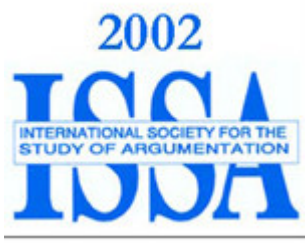
ISSA Proceedings 2002 – Strength And Order In Practical Reasoning: Decision-Guiding Argumentation



Abstract

Beliefs are the only evidence available for an agent making decisions about whether what he wants to do is justified under the circumstances or not. We think the connection between beliefs and goals can be evaluated according to order and strength criteria. Order among supporting reasons constrains the decision-guiding argumentation process to only those decisions that are relevant for the agent while just excluding or postponing the others. Strength determines the expected degree of utility derived from the adoption or non-adoption of a goal. An agent would only be justified in adopting a goal when the reason that supports it remains undefeated.

Introduction

Practical reasoning seems to help the agent in the way of constructing strategies and plans in his pursuit of a better situation for himself. The goals and objectives of an agent can be of diverse nature, from mere intrinsic desires to sub-goals of already intended plans. For instance, to be thirsty is usually a reason for adopting the goal of finding water or some other refreshment to quench one's thirst. Similarly, the obligation of starting work at nine o'clock every morning can be a sufficient reason for adopting the goal of getting up at half past seven daily. Other goals just respond to exigencies arising from intended plans (e.g. getting a ticket for the Symphony Hall can be just a sub-goal of my intended plan for the weekend). One of the tasks of practical reasoning is to cope with *conflict* situations of decision-making among an agent's potential goals. To be sure, sometimes an agent is forced to choose among different relevant options that are jointly incompatible.

Our approach assumes that, though not always, in many cases, the adoption of goals is plan dependent. Generally it happens that a goal cannot be adopted before the agent realizes that he is able to bring a plan about for the occasion. Often an important amount of the value of a goal is directly obtained from the expected utility value of the plan in which it is embedded (Beaudoin 1994). In

more detail, the adoption of a goal would be related to three factors: the value of the goal itself, the possibility of constructing a plan pursuing a previously learnt strategy for that goal, and the agent's commitments to previous plans (Pérez Miranda 1997). Hence to justify the adoption of a goal the rational agent must be able to construct a solid *decision-guiding argument.* That is, the practical argument constructed for the occasion should remain undefeated after the reasoning process. 'Practical reasoning is based on an agent's goals relative to a situation and on his knowledge of what is usually (reasonably expected) to obtain, according to his knowledge of the situation. …Typically, this pragmatic type of reasoning is based on rules or regularities that admit exceptions. Hence the conclusion is based on a kind of plausible reasoning – it represents a type of provisional presumption that could be subject to rejection or revision in the face of the new evidence, or of new developments in the situation' (Walton 1990: 84).
Once the agent has recognized that a potential goal is obtainable, the next step in determining the adoption of a goal is to detect any incompatibilities between that goal and other possible intended goals derived from previous intended plans or single *urgencies* that ought to be accomplished without delay. Hence the agent must look for scenarios in which both potential goals and ongoing adopted goals fit together insofar as fulfilling one may be at odds with fulfilling another or with maximum fulfilment of the overall set. We are concerned with explaining how an agent could manage to make these factors fit together appropriately by adopting a behaviour consisting in, so to speak, following some *rational* patterns.
The evaluative mechanism proposed here is only concerned with those goals that have a motivational or cognitive grounding (see below). According to our model, the rational agent selects only those goals whose supporting reasons are undefeated according to the agent's doxastic states. The mechanism embodies two levels of decision-making depending on the order and strength of the supporting reasons. An agent only would be justified in adopting a goal when the reason that supports that goal remains undefeated.

