Formalizing practical reasoning under uncertainty: An argumentation-based approach

Leila Amgoud
IRIT - CNRS - UPS
118, route de Narbonne
31062 Toulouse Cedex 9, France
amgoud@irit.fr

Henri Prade
IRIT - CNRS - UPS
118, route de Narbonne
31062 Toulouse Cedex 9, France
prade@irit.fr

Abstract

Practical reasoning (PR), as advocated by philosophers is concerned by reasoning about what agents should do. It follows mainly two steps. A deliberation one for identifying the goals to be achieved, and a means-ends reasoning step for choosing the ways of achieving them. The PR literature has mainly proposed informal patterns of inference for describing such a process in simple situations. Moreover, this line of thoughts has influenced some AI researchers who proposed BDI architectures. Namely, agents are supposed to have beliefs and to entertain desires from which they elicit the intentions to be pursued. The interest of such an approach is to emphasize some aspects involved in a decision problem that are not explicitly dealt with by classical approaches, in particular the feasibility of actions, and the generation of agent's goals. However, there is no complete formalization of the whole PR in the BDI literature.

The paper aims at providing an abstract framework for PR. It is based on argumentation techniques for both deliberation and for selecting subsets of compatible actions, possibly in presence of uncertainty. The framework returns a consistent subset of desires as well as ways/actions for achieving them. Such actions are called intentions. We show that these intentions are generated via some decision criteria. Thus, depending on whether the agent has an optimistic or a pessimistic attitude, the set of intentions may not be the same. Indeed, we show that PR leads to a generalized decision making problem, where instead of comparing atomic actions, one compares sets of actions.

1 Introduction

Practical reasoning (PR) [11, 13], is concerned with the generic question "what is the right thing to do for an agent in a given situation". To answer this question, authors (e.g. [16]) have proposed a two steps process. The first step, often called *deliberation*, consists of identifying the goals of the agent. In the second step, they look for ways for achieving those goals, i.e. for actions or plans. Such an approach raises issues such as: how goals are generated? Are actions feasible? Do actions have undesirable consequences? Are sub-plans compatible? Are there alternative plans for achieving a given goal? etc. In [7, 12], it has been argued that this can be done by representing the cognitive states, namely agent's beliefs, desires and intentions (leading to the so-called *BDI* architecture). This requires a rich knowledge/preference representation setting.

What is worth noticing in most works on practical reasoning is the use of argument schemes for providing reasons for choosing or discarding an action. For instance, an action may be considered as potentially useful on the basis of the so-called *practical syllogism* [15]: i) G is a goal for agent X, ii) Doing action A is sufficient for agent X to carry out goal G, iii) Then, agent X ought to do action A. The above syllogism, which would apply to the means-end reasoning step, is in essence already an argument in favor of doing action A. However, this does not mean that the action is warranted, since other arguments (called counter-arguments) may be built or provided against the action. Those counter-arguments refer to critical questions identified in [15] for the above syllogism. In particular, relevant questions are "Are there alternative ways of realizing G?", "Is doing A feasible?", "Has agent X other goals than G?", "Are there other consequences of doing A which should be taken into account?". Recently in [4], the above syllogism has been extended to explicitly take into account the reference to ethical values in arguments.
What is also worth pointing out is that some researchers like [4] have claimed that practical reasoning is essentially a decision making task. This is not completely true if we consider that deliberation and checking the feasibility of sets of plans are pure inference problems. However, selecting

among different feasible sets of plans aiming at achieving justified desires (returned at the deliberation step) is indeed a decision problem, which in our approach will constitute a third step.

The paper presents a formal framework for practical reasoning that works in three steps. At the first step, from a set of conditional desires, a set of arguments supporting them, and a conflict relation among these arguments, one computes 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 should be compatible in the sense that they are achievable together. Such sets of plans are called *extensions*. The input is the set of conditional desires, a set of plans assumed to be known or provided by a planning system (the generation of these plans is outside the scope of the paper), a function specifying for each conditional desire the plans achieving it, and finally a set of conflicting plans. The framework returns different extensions as an output. The third step combines the results of the two first steps in order to return the best extension (according to particular *decision criteria*) that achieves justified desires. The decision criteria may, for instance, privilege the number and/or importance of achievable desires by the extension, the number of plans per desire in the extension (if we are interested in robust solutions with several possible plans for achieving a desire), etc. Thus, we show that PR leads to a generalized decision making problem, where instead of comparing atomic actions, one has to compare sets of actions.

