On the Generation of Bipolar Goals in Argumentation-Based Negotiation

Leila Amgoud[1] and Souhila Kaci[2]

[1]Institut de Recherche en Informatique de Toulouse (I.R.I.T.),
Université Paul Sabatier,
118, route de Narbonne 31062 Toulouse, France
amgoud@irit.fr

[2]Centre de Recherche en Informatique de Lens (C.R.I.L.),
IUT de Lens,
Rue de l'Université SP 16 62307 Lens, France
kaci@cril.univ-artois.fr

Abstract. The notion of agent's goals is crucial in negotiation dialogues. In fact, during a negotiation, each agent tries to make and to accept the offers which satisfy its own *goals*. Works on negotiation suppose that an agent has a set of fixed goals to pursue. However, it is not shown how these goals are computed and chosen by the agent. Moreover, these works handle one kind of goals: the ones that an agent wants to achieve.

Recent studies on psychology claim that goals are bipolar and there are at least two kinds of goals: the *positive goals* representing what the agent wants to achieve and the *negative goals* representing what the agent rejects. In this paper, we present an argumentation-based framework which generates the goals of an agent. The framework returns three categories of goals: the *positive* goals, the *negative* ones and finally the goals in *abeyance*.

Keywords: Negotiation, Argumentation.

1 Introduction

In most agent applications, the autonomous components need to interact with one another because of the inherent interdependencies which exist between them, and negotiation is the predominant mechanism for achieving this by means of an exchange of offers. The purpose of negotiation is to make a deal and each agent aims to maximize its profit. In fact, an agent makes and accepts only offers that satisfy its *goals*.

Works in multi-agents negotiation can be roughly divided into two categories. The first one has mainly focused on the numerical computation of trade-offs in terms of utilities, and the search for concessions which still preserve the possibility of reaching preferred states of affairs e.g.[10, 15]. These works suppose that each agent has a set of fixed goals that it should pursue.

Recently, a second line of research [2, 9, 13] has focused on the necessity of supporting offers by arguments during a negotiation. Indeed, an offer supported by a good argument has a better chance to be accepted by an agent and may lead an agent to *revise*

I. Rahwan et al. (Eds.): ArgMAS 2004, LNAI 3366, pp. 192–207, 2005.

its goals. However, in these works, it is not clear how the goals are handled and updated if necessary.

In sum, in all these approaches, it is not shown how the goals are computed and chosen by the agent and how they can be revised. Moreover, in all the above approaches, only one kind of goals is considered: the ones that an agent wants to achieve. However, in [3] the authors argued that when an agent expresses its goals, it usually does that in a *bipolar* way. On one hand, it expresses what it really wants, what it considers as really satisfactory. These are *positive goals*. They will represent the goals which will be pursued by the agent and each offer satisfying these goals is *rewarded*. On the other hand, it expresses what it definitely rejects, what it considers as unacceptable. These are *negative goals*. They represent the goals which will not be pursued by the agent. This category of goals is very important in a negotiation since each offer satisfying a negative goal will automatically be rejected by the agent. Indeed, reasoning on both what an agent likes and what it rejects, enriches the negotiation process since an offer can be evaluated w.r.t. both kinds of goals. For example an agent may consider an offer better than another if both falsify all its positive goal but the first one does not satisfy any negative goal (i.e., it is not rejected) while the second one satisfies at least one negative goal (i.e., it is rejected).

Beware that positive goals do not just mirror what is not rejected since a goal which is not rejected is not necessarily pursued. This category of goals which are neither negative nor positive are said to be *in abeyance*. Note however that positive and negative goals are related by a coherence condition which says that what is pursued should be among what is not rejected.

This distinction between positive and negative goals is supported by recent studies in cognitive psychology which have shown that these two types of goals are independent and processed separately in the mind [5, 6, 12, 4].

This paper focuses on the computation of goals. It particularly answers the following questions: what is a goal? what is its origin? what is its nature? what are the different kinds of arguments supporting it? and how it is computed or obtained?

We present an argumentation-based framework which returns the three categories of goals, namely *positive goals*, *negative goals* and *goals in abeyance*.

The paper is organized as follows: section 2 studies the nature of a goal. Section 3 introduces two different types of arguments supporting a goal: *explanatory arguments* and *instrumental arguments*. Section 4 presents an argumentation framework which evaluates the explanatory arguments and section 5 presents another framework which evaluates the instrumental arguments. Section 6 computes the positive and the negative goals of an agent. Section 7 shows through an example how the positive and negative goals of an agent may change in a negotiation dialogue.

2 The Nature of a Goal

In this section, we will discuss the nature of a goal according to three different criteria: *subjectivity* which has been already discussed in [11], *bipolarity* and finally the *origin*.

2.1 The Subjective Nature

As it has already been mentioned in [11] goals are considered as *motivational* attitudes of an agent and they are therefore by nature intrinsic to an agent. Indeed, one cannot say, as for beliefs, that a given goal of an agent is correct or incorrect. But, we can attempt to establish that a goal seems unachievable, not useful or unsupported. Let's take the following dialogues between two agents:

Example 1.
P: I would like to fly to Algiers with Algerian Airlines because it is not expensive.
C: But flying with Algeria Airlines means changing your flight arrangements is not flexible.
P: I know that.

In this case even if the argument given by **C** seems *acceptable* and in some sense defeats the argument supporting the goal *to fly with Algerian Airlines*, **P** maintains its goal.