1. Assessment of goals

During a process of deliberation an agent might have to face a conflict among reasons for adopting goals. For instance between a goal the agent desires and one which he ought to carry out because it is valuable morally. Often conflicting reasons affecting a practical resolution may be comparable so that one of them overrides or defeats the others. But on other occasions we can regard reasons as guaranteed by different values or interests which are commensurable though not

through a precisely definable ranking like in decision theory, but following other patterns probably based on experience:
'Why should practical reasoning give the sole role to values of pleasure, avoidance of frustration, or the maximization of coherence in one's life? Why should it not give equal weight to friendship, loyalty, magnanimity, justice, and so on? And above all, why should it be dominated by one value, and deny the independence force of all the others?' (Raz 1999: 52).

So not all conflicts of reasons can be reduced to calculating the relative strength or force of each of them in order to determine which one defeats the other. Concerning this last point Sloman (1990: 235) remarks that:
'Some comparators apply constraint goals in planning, for instance using a 'minimize cost' rule to select the cheaper of two subgoals. Others directly order ends, like a rule that saving life is always more important than any other goal, but not because of some common measure applicable to both. As there are different incommensurable sources of motivation and different bases of comparison, there need not be any optimal resolution of a conflict'.
At this level then what is needed is a way of making *comparisons* among alternative potential goals that enables us to select only those goals that it would be reasonable for the agent to adopt according to his conative dispositions and doxastic states. According to Raz (1975: 35), 'conflicts of reasons for actions can be of many types and are one of the most intricate and complex areas of practical discourse'.

1.1 Conclusive reasons and prima facie reasons
One way of tackling the problem of conflicts among reasons for adopting goals is to pay attention to the different types of reasons an agent can have for adopting a goal. What kinds of reasons do we use when adopting a goal then? To start we can make a distinction between conclusive and prima facie reasons:
Def. 1: A *reason* is an ordered pair áG, *p*ñ, where G is a finite set of premises and p is the conclusion.
Def. 2: *Conclusive reasons* are reasons that logically entail their conclusions.
Def. 3: Prima facie reasons are reasons which create a presumption in favour of a conclusion that is 'defeasible' (Pollock 1991).

In other words, the conclusion is supported but not entailed by the set of premises that captures the evidence the agent has about the situation. We extend the use of these definitions to the case of practical reasoning, where the set G can contain

premises that are mere motivations supporting a conclusion about a goal. For instance, the desire of eating an ice cream can be a prima facie reason for adopting the goal of buying one at the store. In practice, though a certain fact could at first glance be thought a conclusive reason for adopting a goal, sometimes if we add another fact the resulting more complex task is no longer a reason for its adoption. When it is the case we could say that the added fact defeats the reason.

A reason can be a defeater of another reason in two different ways:
Def. 4: Defeaters that attack the conclusion derived from a prima facie reason are *rebutting defeaters*. Formally, if áG, *p*ñ is a prima facie reason, áL, *q*ñ is a rebutting defeater for áG, *p*ñ if and only if áL, qñ is a reason and $q = \neg p$.
Def. 5: Defeaters that attack a prima facie reason without attacking its conclusion are named *undercutting defeaters*. Formally, if áG, *p*ñ is a prima facie reason, áL, qñ is an undercutting defeater for áG, *p*ñ if and only if áL, *q*ñ is a reason and q = ¬(PG»*p*), where PG represents the conjunction of the members of a finite set of premises G and P»Q is a conditional that means that P wouldn't be true unless Q were true (Pollock 1991).

Thus, undercutting defeaters accomplish this by instead attacking the connection between the premises and the conclusion. Suppose that proposition «x is desired by agent A» is a prima facie reason for «A's adoption of x as a goal», but in practice the agent knows that «his desire of achieving such a goal is unrelated to the agent's already intended plans». The last fact can be interpreted as a defeater that attacks the connection between the agent's desire and the agent's adoption of a goal without directly attacking the conclusion. However, since these definitions overlook the importance of the strength of defeaters, in practice they are ineffective for our proposal of computing the defeasibility status of potential goals.

