Practical reasoning as a generalized decision making problem

Leila Amgoud, Henri Prade

Practical reasoning as a generalized decision making problem

Leila Amgoud Henri Prade
amgoud@irit.fr prade@irit.fr

IRIT–CNRS
118, Route de Narbonne
31062 Toulouse cedex 09, FRANCE

Résumé :

La prise de décision, souvent vue comme une forme de raisonnement sur les actions, a été considérée de différents points de vue. La théorie classique de la décision, développée principalement par des économistes, s'est concentrée sur l'identification et la justification de critères, tels que l'utilité espérée, pour comparer différentes alternatives. Cette approche prend en entrée un ensemble d'actions qui sont atomiques faisables, et une fonction qui évalue les conséquences de chaque action. Un trait remarquable mais aussi une limitation de cette approche est la réduction du problème de décision à la disponibilité de deux fonctions : une fonction de distribution de probabilité et une fonction d'utilité. C'est pourquoi certains chercheurs en IA ont préconisé le besoin d'une approche dans laquelle tous les aspects qui interviennent dans un problème de décision (tels que les désirs d'un agent, la faisabilité des actions, etc..) sont explicitement représentés. Dans cette perspective, des architectures BDI (Beliefs, Desires, Intentions) ont été proposées. Elles prennent leur inspiration dans le travail de philosophes sur ce que les anglo-saxons nomment practical reasoning ou le "raisonnement pratique". Le raisonnement pratique traite principalement de la pertinence au contexte, de la faisabilité et finalement des intentions retenues et exécutables. Cependant, ces approches souffrent d'un manque de formulation claire de règles de décision qui combinent les considérations ci-dessus pour décider quelle action exécuter.

Dans cet article, nous montrons que le raisonnement pratique est un problème de la prise de décision généralisé. L'idée fondamentale est qu'au lieu de comparer des actions atomiques, on compare des ensembles d'actions. L'ensemble préféré d'actions devient les intentions retenues par l'un agent.
Le papier présente un cadre unifié qui bénéficie des avantages des trois approches (décision classique, architectures BDI, l'idée générales du raisonnement pratique). Plus précisément, nous proposons un cadre formel qui prend en entrée un ensemble de croyances, un ensemble de désirs conditionnels, et un ensemble de règles présisant comment des désirs peuvent être réalisés, et renvoie en sortie un sous-ensemble cohérent de désirs ainsi que les actions pour les réaliser. De telles actions s'appellent les intentions. En effet, nous montrons que ces intentions sont choisies par l'intermédiaire de quelques règles de décision. Ainsi, selon que l'agent ait une attitude optimiste ou pessimiste, l'ensemble des intentions peut ne pas être le même.

Mots-clés : Raisonnement pratique, Théorie de la décision, Argumentation

Abstract:

Decision making, often viewed as a form of reasoning toward action, has been considered from different points of view. *Classical* decision theory, as developed by economists, has focused mainly on identifying *criteria* such as expected utility for comparing different alternatives. The inputs of this approach are a set of *feasible* atomic actions, and a function that assesses the value of their consequences when the actions are performed in a given state. One of the main practical limitation of this approach is the fact that it reduces the whole decision problem to the availability of two functions: a probability distribution and a utility function. This is why some researchers in AI have advocated the need for a different approach in which all the aspects that may be involved in a decision problem (such as the desires of an agent, the feasibility of actions, etc) are explicitly represented. Hence, BDI architectures have been developed. They take their inspiration in the work of philosophers who have advocated practical reasoning. Practical reasoning mainly deals with the adoption, filling in, and reconsideration of intentions. However, these approaches suffer from a lack of a clear formulation of decision rules that combine the above qualitative concepts to decide which action to perform.

In this paper, we argue that practical reasoning is a generalized decision making problem. The basic idea is that instead of comparing atomic actions, one has to compare sets of actions. The preferred set of actions becomes the intentions of the agent. The paper presents a unified setting that benefits from the advantages of the three above-mentioned approaches (classical decision, BDI, practical reasoning). More precisely, we propose a formal framework that takes as input a set of beliefs, a set of conditional desires, and a set of rules stating how desires can be achieved, and returns a consistent subset of desires as well as ways/actions for achieving them. Such actions are called intentions. Indeed, we show that these intentions are generated via some decision rules. Thus, depending on whether the agent has an optimistic or a pessimistic attitude, the set of intentions may not be the same.