Example 2.
P: I would like to fly to Algiers with Algerian Airlines because it is not expensive.
C: Actually flying with Algeria Airlines can be quite expensive because it is the holiday season.
P: I didn't know that.

In this case, if the agent **P** does not find another company which is cheaper than Algerian Airlines, then we can also imagine that the agent keeps its goal.

2.2 A Bipolar Nature

As we said in the introduction, an original representation of goals called bipolar goals [3] has been proposed. Indeed, we distinguish two independent types of goals: positive goals describing the goals pursued by the agent and negative goals describing the goals rejected, not pursued by the agent. The goals which are neither negative nor positive are called goals in abeyance. In example 1, we can imagine that the flexibility arrangements of a flight is not very important for the agent.

Goals are matter of degrees. Thus an agent expresses its goals by means of two different bases. A base representing what is more or less rejected by the agent and another base representing what is more or less satisfactory for him.

To illustrate the idea of bipolar goals, let us consider the following example (introduced in [8]) of an agent who goes to an agency in order to buy a house or an apartment. It gives the following positive and negative goals to the seller.

Example 3. The agent does not want a house or an apartment with a small surface and which is expensive. This negative goal is encoded as $(small \wedge expensive)$. Another negative goal is that it does not accept a house without a garden $(house \wedge \neg garden)$. However, the agent has some positive goals which are: an apartment with a large surface $(\neg house \wedge large)$[1] and a house with a medium surface and a garden $(house \wedge medium \wedge garden)$.

[1] where $\neg house$ encodes an apartment.

Note that none of positive goals is rejected. Also the goal to have a medium and cheap apartment without garden is neither a negative goal nor a positive one. This is a goal in abeyance.

2.3 The Origins of the Goals

Agent's goals come generally from two different sources:

- from beliefs that justify their existence. So, the agent believes that the world is in a state that warrants the existence of its goals. These goals are called the *initial* ones or also *conditional* goals. They are conditional because they depend on the beliefs.
- an agent can adopt a goal because it allows him to achieve an initial goal. These are called *sub-goals* or *adopted goals*.

Example 4 (Trip to Central Africa).
Let's consider an agent who wants to go to Central Africa because there is a conference there. The goal jca is derived from the belief $Conference$.
The agent believes that to go to Central Africa, it should get tickets (t) and to be vaccinated (vac). To get tickets, the agent can either pass to an agency (ag) or ask a friend of him to get them (fr). Similarly, to be vaccinated, the agent has the choice between *going to a doctor* (dr) or *going to the hospital* (hop). Thus, t, vac, ag, fr, dr and hop become sub-goals of that agent.

3 Arguing About Goals

As mentioned above, there are two kinds of goals: the *initial/conditional* goals and the *adopted* ones called also *sub-goals*. These goals are justified or supported by two different kinds of arguments: *explanatory arguments* and *instrumental arguments*.

Before presenting formally these two types of arguments, we will start by presenting the logical language which will be used throughout this paper. In what follows, $\mathcal{L}$ will denote a propositional language. $\vdash$ denotes classical inference and $\equiv$ denotes logical equivalence.

Definition 1 (Conditional Rules).
A conditional rule *is an expression of the form*

$$R : \phi_1 \wedge \ldots \wedge \phi_n \Rightarrow \phi$$

where R is the name *of the rule and each ϕ_i and ϕ are* literals *of $\mathcal{L}$. The conjunction at the left of the arrow is the* antecedent *and the literal at the right is its* consequent.

A conditional rule expresses the fact that if $\phi_1 \ldots \phi_n$ are true then the agent will have the goal ϕ. Similarly, we will define the planning rules.

Definition 2 (Planning Rules).
A planning rule *is an expression of the form*

$$P : \varphi_1 \wedge \ldots \wedge \varphi_n \mapsto \varphi$$

where P is the name *of the rule and each φ_i and φ are* literals *of $\mathcal{L}$. The conjunction at the left of the arrow is the* antecedent *and the literal at the right is its* consequent.

Such a formulae means that the agent believes that if he realizes $\varphi_1, \ldots, \varphi_n$ then he will be able to achieve φ.

Remark 1. Note that the implications used to define both conditional rules and planning rules are not the material implication.

In what follows, we suppose that the agent's beliefs are more or less certain and that its conditional goals or planning rules may not have equal priority.

Definition 3. *An agent is equipped with three bases* $<\mathcal{B}, \mathcal{B}_c, \mathcal{B}_p>$ *such that:*

- $\mathcal{B} = \{(\alpha_i, a_i) : i = 1, \ldots, n\}$ *with* α_i *is a propositional formulae of the language* $\mathcal{L}$ *and* a_i *its* certainty degree*. This base contains the basic beliefs of the agent.*
- $\mathcal{B}_c = \{(R_j, b_j) : j = 1, \ldots, m\}$ *where* R_j *is a conditional rule and* b_j *represents the* priority degree *of the consequent of* R_j.
- $\mathcal{B}_p = \{(P_k, c_k) : k = 1, \ldots, l\}$ *where* P_k *is a planning rule and* c_k *represents the priority degree of this rule.*

In what follows, we suppose that a_i, b_i and c_i belong to the interval $(0, 1]$. Moreover, we shall denote by $\mathcal{B}^*$, $\mathcal{B}_c^*$ and $\mathcal{B}_p^*$ the corresponding sets when the weights are ignored i.e.