2. Strength and order
Beliefs are the only evidence available to an agent making decisions about whether what he wants to do is justified under the circumstances or not. We think this connection between beliefs (or motivations) and goals can be encoded into an ordered pair, the reason supporting the goal, and be evaluated according to order and strength criteria.
The strength of reasons is one of the decisive factors in determining which goal ought to be adopted. Reasons differ in strength. Some reasons supporting goals

are better than others. Our task is to clarify how those strengths affect interactions between reasons. The decision-theoretical model assumes that an agent with a variety of goals is capable of comparing the satisfaction of these goals so as to come to an overall assessment. Rational choice theory assumes that preferences (or desires) can be ordered on a single scale by comparing the 'utilities' of satisfying them (Jeffrey 1983, Hargreaves *et al.* 1992). From this view it would be possible to select between two competing goals in terms of the expected utility value associated to the achievement of each of those goals. Strength would determine the expected degree of utility derived from adopting or failing to adopt a goal at a certain time given the evidence available.
Even if a potential goal is motivated by an intrinsic desire that is *well-grounded* (in the experience of the relevant states of affairs or the beliefs related to them), its adoption might still produce conflicts with other goals within the agent's overall system of desires, beliefs and interests. For instance, if an agent is persuaded that it is not feasible to satisfy his intrinsic desires, he will generally tend to convert them into wishes, but obviously not into goals for the occasion (Green 1992). We defend the view that *primary grounding* does not warrant goal adoption as seems to happen within instrumentalism. The desirability characteristics of Bach's music can be a reason for listening to Bach this afternoon in the garden, but certainly they can not be a sufficient reason for doing so. Other more encouraging reasons for doing other things can compete and override the former reason. The agent has to evaluate the consequences derived from adopting specific goals while excluding or postponing others. The desideratum in a process of goal adoption should be the fulfilment of the agent's overall interests. In other words, the objective of practical reasoning would then be to render the agent's situation as 'likable' as possible paying attention to both long-term and short-term goals.

Another important aspect generally overlooked is the *order* of the reasons supporting alternative goals. The order among supporting reasons constrains the decision process to only those decisions that are relevant for the agent while just excluding or postponing the others. In particular, high order reasons override low order reasons, ruling them out of the process of assessment. Furthermore, ordering reasons is a way of facing situations of apparent incomparability, for instance, among supporting reasons that are desires and reasons that are beliefs. In this sense, the paper attempts to give a coherent description of why reasons would be of different order. It is not always possible to compare reasons on the

basis of numerical values associated with them, as is done in decision theory. Sometimes, the way of 'measuring' the value of a reason for adopting goals is not quantitative, but qualitative in nature. Suppose that my reason (which arises from my desire) for adopting goal A is stronger than my reason for adopting goal B. At this stage, both A and B are only potential goals. In principle, if I had to make a choice between them, I would select A. Nevertheless, the very influence of competing reasons on goal adoption is not usually determined by their relation to desires, but rather by their relation to the agent's beliefs or to plans already intended by him. That is, apparently conflicting reasons are not only disregarded but also overridden by high order reasons and there is no need to compare them in a quantitative way. Take the case of the tragedy of Antigone. Antigone believes that she is in circumstances in which, legally, she ought not to bury her brother Polyneices but religiously, she ought to do so. Suppose that she believes that the decrees of the gods override the laws of kings. In such a case, she ought to bury her brother. Both supporting reasons have a cognitive grounding as they are based on beliefs; in so far as they are reasons of different order, the conflict is resolved according to the order assigned by the agent to each reason. In this case the agent would not need to calculate the expected utility value associated to each of the goals. On the contrary, the agent could justify her selection saying that she is following a moral or religious norm that defeats any other kind of reason under consideration. There are many everyday situations where agents have good reasons to respect individual rights, cooperate or stick to some collectively accepted rules or norms (Nida-Rümelin 1997). In these cases agents refrain from weighing reasons instrumentally.