Keywords: Practical reasoning, Decision making, Argumentation theory

1 Introduction

Decision making, often viewed as a form of reasoning toward action, has raised the interest of many scholars including philosophers, economists, psychologists, and computer scientists for a long time. Any decision problem amounts to select the best option(s) among different alternatives.

The decision problem has been considered from different points of view. *Classical* decision theory, as developed by economists, has focused mainly on identifying *criteria* for comparing different alternatives. The inputs of this approach are a set of *feasible* actions, and a function that assesses the value of their consequences when the actions are performed in a given state. The output is a preference relation between actions. A decision criterion, such as the classical expected utility [11], should then be justified on the basis of a set of postulates to which the preference relation between action should obey. Note that such an approach considers a group of candidate actions as a whole rather than focusing on a candidate action individually. Moreover, the candidate actions are supposed to be feasible.

More recently, some researchers in AI have advocated the need for a new approach in which the different aspects that may be involved in a decision problem (such as the goals of the agent, the feasibility of an action, its consequences, the conflicts between goals, the alternative plans for achieving the same goal, etc) can be handled. In [5, 6], it has been argued that this can be done by representing the cognitive states, namely agent's beliefs, desires and intentions (thus the so-called *BDI* architecture). The decision problem is then to select among the conflicting desires a consistent and feasible subset that will constitute the intentions. The above line of research takes its inspiration in the work of philosophers who have advocated practical reasoning [10]. Practical reasoning mainly deals with the adoption, filling in, and reconsideration of intentions and plans. It follows two main steps: 1) *deliberation*, in which an agent decides what state of affairs it wants to achieve –that is, its *desires*; and (2) *means-ends reasoning*, in which an agent devises plans for achieving these desires.

In this paper, we argue that practical reasoning is a generalized decision making problem. The basic idea is that instead of comparing atomic actions, one has to compare sets of coherent plans (i.e. plans that can be achieved together) that will achieve the desires computed at the deliberation step. The preferred set of plans becomes the intentions of the agent. The paper presents a formal framework for practical reasoning that works in three steps: at the first step one computes, from a set of conditional desires, a set of arguments supporting them, and a conflict relation among these arguments, a set of what is called *justified desires*. These desires can be pursued provided that they have plans for achieving them. The second step computes sets of plans that can are achievable together. The input is the set of conditional desires, a set of plans (whose structure and origin are not discussed here), a function specifying for each conditional desire the plans for achieving it, and finally a set of conflicting plans. The framework returns extensions of plans. An extension is a set of plans that can be achieved together. Once, these sets identified, one applies decision making techniques for ordering these extensions. The idea is to prefer the set that achieves the most important desires returned at the deliberation level.

The paper is organized as follows: we start by presenting our abstract framework of practical reasoning, then we illustrate it on an example. Then we compare our work with existing works in the literature. The last section is devoted to some concluding remarks and perspectives.

2 General framework for practical reasoning

Practical reasoning is the reasoning toward action. It follows three main steps:

1. Generating desires to be achieved, called also *deliberation*
2. Generating plans for achieving those desires, called *means-end reasoning*
3. Selecting the intentions to be pursued by the agent. The intentions are the plans that will be performed for reaching the generated desires.

In what follows, $\mathcal{L}$ will denote a logical language. From $\mathcal{L}$, we distinguish a finite set $\mathcal{D}$ of potential conditional desires. Desires will be denotes by $d_1, \ldots, d_n$. Some desires may be more important than others. This is captured by a partial preordering $\succeq_d$ on $\mathcal{D}$, thus $\succeq_d \subseteq \mathcal{D} \times \mathcal{D}$.

Similarly, from $\mathcal{L}$, different *arguments* can be built. An argument may provide a reason of generating or adopting a given desire. Let $\mathcal{A}$ denote the set of these arguments whose structure and origin are not known.

Since knowledge bases may be inconsistent, arguments may be conflicting too. These conflicts are captured by a binary relation $\mathcal{R}_a \subseteq \mathcal{A} \times \mathcal{A}$.