The paper is organized as follows: we start by recalling the basic concepts behind argumentation theory, then we propose our abstract framework of practical reasoning, then we illustrate it on an example. We finally compare our work with existing works in the literature before concluding.

2 Argumentation theory: A reminder

Argumentation is a reasoning model based on the construction and the evaluation of interacting arguments. Those arguments are intended to support / explain / attack statements that can be decisions, opinions, etc.

Argumentation has been used in different domains such as nonmonotonic reasoning [14], handling inconsistency in knowledge bases [1, 5], or decision making [3, 6, 9]. In [8], Dung has defined an argumentation system as a pair of a set $\mathcal{A}$ of arguments whose structure and origin are unknown, and a binary relation $\mathcal{R}$ encoding conflicts among elements of $\mathcal{A}$, thus, $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$. Dung has mainly focused on identifying, among the conflicting arguments, the ones that can be considered as acceptable, i.e. the ones with which a dispute can be won. For that purpose, different acceptability semantics have been proposed. All of them are based on two basic concepts: *defence* and *conflict-free*.

Definition 1 (Defence/conflict-free) *Let S be a set of arguments of $\mathcal{A}$. S* defends *an argument a iff each argument that defeats a is defeated in the sense of $\mathcal{R}$ by some argument in S. S is* conflict-free *iff there exist no a, a' in S such that $a \; \mathcal{R} \; a'$.*

Definition 2 (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.*
We will write $\mathcal{E}_1, \ldots, \mathcal{E}_n$ to denote the different extensions under one of those semantics.

Note that there is only one grounded extension that may be empty. 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 can define the status of each argument.

Definition 3 (Argument status) *Let $\langle \mathcal{A}, \mathcal{R} \rangle$ be an argumentation system, and $\mathcal{E}_1, \ldots, \mathcal{E}_x$ its extensions under a given semantics. Let $a \in \mathcal{A}$.*
a is accepted *iff $a \in \mathcal{E}_i$, $\forall \mathcal{E}_i$ with $i = 1, \ldots, x$. a is* rejected *iff $\nexists \mathcal{E}_i$ such that $a \in \mathcal{E}_i$. a is* undecided *iff a is neither accepted nor rejected. This means that a is in some extensions and not in others.*

3 The practical reasoning problem

Practical reasoning is the reasoning toward action. In the literature, authors claim that it is a two steps process: deliberation and means-end reasoning. Moreover, some authors claim that it is a pure decision making problem. In this paper we argue that PR is rather a three steps process: 1) Deliberation which amounts to generate desires to be achieved 2) Means-end reasoning which consists of generating compatible plans for achieving those desires 3) Selecting the intentions to be pursued by the agent. The intentions are the plans that will be performed for reaching the generated desires. The deliberation step is merely an inference problem since it amounts to find a set of desires that are justified on the basis of the current state of the world and of conditional desires. Similarly, checking if a plan is feasible and does not lead to bad consequences is still a matter of inference. A decision problem only occurs when several plans are possible, and one of them has to be chosen at the third step.
In what follows, $\mathcal{L}$ will denote a logical language. From $\mathcal{L}$, we distinguish a finite set $\mathcal{D}$ of 'literals' encoding potential "desires". A desire is a state of affairs that an agent wants to reach, for instance 'to have a picnic'. It may be conditioned

by some beliefs or even by the satisfaction of another desire. Desires will be denoted 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 is a reason for adopting or discarding a given desire. For instance, it is known that today the weather is beautiful, then I can adopt the desire "to have a picnic". Let $\mathcal{A}$rg denote the set of these arguments whose structure and origin are not known. In the illustrative example, these arguments are instantiated. However, we only need to consider them in *abstracto* for presenting our formal framework.

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, $\mathcal{F}_d$: $\mathcal{D} \rightarrow 2^{\mathcal{A}rg}$. For instance, $\mathcal{F}_d(d_1) = \{a_1, \ldots, a_n\}$ with $\{a_1, \ldots, a_n\} \subseteq$ $\mathcal{A}$rg. Note that some desires may not be supported by arguments.