2.1 Exclusionary and high order reasons

In decision theory the *strength* of an alternative goal is represented by a quantitative value or vector, but it is not always necessary for an agent to make use of such a theory when merely noting the consequences of not satisfying a goal is enough to indicate its importance (Beaudoin 1994). No further deliberative reasoning would be required due to the importance of the achieving of that goal for the agent. The agent can believe that the resultant scenario derived from the achievement of the goal is important enough to stop the process of deliberation at that stage and go ahead with that goal. In that case the goal in question would have high priority over any other potential goals at that particular time. Prima facie reasons that correspond to those situations are named exclusionary reasons. When an agent holds any undefeated *exclusionary* reason he is justified in not

adopting any other goals on the balance of possibly previous reasons (Raz 1978). In such a case we regard the content of the agent's mental state not as a reason for adopting a goal, but probably as a reason for disregarding any other reasons for adopting a different goal.

Def. 6: An *exclusionary* reason is a high order reason to refrain from adopting a goal for some reason.
Def. 7: A *high order* reason is any reason to adopt a goal for a reason or to refrain from adopting a goal for some reason.

Exclusionary reasons are of utility in an important range of cases involved in practical deliberation (e.g. a *promise* can be understood as a case of high order reason for adopting a goal that in its turn is exclusionary). An agent's desire to invest in impressionist painting, *p*, could be a prima facie reason for «flying to Paris next week-end and contacting a merchant», goal A. Nevertheless, if we take into account, say, «the agent's promise of spending next week-end by the beach with his family», *q*, then the connection between the former prima facie reason *p* and its related goal A would be defeated by the last fact *q*, which behaves as an exclusionary reason in such circumstances, C. According to the previous scenario, a promise could behave as an *effective* undercutting defeater for any other lower order reasons. It doesn't mean, however, that the defeated prima facie reason couldn't be readopted under other circumstances.
Notice that if *p* is a prima facie reason for *x* to adopt goal A in circumstances C and *q* is an exclusionary reason for him not to adopt that goal on the basis of *p*, then *p* and *q* are not strictly conflicting reasons, because *q* is not a reason for adopting not-A in C. It is a reason for not adopting A in C for the reason that the conflict between p and *q* is a conflict between a low order and a high order reason. In other words, *q* works as an undercutting defeater because it doesn't directly attack conclusion A, but the connection between premise *p* and its associated conclusion A.
An exclusionary reason, of course, may also conflict with and be overridden by another high order reason. Only an undefeated exclusionary reason succeeds in excluding. Imagine that *p* is a prima facie reason for adopting goal A in circumstances C', but *p* is overridden by *q*, a second order reason, which emphasizes the necessity of pursuing B. Nevertheless, *q* is in its turn overridden by another second order reason, *r*, which induces the agent to believe that «B ought not to be adopted». In such situations we need to pay attention to the force

or strength of our second order reasons as we do for conflicts between first order reasons (e.g. following the patterns of decision theory).

High level reasons may preclude an agent from acting on other reasons of lower level. As in our previous example, promises usually are of higher order rather than other kinds of reasons. This doesn't mean, however, that promises couldn't be defeated in their turn by other reasons. We can easily imagine situations where promises are, for instance, cancelled. Suppose that the example is slightly modified, and we also assume that the agent knows that the prices of the paintings will be very low, then it might be sufficient for the subject to change his earlier plans (i.e. break the promise, or negotiate the agreed-on plan), and adopt the new goal. In that case, the new epistemic situation can be a reason not only to cancel the promise, but also to adopt the new goal. In other similar situations to keep the promise may be impossible; or the reason I had to do the promised act can be cancelled. For instance, the person whom I made the promise may release me from it.

The general picture can be summarized as follows:
On the one hand, conflicts of same order reasons are resolved by the relative weight or strength of the conflicting reasons, which determines which of them overrides the other. For instance, in decision theory the relative weight of the conflicting reasons can be calculated using the notion of expected utility value.
On the other hand, possible *conflicts* between lower and higher order reasons are then resolved according to other criteria:
1. importance, urgency or necessity of goals considered (e.g. saving life is always more important than any other goal the agent could be performing at a particular time);
2. whether the goal is part of an intended plan or not (e.g. whether I'm committed to wear sports clothes when playing table tennis; since I plan to play today, I will wear sports clothes for the occasion. I also disregard any other reason for wearing other types of clothing then);
3. level of subordination of goals within the intended plan (goal hierarchy). This specially affects the temporal order for adopting goals (e.g. in my travel to Paris the goal of getting a train ticket is a sub-goal of the main goal of arriving at that city).
So, in this model the resolution of conflicts between reasons located at different levels, just as the resolution of reasons situated at the same level, is described in

terms of one reason prevailing over, or overriding, or being stronger than the other**(i)**.