Let us define a function $\mathcal{F}_d$ that returns for each desire d_i in $\mathcal{D}$ the set of arguments supporting it. Thus,

$$\mathcal{F}_d : \mathcal{D} \rightarrow 2^{\mathcal{A}}$$

for instance, $\mathcal{F}_d(d_1) = \{a_1, \ldots, a_n\}$ with $\{a_1, \ldots, a_n\} \subseteq \mathcal{A}$. Note that some desires may not be supported by arguments. Such desires will not be considered as intentions. We assume that an argument cannot support two or more desires at the same time. Formally: $\forall d_i, d_j$, $\mathcal{F}_d(d_i) \cap \mathcal{F}_d(d_j)$ = $\emptyset$.

We assume that we have a set $\mathcal{P}$ = $\{p_1, \ldots, p_m\}$ of *plans*. A plan is a way of achieving a desire. The structure and the origin of the plans are left unknown.

Plans are related to the desires they achieve by the following function

$$\mathcal{F}_p : \mathcal{D} \rightarrow 2^{\mathcal{P}}.$$

Each plan is assumed to achieve at least one desire, i.e. $\forall d_i, d_j \in \mathcal{D}$, $\mathcal{F}_p(d) \cap \mathcal{F}_p(d')$ = $\emptyset$.

It is very common that a given plan may not be achievable because, for instance, it has a consequence that contradicts the desire it wants to achieve. It is also possible that two or more plans cannot be achievable at the same time since, for instance they yield to conflicting situations. Such conflicts among elements of $\mathcal{P}$ are given by a set $\mathcal{R}_p \subseteq 2^{\mathcal{P}}$. We assume that only minimal conflicts are given in $\mathcal{R}_p$, this means that $\nexists S, S' \in \mathcal{R}_{\mathcal{P}}$ such that $S \subseteq S'$. Let us consider the following example.

Example 1 *Let* $\mathcal{D} = \{d_1, d_2, d_3\}$, $\mathcal{A} = \{a_1, a_2, a_3, a_4\}$, $\mathcal{R}_a = \{(a_1, a_2), (a_2, a_3)\}$, $\mathcal{F}_d(d_1) = \{a_3\}$, $\mathcal{F}_d(d_2) = \{a_4\}$, $\mathcal{F}_d(d_3) = \emptyset$, $\mathcal{P} = \{p_1, p_2, p_3\}$, $\mathcal{F}_p(d_1) = \{p_1\}$, $\mathcal{F}_p(d_2) = \{p_2\}$, $\mathcal{F}_p(d_3) = \{p_3\}$, *and* $\mathcal{R}_p = \{\{p_2\}, \{p_1, p_3\}\}$.

2.1 A general framework for deliberation

This section aims at generating the desires that can be pursued by the agent (in case they are feasible, i.e. they have plans). As shown in the above illustrative example, one may have conditional desires that depend on some beliefs. The idea is to check whether the conditions of these desires hold in the current state of the world. In the above example, both desires d_1 and d_2 are generated since their conditions hold.

In our general framework, we suppose that an argument is built for supporting a desire as soon as the conditions on which it depends hold. However, since a knowledge base may be inconsistent, i.e. the condition may hold but, at the same time there is an information which contradicts it, counter-arguments can be built. Thus, the generated desires, or the outcome of the deliberation step, is the result of a simple argumentation system defined as follows.

Definition 1 (Argumentation system) *An argumentation system for generating desires to be pursued in a pair* $\langle \mathcal{A}, \mathcal{R}_a \rangle$.

In [7], different acceptability semantics have been introduced for computing the status of arguments. These are based on two basic concepts, *defence* and *conflict-free*, defined as follows:

Definition 2 (Defence/conflict-free) *Let* $S \subseteq \mathcal{A}$.

- S defends *an argument* a *iff each argument that defeats* a *is defeated in the sense of* $\mathcal{R}_a$ *by some argument in* S.
- S *is* conflict-free *iff there exist no* a, a' *in* S *such that* $a \; \mathcal{R}_a \; a'$.

Definition 3 (Acceptability semantics) *Let* S *be a conflict-free set of arguments, and let* $\mathcal{T}$: $2^{\mathcal{A}} \rightarrow 2^{\mathcal{A}}$ *be a function such that* $\mathcal{T}(S) = \{a \mid S$ *defends* $a\}$.

