation in logic programs and disjunctive databases. *New Generation Computing*, 9:365–385, 1991.

[Kakas *et al.*, 1998] A.C.Kakas, R.A.Kowalski, and F.Toni. The role of abduction in logic programming. In: *Handbook of Logic in AI and LP*, vol. 5, pp. 235–324, 1998.

[Kakas and Moraitis, 2006] A. Kakas and P. Moraitis. Adaptive agent negotiation via argumentation. In *Proc. 5th AAMAS*, pp. 384–391, 2006.

[Mahon, 2008] J. E. Mahon. Two definitions of lying. *Journal of Applied Philosophy*, 22(2):211–230, 2008.

[Nguyen *et al.*, 2011] N-H. Nguyen, T. C. Son, E. Pontelli and C. Sakama. ASP-Prolog for negotiation among dishonest agents. In *Proc. LPNMR*, LNAI 6645, 2011.

[Sadri *et al.*, 2002] F. Sadri, F. Toni, and P. Torroni. An abductive logic programming architecture for negotiating agents. In *Proc. JELIA*, LNAI 2424, pp. 419–431, 2002.

[Sakama and Inoue, 2000] C. Sakama and K. Inoue. Prioritized logic programming and its application to commonsense reasoning. *Artif. Intell.*, 123(1-2):185–222, 2000.

[Son and Sakama, 2009] T. C. Son and C. Sakama. Negotiation using logic programming with consistency restoring rules. In: *Proc. IJCAI-09*, pp. 930–935, 2009.

[Zlotkin and Rosenschein, 1991] G. Zlotkin and J. S. Rosenschein. Incomplete information and deception in multiagent negotiation. In: *Proc. IJCAI-91*, pp. 225–231, 1991.