3. Means-end Reasoning and Potential Goals

The level of competing reasons for adopting goals depends to a large extent on whether the goals in question are embedded in intended plans or not. If an agent intends a plan, then he probably holds a high order and undefeated exclusionary reason supporting the adoption of goals within that plan. That reason would play two different but complementary roles: first, it would justify the agent's adoption of those goals as part of the overall plan; and second, it would mean the at least temporarily rejection of any other ongoing or incompatible goals sustained by lower or same order reasons. All other things being equal, no further reasoning would be necessary.

It has to be emphasized that many conflicts among alternative goals are resolved not by the strength or force of the competing reasons but by design constraints that determine, for instance, that exclusionary reasons always prevail, when in conflict with lower order reasons. The order of reasons determines the importance of competing goals and usually set up a, sometimes temporal, partial preference order among them (see below).

It is enough to keep in mind that reasons for adopting goals are very varied. If we follow the patterns of decision theory, then we have to deal with the problem of evaluating the agent's relevant alternative goals in terms of a quantitative function. Nevertheless, it could be inadequate to make a choice among goals only taking into account the predictions or projections of their consequences and their utility value, without considering the agent's already selected plans (or adopted goals). Intended plans provide a clear and concrete purpose for means-end reasoning and narrow the scope of deliberation to a limited set of options (Bratman *et al.* 1988; Audi 1991). An agent's intended plans provide a background framework within which deliberation should be performed.

Goals do not necessarily always originate in agent's motivations, desires or urgent needs, since some of them arise from planning or means-end reasoning. In this sense, incoming goals have to be related back to the agent's previously adopted higher-level goals. Hence, both the hierarchy and importance of goals become usually planning dependent. The point is that the evaluation of the expected utility value associated with an action would probably vary according to whether or not that action is considered in connection to a running adopted plan.

Therefore, one of the problems about the decision-theoretic model is that the evaluation of the expected utility value of an action is done *locally* (Pérez Miranda 1997). Obviously, an urgent enough goal can interrupt any plan's execution and cause re-planning in the system towards the achievement of that goal. These preservation goals are supported by high order reasons when deliberating about what to do next. An important issue is to resolve problems of conflict between already adopted goals and potential goals that become of importance as new information is obtained from the world by the agent.

Practical reasoning helps the agent to put himself in a more 'desirable' situation (than otherwise) as he performs actions aimed at satisfying adopted goals. The gent's intended plans drive means-end reasoning. They provide constraints on what potential goals must be considered in the process of deliberation, and they condition the beliefs for further practical reasoning. In our opinion, the degree of *adequacy* of a goal for adoption is conditioned by two factors: the order and strength of the reasons that support that goal. These reasons can respond to many aspects of the agent's mental life (e.g. accepted rules, duties, obligations, desires, sub-goals of intended plans and so forth) that lead him to adopt goals to change the current situation more to his liking.

For instance, having adopted a rule what an agent has to decide is whether to act on it in a particular case. What the agent is not doing is assessing the merits of the case taking all relevant facts into consideration. The agent is not doing this because he has chosen on a rule, that is, he has accepted an exclusionary reason, to guide his behaviour in such cases. He may, of course, occasionally examine the justification of the rule itself. Having rules allows for decisions to be made and revised in conditions other than the occasion for action itself.