- S *is a* complete *extension iff* $S = \mathcal{T}(S)$.
- S *is a* preferred *extension iff* S *is a maximal (w.r.t set* $\subseteq$*) complete extension.*
- S *is a* grounded *extension iff it is the smallest (w.r.t set* $\subseteq$*) complete extension.*

Let $\mathcal{E}_1, \ldots, \mathcal{E}_x$ *denote the different extensions under a given semantics.*

Note that there is only one grounded extension. It contains all the arguments that are not defeated, and those arguments that are defended directly or indirectly by non-defeated arguments.

Now that the acceptability semantics defined, we are ready to define the status of any argument.

Definition 4 (Argument status) *Let* $\langle \mathcal{A}, \mathcal{R}_a \rangle$ *be an argumentation system, and* $\mathcal{E}_1, \ldots, \mathcal{E}_x$ *its extensions under a given semantics. Let* $a \in \mathcal{A}$.

1. a *is* accepted *iff* $a \in \mathcal{E}_i$, $\forall \mathcal{E}_i$ *with* $i = 1, \ldots, x$.
2. a *is* rejected *iff* $\nexists \mathcal{E}_i$ *such that* $a \in \mathcal{E}_i$.
3. a *is* undecided *iff* a *is neither accepted nor rejected. This means that* a *is in some extensions and not in others.*

On the basis of the status of each argument, it is now possible to compute the set of desires that are supposed to be justified in the current state of the world. As said before, this will represent the outcome of the deliberation step.

Definition 5 (Justified desires) *Let $\mathcal{D}$ be a set of potential desires. The* justified desires *are gathered in the set* `Output` $= \{d_i \in \mathcal{D}$ *such that* $\exists a \in \mathcal{A}$*,* a *is accepted, and* $a \in \mathcal{F}_d(d_i)\}$.

Example 2 (Example 1 continued) *Let* $\mathcal{D} = \{d_1, d_2, d_3\}$, $\mathcal{A} = \{a_1, a_2, a_3, a_4\}$, $\mathcal{R}_a = \{(a_1, a_2), (a_2, a_3)\}$, $\mathcal{F}_d(d_1) = \{a_3\}$, $\mathcal{F}_d(d_2) = \{a_4\}$, $\mathcal{F}_d(d_3) = \emptyset$. *In this example, the argumentation system* $\langle \mathcal{A}, \mathcal{R}_a \rangle$ *returns only one grounded extension* $\{a_1, a_3, a_4\}$*. Thus, the output of the deliberation is* $\{d_1, d_2\}$*. The desire* d_3 *is not supported by arguments, thus there is no reason to generate this desire.*

Note that the generated desires will not necessarily be pursued by an agent. They should also be feasible.

2.2 A general framework for means-end reasoning

The second step of practical reasoning consists of looking for plans to achieve desires. Since an agent may have several desires at the same time, then it needs to know not only which desire is achievable, but also which subsets of desires can be achieved together. In what follows, we propose an abstract framework that returns extensions of plans, i.e. sets of coherent plans, and thus subsets of desires that can be pursued at the same time. This framework takes as input the following elements: $\mathcal{D}$, $\mathcal{P}$, $\mathcal{F}_p$, and $\mathcal{R}_p$.

Definition 6 *A framework for generating feasible plans is a pair* $\langle \mathcal{P}, \mathcal{R}_p \rangle$.

Here again, we are looking for groups of plans that are achievable together. This means that the plans should not be conflicting. Thus, the extensions should be *conflict-free*:

Definition 7 (Conflict-free) *Let* $S \subseteq \mathcal{P}$. S *is* conflict-free *iff* $\nexists\, S' \subseteq S$*, such that* $S' \in \mathcal{R}_p$.

Definition 8 (Extension of plans) *Let* $S \subseteq \mathcal{P}$. S *is an* extension *iff:*

- S *is conflict-free*
- S *is maximal for set inclusion among subsets of* $\mathcal{P}$ *that satisfies the first condition.*

$\mathcal{S}_1, \ldots, \mathcal{S}_n$ *will denote the different extensions of plans.*

As for arguments, it also possible to define the status of each plan as follows:

Definition 9 (Status of plans) *Let* $p \subseteq \mathcal{P}$.