Conflicts among arguments may exists and are captured by a binary relation denoted by $\mathcal{R}_a \subseteq \mathcal{A}\text{rg} \times \mathcal{A}\text{rg}$. This relation will satisfy at least the following hypothesis:

Hypothesis 1 *Let $d, d' \in \mathcal{D}$. If $d \equiv \neg d'$ then $\forall a, a' \in \mathcal{A}rg$ such that $a \in \mathcal{F}_d(d)$, $a' \in \mathcal{F}_d(d')$, we have $a \ \mathcal{R}_a \ a'$.*

Let $\mathcal{P} = \{p_1, \ldots, p_m\}$ be a set of *plans*. A plan is a way of achieving a desire. The structure and the origin of the plans are left unknown. Moreover, we assume that these plans are provided by a planning system (not studied here), or are already known. Plans are related to the desires they achieve by the following function $\mathcal{F}_p$: $\mathcal{D} \rightarrow 2^{\mathcal{P}}$. It may be the case that a given plan is assumed to achieve only one desire.

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 $\mathcal{R}_p = \{\{p_2\}, \{p_1, p_3\}\}$.*

The conflict relation should capture at least the fact that contradictory desires should not be feasible at the same time.

Hypothesis 2 *Let $d, d' \in \mathcal{D}$. If $d \equiv \neg d'$ then $\forall p, p' \in \mathcal{A}rg$ such that $p \in \mathcal{F}_p(d)$, $p' \in \mathcal{F}_p(d')$, we have $\{p, p'\} \in \mathcal{R}_p$.*

4 Deliberation

This section aims at generating the desires that can be pursued by the agent (in case there are plans for them). 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 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 4 (Deliberation system) *An* argumentation system for deliberation *is a pair $\langle \mathcal{A}rg, \mathcal{R}_a \rangle$, where $\mathcal{A}rg$ is the set of arguments and $\mathcal{R}_a$ the defeat relation. We will write $E_1, \ldots, E_n$ to denote its extensions under a given Dung's semantics.*

On the basis of the status of each argument (computed as shown in Definition 3), 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` *= $\{d_i \in \mathcal{D} \mid \exists a \in \mathcal{A}rg$, a is accepted, and $a \in \mathcal{F}_d(d_i)\}$.*

Proposition 1 *Let $\langle \mathcal{A}rg, \mathcal{R}_a \rangle$ be a deliberation system. The set* `Output` *is consistent.*

Moreover, we can show that desires that are not supported by arguments will not be considered as justified.

Proposition 2 $\forall d \in \mathcal{D}$. *If $\mathcal{F}_d(d) = \emptyset$, then $d \notin$* `Output`.

Example 2 (Example 1 continued) *Let $\mathcal{D} = \{d_1, d_2, d_3\}$, $\mathcal{A}rg = \{a_1, a_2, a_3, a_4\}$, $\mathcal{R}_a = \{(a_1, a_2), (a_2, a_3)\}$, $\mathcal{F}_d(d_1) = \{a_3\}$, $\mathcal{F}_d(d_2) = \{a_4\}$, $\mathcal{F}_d(d_3) = \emptyset$. In this example, the argumentation system $\langle \mathcal{A}rg, \mathcal{R}_a \rangle$ 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.

5 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* $\nexists\, S' \subseteq S$ *s.t* $S' \in \mathcal{R}_p$.

Definition 8 (Extension of plans) *Let* $S \subseteq \mathcal{P}$. S *is an* extension *iff: 1)* S *is conflict-free, 2)* 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.*

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

Definition 9 *Let* $\mathcal{S}_i$ *be an extension of the framework* $\langle \mathcal{P}, \mathcal{R}_p \rangle$.

$$\texttt{Desires}(\mathcal{S}_i) = \{d_j \in \mathcal{D} \text{ s.t. } \exists p \in \mathcal{S}_i \text{ and } \mathcal{F}_p(d_j) = p\}.$$

Proposition 3 *Let* $\langle \mathcal{P}, \mathcal{R}_p \rangle$ *be a framework and* $\mathcal{S}_1, \ldots, \mathcal{S}_n$ *its extensions of plans.* $\forall \mathcal{S}_i$, $i = 1, \ldots, n$, $\texttt{Desires}(\mathcal{S}_i)$ *is consistent.*

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

Definition 10 (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.*

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.*

6 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 [3], it has been shown that argumentation can be used for defining such a preordering. 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 11 (Argumentation-based 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) pre-ordering 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 12 (The intentions) *The set of* intentions *is* $\mathcal{I} = \{p_i \in \mathcal{S}_j \,|\, p_i \in \mathcal{F}_p(d), d \in \texttt{Output}, \text{ and } \forall \mathcal{S}_k, \mathcal{S}_j \triangleright \mathcal{S}_k\}$.