4. A Single Filtering Mechanism for Goal Adoption

Our aim here is to offer a way of evaluating the status of an agent's potential goals (options) in order to know whether they should be adopted or not according to several adequacy criteria. We have already seen that reasons for goal adoption in planning are defeasible prima facie reasons. All exclusionary reasons are in fact prima facie reasons, but the opposite doesn't hold. Changes in the agent's environment may lead to changes in his beliefs, which in turn may result in his considering new options that can be incompatible with the agent's intended plans; in particular, goals that must be satisfied owing to their importance or urgency without necessity of further deliberation. Since this kind of goals must be adopted as soon as required, they are generally supported by undefeated higher

order prima facie reasons that at the same time are exclusionary.
Once a potential goal has been proposed, either by means-end reasoning or by an opportunity recognizer, it is subject to filtering. Our system resolves the filtering problem evaluating the defeasible status of the reasons which support those potential goals.

Def. 8: A prima facie reason *supports* a potential goal if and only if there is a motivational or cognitive grounding**(ii)** chain which serves as a link between the reason and the goal.
That link embodies a type of inference which proceeds from a set of premises taken as evidence to a conclusion (a potential goal or suitable option) supported by but not entailed by that evidence. The evidence can be just a motivational grounding reason (e.g. the agent's desire to obtain a goal) or a very sophisticated cognitive grounding reason (e.g. the agent's belief that the potential goal in question is a situation-type whose expected outcome is a scenario with high expected utility value).
Although supporting reasons can arise from very different sources, we assume that the agent is able to compare among them in two different ways: (I) ordering the reasons according to their importance and urgency; and (II) assigning numerical values for computing their strength when necessary.
Def. 9: An agent is *justified* in adopting a goal at a particular time t if and only if the prima facie reason that supports it is an undefeated reason then.
In such a situation the agent is quite sure that there is no reason to think that he should be justified in adopting any other goal or none at all.
Let us now define defeat among prima facie reasons then:
Def. 10: A prima facie reason z is defeated at a particular time t by another prima facie reason y, if and only if, at that particular time t, either (1) y is an exclusionary reason for z; or (2) n ≥ k, where n is the strength assigned to y and k is the strength assigned to z, y and z being conflicting reasons of the same order**(iii)**.

It seems that the type of defeaters involved in goal adoption are to a large extent undercutting defeaters. They don't directly attack the goal associated with the reason, but the link between them (see above). As we have seen exclusionary reasons actually proceed as undercutting defeaters of high order. An agent that holds an undefeated exclusionary reason is justified in not adopting any other goals on the balance of possibly previous reasons. By definition an exclusionary

reason can prevent the adoption of a goal without necessarily supporting (the adoption of) any other goal itself. It is necessary, however, to have high order reasons for adopting goals (e.g. the desire to avoid pain, hunger or suffering can be examples of such reasons). Accordingly, these reasons, while excluding the adoption of any other option at a point time, would serve as supporting reasons for, say, high priority goals as well.

It often happens that a reason is *temporarily* defeated, only to be reinstantiated afterwards. For instance, when an urgent goal needs to be accomplished an intended plan might be postponed provisionally, until after the goal has been satisfied.

Def. 11: A prima facie reason is *temporarily defeated* if and only if that reason is defeated until a particular time t and for every t' subsequent to t, t'»t, there is not any defeater for that reason.

Def. 12: A prima facie reason is *definitively defeated* if and only if that reason is defeated at a particular time t and for every t' subsequent to t, t'»t, the reason does not result undefeated.

It would also be reasonable to adopt new goals that don't delay considerably the achievement of other more important goals (e.g. plan's overloading)**(iv)**. In such a case the reason supporting the current goal would be temporarily postponed until the satisfaction of the overloading goal.

Def. 13: A goal is *temporarily postponed* if and only if its supporting reason is temporarily defeated.

Def. 14: A goal is *abandoned* if and only if its supporting reason is definitively defeated or self-defeated.

Undefeated supporting reasons direct the agent to the adoption of goals. The presence of an effective defeater for a reason supporting a goal entails that that goal cannot be adopted until the reason is, if it indeed is, undefeated according to the new epistemic situation.

Def. 15: If a goal is satisfied, then the supporting reason becomes a defeater for itself and is excluded from the evaluative process. In other words, the reason is *self-defeated*. Consequently that reason can't be a defeater for other reasons.