- p *is* feasible *iff* $\exists \mathcal{S}_i$ *such that* $p \in \mathcal{S}_i$
- p *is* unachievable *iff* $\nexists \mathcal{S}_i$ *such that* $p \in \mathcal{S}_i$
- p *is* universally feasible *iff* $\forall \mathcal{S}_i$, $p \in \mathcal{S}_i$. *This means that such a plan is feasible with other plans.*

On the basis of the status of plans, one can define the status of each desire. Four cases are distinguished:

Definition 10 (Status of desires) *Let* $d \subseteq \mathcal{D}$.

- d *is* achievable *iff* $\exists p \in \mathcal{F}_p(d)$ *such that* p *is feasible*

- d *is* unachievable *iff* $\forall p \in \mathcal{F}_p(d)$, *$p$ is unachievable*
- d *is* universally feasible *iff* $\exists p \in \mathcal{F}_p(d)$ *such that p is universally feasible*
- d *is* universally accepted *iff* $\forall p \in \mathcal{F}_p(d)$, *$p$ is universally feasible*

The desires achieved by each extension are returned by a function defined as follows:

Definition 11 *Let $\mathcal{S}_i$ be an extension of the framework $\langle \mathcal{P}, \mathcal{R}_p \rangle$.*
$\texttt{Desires}(\mathcal{S}_i) = \{d_j \in \mathcal{D}$ *s.t.* $\exists p \in \mathcal{S}_i$ *and* $\mathcal{F}_p(d_j) = p\}$.

Example 3 (Example 1 continued) $\mathcal{P} = \{p_1, p_2, p_3\}$, $\mathcal{F}_p(d_1) = \{p_1\}$, $\mathcal{F}_p(d_2) = \{p_2\}$, $\mathcal{F}_p(d_3) = \{p_3\}$, *and* $\mathcal{R}_p = \{\{p_2\}, \{p_1, p_3\}\}$.

The set $\mathcal{R}_p$ means that the plan p_2 is not achievable, and that the two plans p_1, and p_3 cannot be achieved together. Thus, the system $\langle \mathcal{P}, \mathcal{R}_p \rangle$ will return two extensions: $\mathcal{S}_1 = \{p_1\}$, and $\mathcal{S}_2 = \{p_3\}$, with $\texttt{Desires}(\mathcal{S}_1) = \{d_1\}$ *and* $\texttt{Desires}(\mathcal{S}_2) = \{d_3\}$.

It is clear that the desire d_2 is unachievable, and the two desires d_1, d_3 cannot be pursued at the same time. The agent should select only one of them.

2.3 Selecting intentions

In the previous section, we have proposed a framework that returns extensions of plans, i.e. plans that may co-exist together. However, as shown before, several extensions may exist at the same time. One needs to select the one that will constitute the intentions of the agent. A preordering $\triangleright$ on the set $\{\mathcal{S}_1, \ldots, \mathcal{S}_n\}$ is then needed. This is a decision making problem. This latter amounts to defining a pre-ordering, usually a complete one, on a set of possible alternatives, on the basis of the different consequences of each alternative. In [1], it has been shown that argumentation can be used for defining such a pre-ordering. The idea is to construct arguments in favor of and against each alternative, to evaluate such arguments, and finally to apply some principle for comparing pairs of alternatives on the basis of the quality or strength of their arguments. In that framework, atomic actions are ordered. In what follows, we will extend the framework to the case of sets of plans, i.e. instead of ordering atomic actions, we will define a preordering on the set $\mathcal{E} = \{\mathcal{S}_1, \ldots, \mathcal{S}_n\}$.

The main ingredients that are involved in the definition of an argumentation-based decision framework are the following:

Definition 12 (Decision framework) *An* argumentation-based decision framework *is a tuple $\langle \mathcal{E}, \mathcal{A}_e, \succeq_e \rangle$ where:*

- *$\mathcal{E}$ is the set of possible alternatives.*
- *$\mathcal{A}_e$ is a set of arguments supporting/attacking elements of $\mathcal{E}$.*
- *$\succeq_e$ is a (partial or complete) preordering on $\mathcal{A}_e$.*

The output is a preordering $\triangleright$ on $\mathcal{E}$. $\mathcal{S}_i \triangleright \mathcal{S}_j$ means that the extension $\mathcal{S}_i$ is preferred to the extension $\mathcal{S}_j$.

Once the relation $\triangleright$ is identified, one can compute the intentions of an agent. The intentions are the set of plans belonging to the most preferred extension w.r.t. $\triangleright$, and which achieve generated desires.