- $\mathcal{B}^* = \{\alpha_i : (\alpha_i, a_i) \in \mathcal{B}, i = 1, \ldots, n\}$
- $\mathcal{B}_c^* = \{R_j : (R_j, b_j) \in \mathcal{B}_c, j = 1, \ldots, m\}$
- $\mathcal{B}_p^* = \{P_k : (P_k, c_k) \in \mathcal{B}_p, k = 1, \ldots, l\}$

Once the language is introduced, we are now able to define formally the *potential initial goals* and the *sub-goals*.

Definition 4 (Initial goal — Sub-goal).
Let an agent be equipped with $<\mathcal{B}, \mathcal{B}_c, \mathcal{B}_p>$.

- $\mathcal{IG} = \{\phi$ *s.t* $\exists\ \phi_1 \wedge \ldots \wedge \phi_n \Rightarrow \phi \in \mathcal{B}_c^*\}$ *is the set of potential* initial goals *of the agent.*
- $Sub\mathcal{G}$ *is the set of potential* sub-goals *of the agent: A literal* $\varphi' \in Sub\mathcal{G}$ *iff there exists a rule* $\varphi_1 \wedge \varphi' \ldots \wedge \varphi_n \mapsto \varphi \in \mathcal{B}_p^*$ *with* $\varphi \in \mathcal{IG}$ *or* $\varphi \in Sub\mathcal{G}$*. In that case,* φ' *is a sub-goal of* φ.

Remark 2. Note that in the above definition, we speak about potential initial goals of the agent. The reason is that we are not sure that the antecedents of the corresponding rules are true. Consequently, if a potential goal is not adopted by the agent (the antecedents are not true), then it is not useful for the agent to realize its plan.

Example 5 (Trip to Central Africa).
Let us consider an agent who has the two following goals:

1. To go on a journey to central Africa if there is a Conference there. (jca)
2. To finish a publication before going on a journey. (fp)

Thus, $\mathcal{B}_c = \{(conf \Rightarrow jca, 0.6), (\Rightarrow fp, 0.8)\}$.
In addition to the goals, the agent is supposed to have beliefs on the way of achieving a given goal (we suppose that all the rules have equal priority):

$$\mathcal{B}_p^* = \begin{cases} t \wedge vac \mapsto jca \\ w \mapsto fp \\ ag \mapsto t \\ fr \mapsto t \\ hop \mapsto vac \\ dr \mapsto vac \end{cases}$$

and $\mathcal{B} = \{(w \rightarrow \neg ag, 0.8), (w \rightarrow \neg dr, 0.8), (conf, 0.8), (can, 0.4), (can \rightarrow \neg conf, 1)\}$.
with: conf = "a conference", can = "to be canceled", t = “to get the tickets”, vac = “to be vaccinated”, w = “to work”, ag = “to go to the agency”, fr = “to have a friend who may bring the tickets”, hop = “to go to the hospital”, dr = “to go to a doctor”.
In this example, $\mathcal{IG}$ = {jca, fp} and $Sub\mathcal{G}$ = {t, vac, ag, fr, dr, hop, w}.

4 Explanatory Arguments

Explanatory arguments are used to explain / to give a reason of adopting a given goal. They are also used to give reasons for and against beliefs. In this section, we will propose an argumentation system which constructs explanatory arguments from the different bases of an agent and which evaluates them.

4.1 Basic Definitions

Definition 5 (Explanatory argument).
An explanatory argument is a pair <H, h> such that:

- $H \subseteq \mathcal{B}^* \cup \mathcal{B}_c^*$.
- $H \vdash h$.
- H *is consistent and minimal (for set inclusion).*

$\mathcal{A}_g$ denotes the set of all arguments such that $h \in \mathcal{IG}$. In other terms, it gathers all the arguments supporting initial goals. $\mathcal{A}_b$ gathers all the arguments supporting beliefs (i.e $h \notin \mathcal{IG}$). Finally, $\mathcal{A} = \mathcal{A}_g \cup \mathcal{A}_b$.

Definition 6 (Sub-argument).
Let <H, h>, $< H', h' > \in \mathcal{A}$. $< H, h >$ is a sub-argument *of* $< H', h' >$ *iff* $H \subseteq H'$.

Example 6 (Trip to Central Africa).
The arguments $<\{conf, conf \Rightarrow jca\}, jca>$ and $<\{\Rightarrow fp\}, fp> \in \mathcal{A}_g$. However, the argument $<\{can, can \rightarrow \neg conf\}, \neg conf> \in \mathcal{A}_b$.

Remark 3. Note that the implication used in conditional rules is not material and it has no contrapositive. So, for example the set $\{x, x \rightarrow \neg y, g \Rightarrow y\}$ does not infer $\neg g$. Consequently, $<\{x, x \rightarrow \neg y, g \Rightarrow y\}, \neg g>$ is not an explanatory argument.

4.2 The Strength of Explanatory Arguments

As mentioned before, each of the three bases $<\mathcal{B}, \mathcal{B}_c, \mathcal{B}_p>$ is pervaded with certainty or priority. From the certainty degrees, we define the certainty level of an argument.