4.1 Description of the filtering mechanism

If the reason supporting goal1 (G1) is of higher order (O) than the reason supporting goal2 (G2) at t, or, both being reasons of the same order, the strength (S) of the former reason is equal or higher than the strength of the latter one, then an agent is justified in adopting a goal1 over another goal2 at a point time t.

If O(Reason1 áevidence, G1ñ, t) > O(Reason2 áevidence, G2ñ, t) or, both being reasons of the same order,
S(Reason1áevidence, G1ñ, t) ≥ S(Reason2 áevidence, G2ñ, t), then Adopt (Agent, G1, G2, t)

The filtering mechanism chooses only those goals whose reasons are undefeated by looking, first, at the order of reasons, and second, if necessary, at their strength**(v)**. It is not always that case that as a result of computing the status of available reasons a goal is ready for adoption. It can just so happen that the only resultant justified reason does not support any goal at all (e.g. a plain exclusionary reason).
The model's assumption is that goal adoption is a selection task, where an agent is presented with a set of goals and one or some has to be selected. Furthermore, by posing goal adoption as a selection task, goal and planning tweaking (adaptation) is within the scope of the model. For example, an agent facing the above dilemma (see section two) might renegotiate his promise to the family, and they might decide to go to Paris together instead of going to the beach. It is quite clear how the proposed model would produce similar behaviour. The filtering mechanism would generate a new higher order goal that makes it possible to unify previously intended goals but in an easier way than otherwise. So, if the reason supporting the new goal is of higher order than the rest of relevant reasons available, then the agent would be justified in adopting the new goal. Indeed, this is not but a special case of overloading.

Conclusions
We emphasize the necessity of incorporating into our models three distinctive factors: time, strength and order. Both doxastic states and conative dispositions are conditioned by time insofar as beliefs and motivations change along the agent's life. The evaluative mechanism proposed only concerns those goals that have a motivational or cognitive grounding (or both at the same time). Beliefs are the only evidence available to an agent making decisions about whether what he wants to do is justified under the circumstances or not. We think this connection between beliefs (or motivations) and goals can be encoded into an ordered pair, the reason supporting the goal, and be evaluated according to order and strength criteria. Order among supporting reasons constrains the decision process to only those decisions that are relevant for the agent while just excluding or postponing the others. In particular, high order reasons override low order reasons, ruling

them out of the process of assessment.
Furthermore, ordering reasons is a way of facing situations of apparent incomparability, for instance, among supporting reasons that are desires and reasons that are beliefs. Strength determines the expected degree of utility derived from adopting or not adopting a goal at a particular time given the evidence available. Although this is still an open question for ongoing research this value would in principle be calculated following the main principles of decision theory.

Acknowledgments
This work has been supported by the Research Project 9/UPV 00003.230-13907/2001.

NOTES
i. An interesting point that we don't pursue here is that one reason is not always stronger than or overrides the other reason. A more plausible model is necessary for dealing with such a situation. The model does not deal with issues of resource requirements for pursuing a goal (i.e., when two goals might be supported by equally strong reasons, and they might not be contradictory, yet pursuing both would not be possible due to resource limitations).
ii. Some people defend the view that situation-likings provide the ultimate starting point for deliberation in practical reasoning and that those are not representational states, but something closer to feelings or emotions than to propositional attitudes (Simon 1967, Green 1992, Pollock 1995). Other cognitive scientists recognise the importance of 'motivators' (transient or long term) in producing, modifying or selecting among actions, considering beliefs (Sloman 1990, Wright 1994). In our opinion, these sources could be generically understood as mental states that provide either motivational grounding or cognitive grounding for the agent's potential goals. By motivational grounding we refer to mental states which, though generators of goals, are not however based on the agent's beliefs (e.g. an agent can desire to listen to Bach this afternoon in virtue of experiencing the desirability characteristics of such listening, independently of whether the agent has or does not have beliefs to the effect that his listening to Bach has these qualities). By contrast, when the relevant beliefs are reasons for potential goals, we speak of cognitive grounding (e.g. an agent can believe that it is appropriate to listen to Bach's baroque music because it is a worthwhile or pleasant experience).