Definition 13 (The intentions) *The set of* intentions *is* $\{p_i \in \mathcal{S}_j | \ p_i \in \mathcal{F}_p(d), \ d \in \texttt{Output}$, *and* $\forall \mathcal{S}_k, \ \mathcal{S}_j \triangleright \mathcal{S}_k\}$.

Arguments. A decision may have arguments in its favor (called PROS), and arguments against it (called CONS). Arguments PROS point out the existence of good consequences for a given decision. In our application, an argument PRO an extension $\mathcal{S}_i$ points out the fact it achieves a generated desire, i.e. an element of the set `Output`. Formally:

Definition 14 (Arguments PROS) *Let $\mathcal{S}_i \in \mathcal{E}$. An* argument in favor of*, or PRO, the extension $\mathcal{S}_i$ is a triple $A = \langle p_j, \mathcal{S}_i, d_k \rangle$ such that $p_j \in \mathcal{S}_i$, $p_j \in \mathcal{F}_p(d_k)$, and $d_k \in$* `Output`.
Let Arg_P be the set of all such arguments that can built.

Note that there are as many arguments as plans to carry out the same desire. Arguments CONS highlight the existence of bad consequences for a given decision, or the absence of good ones. Arguments CONS are defined by exhibiting a generated desire that is not achieved by the extension. Formally:

Definition 15 (Arguments CONS) *Let $\mathcal{S}_i \in \mathcal{E}$. An* argument against*, or CONS, the extension $\mathcal{S}_i$ is a pair $A = \langle \mathcal{S}_i, d_k \rangle$ such that $\nexists p_j \in \mathcal{S}_i$, $p_j \in \mathcal{F}_p(d_k)$, and $d_k \in$* `Output`.
Let Arg_C be the set of all such arguments that can built.

Note that some arguments may be stronger than others. For instance, an argument $A = \langle p_j, \mathcal{S}_i, d_k \rangle$ in favor of the extension $\mathcal{S}_i$ may be preferred to an argument $B = \langle p'_j, \mathcal{S}_i, d_l \rangle$ if the desire d_k is preferred to the desire d_l. In this case, the preference relation $\succeq_e$ is based on a preference relation $\succeq_d$ between the potential desires of $\mathcal{D}$. The relation $\succeq_e$ can also be defined on the basis of the plans themselves. For instance, one may prefer the argument A over the argument B if the cost of p_j is lower than the cost of the plan p'_j.

Some decision criteria. Different criteria for defining the preordering $\rhd$ on $\mathcal{E}$ can be defined. In what follows, we will present some examples borrowed from [1], and adapted to our application, i.e. ordering sets of plans.

In what follows, $\texttt{Goals}_X(\mathcal{S}_i)$ be a function that returns for a given decision or extension $\mathcal{S}_i$, all the desires for which there exists an argument of type X (i.e. PROS or CONS) with conclusion $\mathcal{S}_i$.
Let $\mathcal{S}_i, \mathcal{S}_j \in \mathcal{E}$.

$$\mathcal{S}_i \rhd_1 \mathcal{S}_j \text{ iff } \texttt{Goals}_P(\mathcal{S}_i) \neq \emptyset, \text{ and } \texttt{Goals}_P(\mathcal{S}_j) = \emptyset \quad (1)$$

The above criterion prefers the extension that achieves generated desires. This can be refined as follows:

$$\mathcal{S}_i \rhd_2 \mathcal{S}_j \text{ iff } \texttt{Goals}_P(\mathcal{S}_i) \supset \texttt{Goals}_P(\mathcal{S}_j) \quad (2)$$

The above criterion prefers the extension that achieves more generated desires. This partial preorder can be further refined into a complete preorder as follows:

$$\mathcal{S}_i \rhd_3 \mathcal{S}_j \text{ iff } |\texttt{Goals}_P(\mathcal{S}_i)| > |\texttt{Goals}_P(\mathcal{S}_j)| \quad (3)$$

3 Illustrative example

Let us consider an agent who has the two following conditional desires:

1. To go on a journey to central Africa if he is in holidays. ($hol \rightarrow jca$)
2. To finish a publication if there is a deadline of a conference. ($conf \rightarrow fp$)

In addition to the desires, the agent is supposed to have beliefs on the way of achieving a given desire:

$$\begin{cases} t \wedge vac & \rightarrow & jca \\ w & \rightarrow & fp \\ ag & \rightarrow & t \\ fr & \rightarrow & t \\ hop & \rightarrow & vac \\ dr & \rightarrow & vac \end{cases}$$

with: 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”.