Definition 7 (Certainty level of an explanatory argument).
Let $A = <H, h> \in \mathcal{A}$*. The* certainty level *of* $<H, h>$*, denoted by* $level(A) = min\{a_i \mid \varphi_i \in H \cap \mathcal{B}^*$ *and* $(\varphi_i, a_i) \in \mathcal{B}\}$*. If* $H \cap \mathcal{B}^* = \emptyset$ *then* $level(A)$ *= 1.*

Remark 4. Note that the priority degree of a given conditional goal is not taken into account in the definition of the strength of its supporting argument. In fact, the intuition behind a conditional goal is that: “the agent will adopt the goal, with its associated priority degree, if the condition is satisfied”. So even if the goal is very desired by the agent, if the conditions are not satisfied, then that goal will not be pursued.

The certainty level of the arguments makes it possible to compare different arguments as follows:

Definition 8 (Preference relation).
Let A_1 *and* A_2 *be two arguments in* $\mathcal{A}$*.* A_1 *is* preferred *to* A_2*, denoted by* $A_1 \succ A_2$*, iff* $level(A_1) > level(A_2)$*.*

Example 7 (Trip to Central Africa).
In the above example, the certainty level of the argument $<\{conf, conf \Rightarrow jca\}, jca>$ is 0.8. Whereas, the certainty level of the argument $<\{can, can \rightarrow \neg conf\}, \neg conf>$ is 0.4. The certainty level of the argument $<\{\Rightarrow fp\}, fp>$ is 1. Thus, $<\{\Rightarrow fp\}, fp> \succ <\{conf, conf \Rightarrow jca\}, jca> \succ <\{can, can \rightarrow \neg conf\}, \neg conf>$.

4.3 Conflicts Between Explanatory Arguments

An explanatory argument can be defeated either on one of its beliefs or one of its conditional goals. For example, the argument supporting the goal of going to Central Africa because there is a conference can be defeated by another argument stating that the conference has actually been canceled. This kind of defeat is modeled by the relation of "undercut" defined as follows:

Definition 9 (Undercut relation).
Let $<H, h>$*,* $<H', h'> \in \mathcal{A}$*.* $<H, h>$ undercuts $<H', h'>$ *iff* $\exists\, h'' \in H' \cap \mathcal{B}$ *such that* $h \equiv \neg h''$*.*

A conditional goal can also be defeated. For instance, the argument of going to Central Africa can be defeated by an argument stating that there is no money and if there is no money then the agent cannot go to Central Africa $< \{NoMoney, NoMoney \rightarrow \neg jca\}, \neg jca >$. This kind of defeat is modeled by the following relation of "rebut".

Definition 10 (Rebut relation).
Let $<H, h>$*,* $<H', h'> \in \mathcal{A}$*.* $<H, h>$ rebuts $<H', h'>$ *iff* $h' \in \mathcal{IG}$ *and* $h \equiv \neg h'$*.*

The two relations of conflicts are brought together in a unique relation called *attack*:

Definition 11 (Attack relation).
Let $<H, h>$ *and* $<H', h'> \in \mathcal{A}$. $<H, h>$ attacks $<H', h'>$ *iff:*

- $<H, h>$ *undercuts* $<H', h'>$ *and not(*$<H', h'> \succ <H, h>$*) or*
- $<H, h>$ *rebuts* $<H', h'>$ *and not(*$<H', h'> \succ <H, h>$*) or*
- $<H, h>$ *rebuts a sub-argument of* $<H', h'>$ *and not(*$<H', h'> \succ <H, h>$*).*

4.4 The Acceptability of Explanatory Arguments

We can now define the argumentation system we will use to evaluate our arguments:

Definition 12 (Argumentation system).
An argumentation system *(AS) is a pair* $\langle \mathcal{A}, Attack \rangle$ *such that* $\mathcal{A}$ *is the set of all explanatory arguments built from* $\mathcal{B} \cup \mathcal{B}_c$.

This system will return three categories of explanatory arguments:

- The class $\underline{\mathrm{S}}$ of *acceptable* explanatory arguments. Goals supported by such arguments are really justified and they may be the "positive goals" that an agent will pursue, if they are achievable.
- The class $\underline{\mathrm{R}}$ of *rejected* arguments. An argument is rejected if it is attacked by an acceptable one. Goals supported only by such arguments will be rejected by the agent even if they can be achieved. They will represent the negative goals of the agent.
- The class $\underline{\mathrm{C}}$ of arguments *in abeyance*. Such arguments are neither acceptable nor rejected. $\underline{C} = \mathcal{A} \setminus (\underline{S} \cup \underline{R})$.

In what follows, we will define the class of acceptable arguments. For that purpose, we will start by presenting the notion of defence introduced in [7].

Definition 13 (Defence).
Let A, B *be two arguments of* $\mathcal{A}$ *and* $S \subseteq \mathcal{A}$. S defends A *iff for every argument* B *where* B *attacks* A*, there is some argument in* S *which attacks* B.

Henceforth, the set C_{Attack} will gather all non-attacked arguments. We can show that the set $\underline{\mathcal{S}}$ of acceptable arguments of the argumentation system $\langle \mathcal{A}, Attack \rangle$ is the least fixpoint of a function $\mathcal{F}$:

$$\mathcal{S} \subseteq \mathcal{A},$$
$$\mathcal{F}(\mathcal{S}) = \{(H, h) \in \mathcal{A}(\Sigma) | (H, h) \text{ is defended by } \mathcal{S}\}.$$

Proposition 1. *Let* $\langle \mathcal{A}, Attack \rangle$ *be an an argumentation system. The set of its* acceptable *arguments is:*

$$\underline{\mathcal{S}} = \bigcup \mathcal{F}_{i \geq 0}(\emptyset) = C_{Attack} \cup [\bigcup \mathcal{F}_{i \geq 1}(C_{Attack})].$$

Example 8 (Trip to Central Africa).
In this example, the argument $<\{\Rightarrow fp\}, fp>$ is not attacked then it is acceptable. The argument $<\{conf, conf \Rightarrow jca\}, jca>$ is preferred to its unique undercutting argument $<\{can, can \rightarrow \neg conf\}, \neg conf>$. Then it is not attacked and consequently it is also acceptable.