iii. Notice that the reason supporting a goal could be an 'argument'. In that case, if one of the lines of such an argument is attacked by an undefeated reason at a point time t, it would also be unjustified for the agent to adopt the goal in question then (Pollock 1995).

iv. Sometimes the overloading of a plan – introducing new courses of action compatible with it – is possible in order to satisfy new goals previously unforeseen while the agent continues performing his prior plan (Wilensky 1980, Pollack 1992).

v. Sometimes connections between inputs and outputs are learned responses or innate reflexes. These sorts of connections are called condition-action rules (productions or if-then rules) (Russell and Norving 1995). Since condition-action rules, when fired, are to be applied without necessity of further deliberation, they can be conventionally interpreted as exclusionary reasons of higher order than the rest of reasons available at a point time.

REFERENCES

Audi, R. (1991). Intention, Cognitive Commitment, and Planning. *Synthese* 86, 361-378.

Bratman, M.E., D.J. Israel & M.E. Pollack (1988). Plans and Resource-Bounded Practical Reasoning. *Computational Intelligence*, 4 (4), 349-355.

Beaudoin, L. P. (1994). *Goal Processing in Autonomous Agents*. Birmingham: PhD thesis, School of Computer Science, The University of Birmingham.

Green, O. H. (1992). *The Emotions*. Dordrecht: Kluwer.

Hargreaves, S., M. Hollis, B. Lyons, R. Sugden & A. Weale (1992). *The Theory of Choice*. Oxford: Blackwell.

Jeffrey, R. (1983). *The Logic of Decision*. Chicago: University of Chicago Press.

Nida-Rümelin, J. (1997). Why Consequentialism Fails. In: G. Holmström Hintikka & R. Tuomela(Eds.), *Contemporary Action Theory* (pp. 295-308, Vol. II). Dordrecht: Kluwer Academic Publishers.

Nozick, R. (1993). *The Nature of Rationality*. Princeton: Princeton University Press.

Pérez Miranda, L.A. (1997). Deciding, Planning, and Practical Reasoning: Elements Towards a Cognitive Architecture. *Argumentation 11*, 435-461.

Pollack, E. M. (1992). The Uses of Plans. *Artificial Intelligence* 57, 43-68.

Pollock, J. (1991). A Theory of Defeasible Reasoning. *International Journal of Intelligent Systems* 6, 33-54.

Pollock, J. (1995). *Cognitive Carpentry: a Blueprint for How to Build a Person*.

Cambridge, MA: The MIT Press.
Raz, J. (1975). *Practical Reason and Norms*. London: Hutchinson and Co.
Raz, J. (1978). Reasons for Actions, Decisions, and Norms. In: J. Raz (Ed.), *Practical Reasoning* (pp. 128-143), Oxford: Oxford University Press.
Raz, J. (1999). *Engaging Reason. On the Theory of Value and Action*. Oxford: Oxford University Press.
Russell, S. & P. Norving (1995). *Artificial Intelligence*. New Jersey: Prentice Hall.
Simon, H. A. (1967). Motivational and Emotional Controls of Cognition. In: *Models of Thought*, Yale: Yale University Press.
Sloman, A. (1990). Motives mechanisms and emotions. In: M.A. Boden (Ed.), *The Philoshophy of Artificial Intelligence* (pp. 231-247), Oxford: Oxford University Press.
von Wright, G. H. (1971). Explanation and Understanding. Cornell: Cornell University Press.
Walton, D. N. (1990). *Practical Reasoning. Goal-Driven, Knowledge-Based, Action-Guiding Argumentation.* Maryland: Rowman and Littlefield Publishers.
Wilensky, R. (1980). Meta-Planning: Representing and using knowledge about planning in problem solving and natural language understanding. No. UCB/ERL/M80/33, Berkeley: Electronic Research Laboratory.