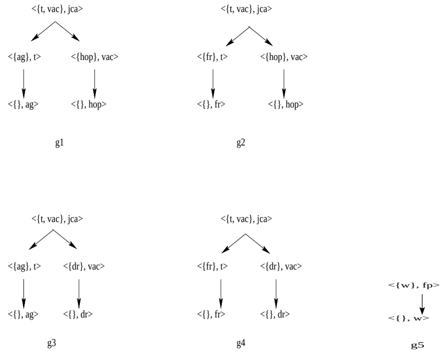

Figure 1: Complete plans

For example, the rule $t \wedge vac \rightarrow jca$ means that the agent believes that if he gets tickets and he is vaccinated then he will be able to go on a journey in central Africa. The rule $w \rightarrow fp$ expresses that the agent believes that if he works then he will be able to finish his paper. To get tickets, the agent can either visit an agency or ask a friend of him to get them. Similarly, to be vaccinated, the agent has the choice between *going to a doctor* or *going to the hospital*. In these two last cases, the agent has two ways to achieve the same desire.

An agent may have also another kind of beliefs representing integrity constraints and facts. In our example, we have:

$$\begin{cases} hol & & \\ conf & & \\ w & \rightarrow & \neg ag \\ w & \rightarrow & \neg hop \end{cases}$$

The two latter rules mean that the agent believes that if he works, he can neither visit an agency nor go to a doctor. In this example, the two conditional desires jca and fp are justified in the current state of the world since the they depend on beliefs (respectively hol and $conf$) that are true. Moreover, both desires have at least a plan for achieving them. However, some ways of achieving the desires are conflicting.

Of course, it would be ideal if all the desires can become intentions. As our example illustrates, this may not always be the case. We will answer the following questions: which desires will be *pursued* by the agent and with *which plans*?

In this example, we have two arguments in favor of the conditional desires jca and fp. Let $A = \langle\{hol, hol \rightarrow jca\}, jca\rangle$ and $B = \langle\{conf, conf \rightarrow fp\}, fp\rangle$. These arguments are not defeated at all, thus they belong to the grounded extension of the argumentation system. Consequently, `Output` = $\{jca, fp\}$.

there are four complete plans (g_1, g_2, g_3, g_4) for the desire ‘going on a journey to central africa’ and exactly one complete plan g_5 for the desire ‘finishing the paper’. These are given in figure 1. Moreover, g_5 attacks g_1, g_2 and g_3. Thus, there are exactly two extensions:

- $\mathcal{S}_1 = \{g_1, g_2, g_3, g_4\}$
- $\mathcal{S}_2 = \{g_4, g_5\}$

The extension $\mathcal{S}_1$ is supported by four arguments:

- $A_1 = \langle g_1, \mathcal{S}_1, jca\rangle$
- $A_2 = \langle g_2, \mathcal{S}_1, jca\rangle$

- $A_3 = \langle g_3, \mathcal{S}_1, jca \rangle$
- $A_4 = \langle g_4, \mathcal{S}_1, jca \rangle$

The four arguments exhibits the same desire jca. However, the extension $\mathcal{S}_2$ is supported by only two arguments:

- $B_1 = \langle g_4, \mathcal{S}_1, jca \rangle$
- $B_2 = \langle g_5, \mathcal{S}_1, fp \rangle$

However, the two arguments refer to two different desires. According to criterion (2), it is clear that $\mathcal{S}_2$ is preferred to $\mathcal{S}_1$ since $\texttt{Goals}_P(\mathcal{S}_2) \supset \texttt{Goals}_P(\mathcal{S}_1)$, knowing that $\texttt{Goals}_P(\mathcal{S}_1) = \{jca\}$, and $\texttt{Goals}_P(\mathcal{S}_1) = \{jca, fp\}$. The intensions to be pursued by the agent are then $\{g_4, g_5\}$.

4 Related works

Recently, a number of attempts have been made to use formal models of argumentation as a basis for practical reasoning. Some of these models (e.g. [2, 3, 8]) are instantiations of the *abstract* argumentation framework of Dung [7]. Others (e.g. [9, 12]) are based on an encoding of argumentative reasoning in logic programs. Finally, there are frameworks based on empirical approaches to practical reasoning and persuasion (e.g. [4, 13]). Our framework builds on the former, and is therefore a contribution towards formalising practical reasoning using abstract argumentation systems.