Let T be a set of arguments. The function $Supp(T) = \cup H_i$ such that $<H_i, h_i> \in T$. In other terms, the function $Supp$ returns the union of all the supports of arguments of T. We can show the following result:

Proposition 2. *Let $\langle \mathcal{A}, Attack \rangle$ be an argumentation framework and $\underline{S}$ its set of acceptable arguments. Then $Supp(\underline{S})$ is* consistent.

Property 1. Let $A \in \mathcal{A}$. If A is acceptable then each sub-argument B of A is also acceptable.

5 Instrumental Arguments

An agent may have another kind of goals. These last are not derived from the current beliefs of the agent, but from the plans to achieve the initial goals. In fact, they are justified by the fact that they will contribute to the achievement of initial goals. They are thus considered as sub-goals of the initial goals.

In [1], an argumentation framework which handles conflicting goals has been developed. This framework takes as input a set of initial goals, a belief base and a base of planning rules and returns the goals which can be achieved together, as well as the appropriate plans (i.e. the sub-goals). In this section, we will present an extended version of that framework which takes into account the priorities of the goals.

5.1 Basic Definitions

An agent may have one or several ways to achieve a given goal. We bring the two notions together in a new notion of *partial plan*.

Definition 14 (Partial plan).
A partial plan *is a pair $a = <H, h>$ such that:*

- *h is an initial goal or a sub-goal.*
- *$H = \{\varphi_1, \ldots, \varphi_n\}$ if there exists a rule $\varphi_1 \wedge \ldots \wedge \varphi_n \mapsto h \in \mathcal{B}_p^*$, $H = \emptyset$ otherwise.*

The function $Goal(a) = h$ returns the initial goal or sub-goal of a given partial plan "a" and the function $Plan(a) = H$ returns the support of the partial plan. $\aleph$ will gather all the partial plans that can be built from $<\mathcal{IG}, \mathcal{B}, \mathcal{B}_p>$.

Note that a goal may have several partial plans.

Remark 5. Let $a = <H, h>$ be a partial plan. Each element of the support H is a sub-goal of h.

Definition 15. *A partial plan* $a = <H, h>$ *is* elementary *iff* $H = \emptyset$.

Remark 6. If there exists an elementary partial plan for a goal h, it means that the agent knows how to achieve directly h.

A partial plan shows the actions that should be performed in order to achieve the corresponding goal (or sub-goal). However, the elements of the support of a given partial plan are considered as sub-goals that must be achieved at their turn by another partial plan. The whole way to achieve a given goal is called in [1] a *complete plan*. A *complete plan* for a given goal g is an *AND* tree. Its nodes are partial plans and its arcs represent the sub-goal relationship. The root of the tree is a partial plan for the goal g. It is an AND tree because all the sub-goals of g must be considered. When for the same goal, there are several partial plans to carry it out, only one is considered in a tree.

Definition 16 (Instrumental argument).
An instrumental argument *is a pair* $<G, g>$ *such that* $g \in \mathcal{IG}$ *and* G *is a finite tree such that:*

- *The root of the tree is a partial plan* $<H, g>$ *.*
- *A node* $<\{\varphi_1, \ldots, \varphi_n\}, h'>$ *has exactly* n *children* $<H'_1, \varphi_1>, \ldots, <H'_n, \varphi_n>$ *where* $<H'_i, \varphi_i>$ *is a partial plan for* φ_i.
- *The leaves of the tree are elementary partial plans.*

The function $Nodes(G)$ *returns the set of all the partial plans of the tree* G. $\mathcal{A}'$ *will denote the set of all the instrumental arguments that can be built from the bases* $<\mathcal{IG}, \mathcal{B}, \mathcal{B}_p>$.

Example 9. The goal jca has four instrumental arguments: $<g_1, jca>$, $<g_2, jca>$, $<g_3, jca>$ and $<g_4, jca>$. The goal fp has only one instrumental argument $<g_5, fp>$ (see figure 1 for the trees g_i).

5.2 The Strength of Instrumental Arguments

As mentioned before, the base $\mathcal{B}_c$ is pervaded with priority. From the priority degrees, we define the weight of an instrumental argument.

Definition 17 (Weight of an instrumental argument).
Let $A = <G, g> \in \mathcal{A}'$. *The* weight *of* $<G, g>$ *is* $Weight(A) = min\{b_i\}$ *such that* $(R_i, b_i) \in \mathcal{B}_c$ *and* $R_i = \varphi_1 \wedge \ldots \wedge \varphi_n \Rightarrow g$.

In other words, the weight of an instrumental argument is exactly the degree of priority/importance of the corresponding goal. The weights make it possible to compare different arguments as follows:

Definition 18 (Preference relation).
Let A_1 *and* A_2 *be two arguments in* $\mathcal{A}'$. A_1 *is* preferred *to* A_2, *denoted by* $A_1 \succ A_2$, *iff* $Weight(A_1) > Weight(A_2)$.