Proposition 4 *The set* $\texttt{Desires}(\mathcal{I})$ *is consistent.*

6.1. 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 13 (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* $d_k \in \texttt{Output}$.
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 14 (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, or if the certainty of success of p_j is greater than the one of p'_j.

6.2. Some decision criteria

Different criteria for defining the preordering $\rhd$ on $\mathcal{E}$ can be defined. In what follows, we will show some examples borrowed from [3], and adapted to our application, i.e. ordering sets of plans instead of ordering single actions. Indeed, this shows clearly that our practical reasoning framework is a true generalization of classical decision making problems handled in an argumentative way as in [3], where a preference relation between single actions relies on the strengths of arguments PROS and CONS, expressed in terms of the certainty level with which the goals with high priorities are satisfied.
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 $\mathcal{S}_i, \mathcal{S}_j \in \mathcal{E}$.

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

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

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

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

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

When the strength of arguments is taken into account in the decision process, one may think of preferring a choice that has a dominant argument, i.e. an argument PROS that is preferred to any argument PROS the other choices. This principle is called promotion focus in [3].

$$\mathcal{S}_i \rhd_4 \mathcal{S}_j \text{ iff } \exists \langle p_k, \mathcal{S}_i, d_m \rangle \text{ such that } \forall \langle p'_k, \mathcal{S}_j, d'_m \rangle, \; \langle p_k, \mathcal{S}_i, d_m \rangle \succeq_e \langle p'_k, \mathcal{S}_j, d'_m \rangle \; (4)$$

Similarly, one may prefer the choice that has the weakest argument CONS.

7 Illustrative example

The illustrative example involves the set $\mathcal{D}$ of desires of an agent, its knowledge base $\mathcal{K}$, a set Ac of actions that it may perform, and a factual base F describing the current state of the world. In the encoding of the example, we use the following convention: capital letters for desires, lower case letters for propositions describing states of the world, and bold lower case letters for actions (thus **a** denotes the action a, while *a* expresses the fact that the action is been realized.

The agent has the following desires:

- "Not to get a cold" ($\neg C$)
- "Not to get a headache" ($\neg H$)
- "Get work finished in a acceptable way" (WA)
- "Get work finished in a perfect way" (WP)
- "If tired get a nap or get fresh air" ($tir \rightarrow N \vee F$)

Priorities between these desires will be introduced later.

The actions that the agent may perform are the following: $Ac = \{$"do nothing" (**DoNo**), "go outside" (**go**), "expedite work" (**ew**), "work carefully" (**wc**), "check work" (**chw**), "get a nap" (**n**), "to take aspirin" (**ta**), ¬**go**, ¬**ew**, ¬**wc**, ¬**chw**, ¬**n**, ¬**ta**$\}$.

The agent has the following knowledge base $\mathcal{K}$:

- ''doing nothing leads to not have cold" (**DoNo** $\rightarrow \neg C$)
- "getting a nap requires extra time" (**n** $\rightarrow N \wedge \neg eti$)
- "if it rains and the agent goes outside, there is a risk to get a cold" ($r \wedge go \rightarrow C$)
- "if the agent checks work, it may get an headache" (**chw** $\rightarrow chw \wedge H$)
- "to expedite a work leads to an acceptable finished work" (**ew** $\rightarrow ew \wedge WA$)
- "to work carefully is incompatible with a lack of extra time" ($\neg eti \rightarrow$ ¬**wc**)

- "to work carefully and then to check leads to a perfect finished work for sure" (**wc** $\wedge$ **chw** $\rightarrow wc \wedge chw \wedge WP$)
- "to get fresh air the agent has to go outside" (**go** $\rightarrow go \wedge F$)
- "to check an acceptable work leads in general to a perfect work" (**chw** $\wedge WA \rightarrow WP$)
- "in case of headache, the agent may take aspirin to cure it" (**ta** $\rightarrow ta \wedge \neg H$)

In addition to the above rules, we have the following facts: $F = \{r, tir\}$.