Amgoud [2] presented an argumentation framework for generating consistent plans from a given set of desires and planning rules. This was later extended with argumentation frameworks that generate the desires themselves (see below).

Amgoud and Kaci [3] have a notion of "conditional rule", which is meant to generate desires from beliefs. Our framework is more general in the sense that we don't specify how arguments are built from bases. Indeed, the structure and the origin of the arguments are left unknown and can be instantiated with any logic. Moreover, in that work it is not clear how intentions are chosen.

Hulstijn and van der Torre [8], on the other hand, have a notion of "desire rule", which contains only desires in the consequent. But their approach is still problematic. It requires that the selected goals are supported by goal trees which contain both desire rules and belief rules that are deductively consistent. This consistent deductive closure again does not distinguish between desire literals and belief literals (see Proposition 2 in [8]). This means that one cannot both believe $\neg p$ and desire p. Here again, the selection of intention is left unsolved.

5 Conclusion

This paper has presented the first general and abstract framework for practical reasoning. It shows that this latter generalizes the decision making problem.

We presented a formal model for reasoning about desires (generating desires and plans for achieving them) based on argumentation theory. We adapted the notions of attack and preference among arguments in order to capture the differences in arguing about desires and plans.

One of the main advantages of our framework is that, being grounded in argumentation, it lends itself naturally to facilitating dialogues about desires and plans. Indeed, we are currently extending our framework with dialogue game protocols in order to facilitate negotiation and persuasion among agents. Another interesting area of future work is investigating the relationship between our framework and axiomatic approaches to BDI agents.

References

[1] L. Amgoud and H. Prade. Explaining qualitative decision under uncertainty by argumentation. In *Proc. of the 21st National Conference on Artificial Intelligence, AAAI'06*, pages 219–224, 2006.

[2] Leila Amgoud. A formal framework for handling conflicting desires. In Thomas D. Nielsen and Nevin Lianwen Zhang, editors, *Proc. ECSQARU*, volume 2711 of *LNCS*, pages 552–563. Springer, Germany, 2003.

[3] Leila Amgoud and Souhila Kaci. On the generation of bipolar goals in argumentation-based negotiation. In Iyad Rahwan et al, editor, *Proc. 1st Int. Workshop on Argumentation in Multi-Agent Systems (ArgMAS)*, volume 3366 of *LNCS*. Springer, Germany, 2005.

[4] Katie Atkinson, Trevor Bench-Capon, and Peter McBurney. Justifying practical reasoning. In C. Reed F. Grasso and G. Carenini, editors, *Proc. Workshop on Computational Models of Natural Argument (CMNA)*, pages 87–90, 2004.

[5] M. Bratman. *Intentions, plans, and practical reason.* Harvard University Press, Massachusetts., 1987.

[6] M. Bratman, D. Israel, and M. Pollack. *Plans and resource bounded reasoning.*, volume 4. Computational Intelligence., 1988.

[7] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321–358, 1995.

[8] Joris Hulstijn and Leendert van der Torre. Combining goal generation and planning in an argumentation framework. In Anthony Hunter and Jerome Lang, editors, *Proc. Workshop on Argument, Dialogue and Decision, at NMR*, Whistler, Canada, June 2004.

[9] Antonis Kakas and Pavlos Moraitis. Argumentation based decision making for autonomous agents. In *Proc. 2nd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 883–890, Melbourne, Australia, 2003.

[10] J. Raz. Practical reasoning. *Oxford, Oxford University Press*, 1978.

[11] L. J. Savage. *The Foundations of Statistics*. Dover, New York, 1954. Reprinted by Dover, 1972.

[12] Guillermo R. Simari, Alejandro J. Garcia, and Marcela Capobianco. Actions, planning and defeasible reasoning. In *Proc. 10th International Workshop on Non-Monotonic Reasoning*, pages 377–384, 2004.

[13] Yuqing Tang and Simon Parsons. Argumentation-based dialogues for deliberation. In Frank Dignum et al, editor, *Proc. AAMAS, Utrecht, The Netherlands*, pages 552–559, New York NY, USA, 2005. ACM Press.