Example 10. The weight of the four instrumental arguments $A_1 = <g_1, jca>$, $A_2 = <g_2, jca>$, $A_3 = <g_3, jca>$ and $A_4 = <g_4, jca>$ is 0.6 whereas the weight of the argument $A_5 = <g_5, fp>$ is 0.8. Hence, $<g_5, fp> \succ <g_1, jca>, <g_2, jca>, <g_3, jca>$ and $<g_4, jca>$.

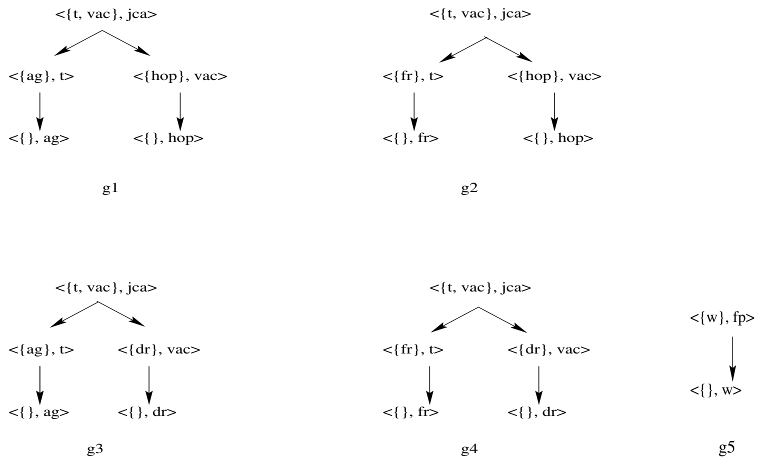

Fig. 1. Complete plans

5.3 Conflicts Between Instrumental Arguments

In [1], it has been shown that partial plans may be conflicting for several reasons.These different kinds of conflicts are brought together in a unique relation of *conflict* defined as follows:

Definition 19 (Conflict).
Let a_1 *and* a_2 *be two partial plans of* $\aleph$. a_1 conflicts *with* a_2 *iff:*

$$\{Goal(a_1), Goal(a_2)\} \cup Plan(a_1) \cup Plan(a_2) \cup \mathcal{B}^* \cup \mathcal{B}_p^* \vdash \bot.$$

Example 11. In example 5, a_{11a} = $<\{$ag$\}$, t$>$ conflicts with a_2 = $<\{w\}$, fp$>$. Indeed, Plan(a_{11a}) $\cup \mathcal{B}^* \vdash \{\neg w\}$ and Plan(a_2) = $\{w\}$.

More generally, a set of partial plans may be conflicting.

Definition 20. *Let* $S \subseteq \aleph$. *S is conflicting iff*

$$\bigcup_{a \in S} (\{Goal(a)\} \cup Plan(a)) \cup \mathcal{B}^* \cup \mathcal{B}_c^* \vdash \bot.$$

Since partial plans may be conflicting, two instrumental arguments may be conflicting too.

Definition 21 (Defeat).
Let $<G_1, g_1>$, $<G_2, g_2> \in \mathcal{A}'$. $<G_1, g_1>$ *defeats* $<G_2, g_2>$ *iff* $\exists a_1 \in Nodes(G_1)$ *and* $\exists a_2 \in Nodes(G_2)$ *such that* a_1 *conflicts with* a_2 *and not(*$<G_2, g_2> \succ <G_1, g_1>$*).*

More generally we are interested in conflict-free sets of instrumental arguments.

Definition 22 (Conflict-free).
Let $S = \{<G_i, g_i>: i = 1, \ldots, n\} \subseteq \mathcal{A}'$. *S is* conflict-free *iff*

$$[\bigcup_{<G,g> \in S} [\bigcup_{a \in\ Nodes(G)} (Plan(a) \cup \{Goal(a)\})] \cup \mathcal{B}^* \cup \mathcal{B}_p^* \nvdash \bot].$$

If $S = \{< G, g >\}$ *then we say that the argument* $< G, g >$ *is conflict-free.*

Proposition 3. *Let $S \subseteq \mathcal{A}'$. If S is conflict-free then $\nexists$ A_1 and $\nexists$ A_2 in S such that A_1 defeats A_2.*

The following example shows that we can find an instrumental argument which is not conflict-free even if it does not defeat itself.

Example 12. X is an agent equipped with the following bases:

$$\mathcal{IG} = \{d\},\ \mathcal{B}^* = \{b' \wedge c' \rightarrow \neg a\} \text{ and } \mathcal{B}_p^* = \begin{cases} a' & \mapsto a \\ b' & \mapsto b \\ c' & \mapsto c \\ a \wedge b \wedge c & \mapsto d \end{cases}$$

There is a unique instrumental argument for d whose set of nodes is conflicting.

Obviously a goal which has no conflict-free instrumental argument will be called *unachievable*. This means it is impossible to carry out such a goal.

5.4 Acceptability of Instrumental Arguments

From the preceding definitions, we can now present the formal system for handling instrumental arguments.