In our example, all the desires are justified, i.e. they belong to the set `Output`. Indeed, the desires $\neg C, \neg H, WA$, and WP are unconditional, thus they are justified. However, the desires N and F are conditional, but their disjunction is supported by the argument $\langle tir, tir \rightarrow N \vee F \rangle$ which is not defeated at all.

Regarding the feasibility of these desires, the following plans are built for achieving them:

- P_1: $\langle$ **DoNo** $\rightarrow \neg C \rangle$ for the desire $\neg C$
- P_2: $\langle$ **ta** $\rightarrow ta \wedge \neg H \rangle$ for the desire $\neg H$
- P_3: $\langle$ **ew** $\rightarrow ew \wedge WA \rangle$ for the desire WA
- P_4: $\langle$ **wc** $\wedge$ **chw** $\rightarrow wc \wedge chw \wedge WP \rangle$ for the desire WP
- P_5: $\langle$ **ew** $\rightarrow ew \wedge WA$, **chw** $\wedge WA \rightarrow WP \rangle$ for the desire WP
- P_6: $\langle$ **n** $\rightarrow N \wedge \neg eti \rangle$ for the desire N
- P_7: $\langle$ **go** $\rightarrow go \wedge F \rangle$ for the desire F

From the previous bases, the following set of conflicts can be built: $\mathcal{R}_p = \{\{P_1, P_7\}, \{P_2, P_4\}, \{P_2, P_5\}, \{P_4, P_5\}, \{P_4, P_6\}\}$. Thus, one can build six extensions of plans with their associates sets of desires:

- $S_1 = \{P_1, P_2, P_3, P_6\}$, $\texttt{Desires}(S_1) = \{\neg C, \neg H, WA, N\}$.
- $S_2 = \{P_1, P_3, P_4\}$, $\texttt{Desires}(S_2) = \{\neg C, WA, WP\}$.
- $S_3 = \{P_1, P_3, P_5, P_6\}$, $\texttt{Desires}(S_3) = \{\neg C, WA, WP, N\}$.
- $S_4 = \{P_2, P_3, P_6, P_7\}$, $\texttt{Desires}(S_4) = \{\neg H, WA, N, F\}$.
- $S_5 = \{P_3, P_4, P_7\}$, $\texttt{Desires}(S_5) = \{WA, WP, F\}$.
- $S_6 = \{P_3, P_5, P_6, P_7\}$, $\texttt{Desires}(S_6) = \{WA, WP, N, F\}$.

If one does not take into account neither the uncertainty, nor priorities between desires, applying decision criterion (2), the extension S_3 is preferred to S_2, and S_6 is preferred to S_5. However, the other extensions are not comparable. Using criterion (3), we have a complete preorder on the extension, and the best ones are S_1, S_3, S_4 and S_6.

Introducing priorities and uncertainty will allow us to refine the preordering. Let us now suppose that the desires may not have the same priority. We assume the following preference: $\neg C \succeq_d WA \succeq_d WP \succeq_d \neg H \succeq_d N \succeq_d F$.

In this case, it is clear that the extension S_3 is the best one since it satisfies the three most important desires and some other (missed by S_2, which is the second preferred extension), thus S_3 is the intention set.

This example exhibits three pieces of uncertain information in $\mathcal{K}$. In fact, one can distinguish between two types of uncertainty: i) the one pertaining to *side* effects of actions (getting cold when going outside, getting headache when checking), and ii) the one referring to the lack of certainty of satisfying the desire to which the action directly refers (checking an acceptable work does not always lead to a perfect work). Thus, this has consequences for plans P_4 (one may get H as a side effect), P_5 (one may get H as a side effect, or $\neg WP$), P_7 (one may get C as a side effect). This here gives birth to arguments CONS these plans. Consequently, the set of satisfied desires associated to the different extensions is now affected by this uncertainty. Namely, in S_2, S_3, S_4, S_5, and S_6, on may still have H if one is lucky enough, and one may still have C in S_4, S_5, and S_6 if one is lucky enough. In S_3 and S_6, one may have $\neg WP$ (instead of WP), if one is unlucky.