Definition 23. *Let's consider a triple <$\mathcal{B}$, $\mathcal{B}_c$, $\mathcal{B}_p$>. The pair <$\mathcal{A}'$, $Defeat$> will be called a system for handling instrumental arguments.*

As for explanatory arguments, this system will return three categories of instrumental arguments:

- The *acceptable* instrumental arguments. These arguments represent the *good plans* to achieve their corresponding goals.
- The class <u>R'</u> of *rejected* instrumental arguments. This class gathers the arguments which are not conflict-free and those defeated by acceptable arguments. Goals supported only by such arguments are *unachievable*.
- The class <u>C'</u> of arguments *in abeyance*. Such arguments are neither acceptable nor rejected.

Unlike the previous system, we may have here several acceptable sets. Each of them will correspond to a set of goals which can be achieved together.

Definition 24. *Let <$\mathcal{A}'$, $Defeat$> be a system and $S \subseteq \mathcal{A}'$. S is an* acceptable *set of arguments iff:*

- *S is conflict-free.*
- *S is maximal (for set inclusion).*

Let $\underline{S_1}$, . . ., $\underline{S_n}$ be the different sets of acceptable arguments.

Example 13. In example 5, there are four instrumental arguments (A_1, A_2, A_3, A_4) for the goal "going on a journey to Central Africa" and exactly one argument A_5 for the goal "finishing the paper". Moreover, A_5 defeats A_1, A_2 and A_3. We have exactly two acceptable sets of arguments:

- $\underline{S_1} = \{A_1, A_2, A_3, A_4\}$,
- $\underline{S_2} = \{A_4, A_5\}$

The purpose of an agent is to achieve a maximal subset of $\mathcal{I}G$. Consequently, among the sets of acceptable arguments, we will choose the ones which achieve maximal sets of desires (for set inclusion). In the above example, we will choose the set $\underline{S_2}$.

6 Computing Bipolar Goals

Once we have defined the two frameworks which evaluate the different arguments supporting goals, we are now able to define among the potential initial goals and sub-goals the positive goals of the agent, the negative ones and finally the goals in abeyance.

Definition 25 (Positive goals).
Let $g \in \mathcal{I}G$. g is a positive goal *iff:*

- $\exists\ (H, g) \in \mathcal{A}_g$ *such that* $(H, g) \in \underline{\mathcal{S}}$, *and*
- $\exists\ (G, g) \in \mathcal{A}'$ *and* $\exists\ \underline{S_i}$ *such that* $(G, g) \in \underline{S_i}$.

This means that a goal is positive if it is justified and it is achievable.
Note that the sub-goals of a positive initial goal are also considered as positive.

Definition 26 (Negative goals).
Let $g \in \mathcal{I}G$. g is a negative goal *iff:*

- $\forall\ (H, g) \in \mathcal{A}_g$, $(H, g) \in \underline{\mathcal{R}}$, *or*
- $\forall\ (G, g) \in \mathcal{A}'$, $(G, g) \in \underline{\mathcal{R}'}$

Indeed, a goal is negative if it is not justified or if it is unachievable.

Definition 27 (Goals in abeyance).
Let $g \in \mathcal{I}G$. g is in abeyance *iff it is neither positive nor negative.*

Example 14 (Trip to Central Africa).
In this example, the two initial potential goals of the agent (jca, fp) become positive goals of the agent. Moreover, the sub-goals (t, vac, fr, hop, w) are also positive goals. However, the following sub-goals are in abeyance: (dr, ag).

7 Handling Bipolar Goals

Let's consider an agent P who has the following belief base: $\mathcal{B} = \{(\neg age > 40, 1), (PhD, 0.5)\}$. Suppose that this agent has two potential goals: to be a president and/or to be a professor. $\mathcal{B}_c = \{(age > 40 \Rightarrow president, 1), (PhD \Rightarrow professor, 1)\}$.

To be a president, the agent knows that he should have more than 40 years old however he has less than 40 years. According to its beliefs, there is no argument in favor of this goal. Consequently, the agent will keep this goal in *abeyance*.

Concerning the goal of becoming a professor, it has the following explanatory argument: $A_1 =< \{PhD, PhD \Rightarrow professor\}, professor >$. This argument is not attacked at all, thus it is acceptable. In this example, there are no instrumental arguments. So, the goal $professor$ is a *positive* one. Note that at this stage, the agent has one positive goal, one goal in abeyance an no negative goals.

Suppose that one year later, the agent has more than 40 years old then $\mathcal{B}$ is updated as follows: $\mathcal{B} = \{(age > 40, 1), (PhD, 0.5)\}$.

Using this new base, we can find an argument in favor of the goal $president$, namely: $A_2 =< \{age > 40, age > 40 \Rightarrow president\}, president >$ which is acceptable since it is not attacked. Thus, the goal $president$ which was in abeyance will become a positive one. In sum, the agent has now two positive goals that it will pursue.

This agent applies for the job of president and starts a negotiation with the appropriate services. Let's imagine the following dialogue:

X: I want to become a president.
S: This entails that you will leave your actual job.
X: But I want to be a professor too. I can do both jobs.
S: It's impossible. You should not have another job.
X: Okey.

When the agent receives the new information which says that he cannot have two jobs so he pursues either the goal *president* or the goal *professor* but not both, he updates its belief base: $\mathcal{B} = \{(president \rightarrow \neg professor, 1), (age > 40, 1), (PhD, 0.5)\}$.
Due to this change of beliefs we have the following arguments: A_1, A_2 computed above and $A_3 =< \{age > 40, age > 40 \Rightarrow president, president \rightarrow \neg professor\}, \neg professor >$, $A_4 =< PhD, PhD \Rightarrow professor, president \rightarrow \neg professor\}, \neg president >$ and $A_5 =< \{age > 40, age > 40 \Rightarrow president, PhD, PhD \Rightarrow professor\}, \neg(president \rightarrow \neg professor) >$.

The certainty level of A_1, A_4 and A_5 is 0.5 and the certainty level of A_2 and A_3 is 1. Thus, A_2, $A_3 \succ A_1$, A_4 and A_5. We can check easily that A_4 rebuts A_2 but since A_2 is preferred to A_4 then A_2 is not attacked and consequently it is acceptable. The argument A_3 rebuts and defeats A_1. Moreover, A_3 is preferred to its unique undercutting argument A_5. Then A_3 is also acceptable. Consequently, A_1 is rejected. The goal of being president is supported by an acceptable argument A_2 then this goal is positive. However, the goal of being a professor is supported only by the rejected argument A_1 then it is a negative goal.

8 Conclusion

In most negotiation literature, each negotiating agent is supposed to have a set of fixed and predefined goals. It is not clear where do these goals come from and how an agent selects them. Argumentation-based negotiation makes an advance by supposing that the goals are not fixed during a negotiation and may change. However, even in these works the goals are predefined and it is not clear how they are changed. In fact, since there is no work on how goals are computed, it seems difficult to model the way in which they are updated.

The aim of this paper is twofold. First, it presents the goals in a bipolar way. In fact, an agent has positive goals that it will pursue and also negative goals that it does not want to achieve. This second category of goals is very important in negotiation since the offers that satisfy such goals will be rejected by the agent. The second aim of this paper is to present a formal framework which computes the goals of an agent. We have shown through an example how an agent may change its goals during a negotiation.

The principle of goals generation proposed in this paper is close to the one proposed in [14] in the context of planning, where an argument-based generation of goals is implicitly used. However in this latter, the author *only* generates the positive goals (called *wants*), those that the agent will pursue, from a set of initial conditional goals called *wishes*. Moreover, the set of beliefs on which initial conditional goals are based is flat (i.e., all the beliefs are equally certain) then there is no evaluation of arguments when they are conflicting.

In this paper, we have shown that a goal may be supported by two different kinds of arguments: the exaplanatory arguments and the instrumental ones. We have then presented two different systems for handling each category of arguments. An extension of our work will be to handle the two kinds of arguments in a unique framework. We are actually working in this direction. We are also planning to investigate more deeply the handling of bipolar goals in a negotiation dialogue.

Acknowledgments

This work was supported by the Commission of the European Communities under contract IST-2004-002307, ASPIC project "Argumentation Service Platform with Integrated Components".

References

1. L. Amgoud. A formal framework for handling conflicting desires. In *Proceedings of the 7th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU'2003*, pages 552–563, 2003.
2. L. Amgoud, S. Parsons, and N. Maudet. Arguments, dialogue, and negotiation. In *Proceedings of the 14th European Conference on Artificial Intelligence (ECAI'00)*, pages 338–342, 2000.
3. S. Benferhat, D. Dubois, S. Kaci, and H. Prade. Bipolar representation and fusion of preferences in the possibilistic logic framework. In *Proceedings of the eighth International Confenrence on Principle of Knowledge Representation and Reasoning (KR'02)*, pages 158–169, 2002.
4. J. Borod. The neuropsychology of emotion. Oxford University Press,, 2000.
5. J. Cacioppo and G. Bernston. The affect system: Architecture and operating characteristics. *Current Directions in Psychological Science*, 8, 5:133–137, 1999.
6. J. Cacioppo, W. Gardner, and G. Bernston. Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review*, 1, 1:3–25, 1997.
7. P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.
8. S. Kaci and H. Prade. Bipolar goals. A possibilistic logic characterization of preferred choices. In *Proceedings of Workshop on Local Computation for Logics and Uncertainty, in conjunction with ECAI'04*, 2004.

9. S. Kraus, K. Sycara, and A. Evenchik. *Reaching agreements through argumentation: a logical model and implementation*, volume 104. Journal of Artificial Intelligence, 1998.
10. X. Luo, N. Jennings, N. Shadbolt, H. fung Leung, and J. H. man Lee. A fuzzy constraint based model for bilateral, multi-issue negotiations in semi-competitive environments. In *Artificial Intelligence, 148(1-2)*, pages 53–102, 2003.
11. I. Rahwan, L. Sonenberg, and F. Dignum. Towards interest-based negotiation. In *Proceedings of the Second International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'03)*, pages 773–780, 2003.
12. E. Rolls. Precis of "brain and emotion". *Behavioral and Brain Sciences*, 23(2):177–234, 2000.
13. C. Sierra, N. Jennings, P. Noriega, and S. Parsons. A framework for argumentation-based negotiation. In *Proceedings of the fourth Workshop on Agent Theories, Architectures and Languages (ATAL'97)*, pages 167–182, 1997.
14. R. Thomason. Desires and defaults: A framework for planning with inferred goals. In *Proceedings of the seventh International Confenrence on Principle of Knowledge Representation and Reasoning (KR'00)*, pages 702–713, 2000.
15. O. Wong and R. Lau. Possibilistic reasoning for intelligent payment agents. In *Proceedings of the second Workshop on AI in Electronic Commerce AIEC*, pages 170–180, 2000.