This may be the basis for defining a pure pessimistic attitude considering that nothing lucky happens and anything unlucky does not happen, and a pure optimistic attitude where this is the converse. For instance, an optimistic agent will consider that in S_6 checking will not lead to headache, nor going outside to cold, and then will consider that all the desires even both elements in the disjunction $N \vee F$ may be achieved. A pessimistic agent, on the contrary, will consider for instance that in S_3, WP is unsure and that only $\neg C$, WA and N are reached for sure, and so on. This may be further refined by distinguishing different levels of uncertainty following the approach presented in [3].

8 Related works and discussion

As already said, there has been mainly informal philosophy-oriented discussion on practical reasoning. Only recently, some AI researchers have advocated the need of formalizing this kind of reasoning, especially since it is in the core of agent's interaction, for instance deliberation has made by humans. Unfortunately, there has been a big mess about the exact nature of practical reasoning. Several researchers [4] consider it a pure inference problem. Others think that it is rather a pure decision making problem. In this paper, we argue that it is a three steps process involving two inference steps and one decision step. To the best of our knowledge, this is the first work that completely articulates the different steps of practical reasoning, and even identifies the main ingredients involved in such a problem. Moreover, it appear that the decision part of practical reasoning is a classical one (up to the fact one consider sets of actions rather than single actions). This paper has also provided a first general framework for practical reasoning based on an abstract argumentative machinery.

Due to a lack of a complete analysis of the whole practical reasoning process, the few existing attempts at formalising PR had until now only focused on one particular step, either the first or the second one. This is the case for the models (e.g. [2, 10]) that are instantiations of the *abstract* argumentation framework of Dung [8]. Along this line, there are also frameworks based on completely new theories of practical reasoning and persuasion (e.g. [4]). This is also true for the model developed by Hulstijn and van der Torre [10]. Their approach is even problematic. It requires that the selected desires are supported by desire 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. This means that one cannot both believe $\neg p$ and desire p. Here again, the selection of intentions is left unsolved.

An extension of this work would be the study of the formal properties of the general approach proposed here. Another worth considering idea would be to propose proof theories for the model. Indeed, instead of computing the whole extensions, it would be desirable to find out directly whether a desire can be achieved by the intention set of not. Another obvious line of research would be the formal introduction of the stratified possibilistic approach to qualitative decision for handling both the uncertainty and the priorities between desires as suggested by our illustrative example.

References

[1] L. Amgoud and C. Cayrol. Inferring from inconsistency in preference-based argumentation frameworks. *International Journal of Automated Reasoning*, Volume 29 (2):125–169, 2002.

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

[3] 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.

[4] K. Atkinson, T. Bench-Capon, and P. McBurney. Justifying practical reasoning. In *Proceedings of the Fourth Workshop on Computational Models of Natural Argument (CMNA 2004)*, pages 87–90, 2004.

[5] P. Besnard and A. Hunter. A logic-based theory of deductive arguments. *Artificial Intelligence*, 128 (1-2):203–235, 2001.

[6] B. Bonet and H. Geffner. Arguing for decisions: A qualitative model of decision making. In F. J. e. E. Horwitz, editor, *Proc. 12th Con. on Uncertainty in Artificial Intelligence (UAI'96)*, pages 98–105, Portland, Oregon, 1996.

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

[8] P. M. 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.

[9] J. Fox and S. Parsons. On using arguments for reasoning about actions and values. In *Proceedings of the AAAI Spring Symposium on Qualitative Preferences in Deliberation and Practical Reasoning, Stanford*, 1997.

[10] J. Hulstijn and L. van der Torre. Combining goal generation and planning in an argumentation framework. In A. Hunter and J. Lang, editors, *Proc. Workshop on Argument, Dialogue and Decision, at NMR*, Whistler, Canada, June 2004.

[11] J. Pollock. The logical foundations of goal-regression planning in autonomous agents. *Artificial Intelligence*, 106(2):267–334, 1998.

[12] A. S. Rao and M. P. Georgeff. Bdi agents: from theory to practice. In *Proceedings of the 1st International Conference on Multi Agent Systems*, pages 312–319, 1995.

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

[14] G. A. W. Vreeswijk. Abstract argumentation systems. *Artificial Intelligence*, 90:225–279, 1997.

[15] D. Walton. *Argument schemes for presumptive reasoning*, volume 29. Lawrence Erlbaum Associates, Mahwah, NJ, USA, 1996.

[16] M. J. Wooldridge. *Reasoning about rational agents*. MIT Press, Cambridge Massachusetts, London England, 2000.