Decision Theoretic Planning and the Bounded Rationality of BDI Agents

Guido Boella
Dipartimento di Informatica - Università di Torino
Cso Svizzera 185 Torino ITALY - email: guido@di.unito.it

Abstract

We reconsider the BDI agents model by exploiting a *decision theoretic planning* framework together with a proposal for bounded rationality. In particular, we show how the interplay between a planner based on an *anytime algorithm* and a *meta-deliberation* module for dealing with bounded resources sheds light on the notion of *intention* as proposed by Michael Bratman.

Introduction to the agent architecture

Studies about bounded rationality have flourished in the last years and have shed light on the meaning itself of Artificial Intelligence: an intelligent program should not (or even cannot) provide the optimal solution, but the best solution net of the computational costs ((Russell & Wefald 1991), (Zilberstein & Russell 1993), (Boddy & Dean 1994), (Wooldridge & Schut 2001)). As (Doyle 1987) expresses this aim: "the agent reflects on its circumstances, abilities, and limitations to rationally guide its own reasoning and internal organization [...]. Rational self-government includes direct control of the path and amount of reasoning performed".

It is now possible to design "systems that are capable of taking their own computational resources into consideration during planning and problem solving" ((Boddy & Dean 1994), p.245).

In particular, the role of *meta-reasoning* has been formalized: it "consists of running a decision procedure whose purpose is to determine what other decision procedures [e.g., planners] should run and when" (*ibid.*, p.248). The time used by the decision procedure is important since, in *real time* problem situations "the utility of a given action varies significantly over the time necessary for a complete solution of the decision problem" (Russell & Wefald 1991).

In parallel, the agent paradigm has provided a unifying point of view over many branches of AI. Some of the agent models have received particular attention from the formal point of view: in particular, the BDI agent model, a 'cognitive' model of *agenthood* which assumes that agents, as well as humans, are driven by their mental representations; the leading mental attitudes are beliefs, desires and intentions. To face the resource boundedness problem, (Bratman 1987) proposes to use the currently intended plans to limit the set of actions to be considered.

Surprisingly enough, there is almost no connection between the formal models of bounded rationality and the BDI agents models. Since the notion of intention has been introduced by Bratman to deal with the resource boundedness of agents, a reason must be found for the missing of this link.

As an exception to the distance existing between formal studies of bounded rationality and the BDI paradigm, consider the position of (Wooldridge & Schut 2001). They see the BDI agents model "not as another time-dependent planning model [i.e., a bounded rationality model], but rather as a model in which time-dependent planning is useful to incorporate." However, for (Wooldridge & Schut 2001), the role of time-dependent planning is to provide "an implementation for intention reconsideration" (*ibid.*, p.20) and not a general conceptual instrument for analyzing the BDI model.

As (Horvitz & Zilberstein 2001) note "elusive principles of intelligence might actually be founded in developing a deeper understanding of how systems might grapple with scarce, varying or uncertain time and memory resources". For this reason, we proceed in (Wooldridge & Schut 2001)'s direction, reconciling both fields, so to consider the bounded rationality model not only a component of the BDI model, rather, as an explanatory concept for understanding it.

Our approach is grounded by the use of a *decision theoretic planning paradigm* (DTP). However, introducing decision theoretic planning is not *per se* sufficient, nor it is a feasible choice: decision theory has been proven too complex for dealing with real time agents.

Here comes into play the bounded rationality model. However, most works on rationally bounded agents focus on particular tasks like game playing, mail sorting, path discovery, and even on general programming. On the other hand, the BDI model refers only to planning and in particular hierarchically structured planning, since it corresponds to how humans reason.

The new conceptual instruments resulting from the marriage between DTP and bounded rationality will be used to reconsider Bratman's analysis of the concept of intention and of the related BDI agents model. In fact, Bratman's view of intentions is strictly related to that of planning: "Our commonsense conception of intention is inextricably tied to the phenomena of plans and planning" ((Bratman 1987), p.30).

Commitment has been proposed by Bratman as the main property which distinguishes intentions from other attitudes such as goals. At the time of Bratman's work, the end of the '80s, the state of the art planning technique, which he necessarily refers to, was hierarchical decomposition planning.

In contrast, decision theoretic planning allows finding a plan which is optimal with respect to the degree of achievement of the goal and to resource consumption. In particular, some DTP planners (Ha & Haddawy 1996) use a hierarchical organization of plans, a requirement posed by Bratman: further, *abstract* actions in the DTP plan hierarchy are based on a sound theory of abstraction, and provide approximate estimates of the costs and benefits of the primitive plans they subsume. Finally, hierarchical DTP is implemented by *anytime algorithms* (Zilberstein & Russell 1993): the quality of the solutions provided depends on the time allocated to the processing, and the algorithms exhibit a time/quality trade-off defined by a *performance profile* (Boddy & Dean 1994).

In particular, we analyze the concept of *intention* in terms of *desires* and *beliefs*, without introducing a primitive notion of *commitment*. This form of reduction is rejected by Bratman after a tentative analysis of the definition of intentions as *predominant desires*: "intentions are distinctive states of mind on a par with desires and beliefs, [...] I propose to consider this network of dispositions and functional relations on its own terms, without trying somehow to reduce it to ordinary desires and beliefs" ((Bratman 1987), p.20). Bratman argues that desires, even *predominant* ones, do not imply that the agent settled on a decision. In our model, this argument is countered by the fact that devising a different solution with respect to the currently intended plan is blocked by the *meta-reasoning* component of the agent since it has to deal with the bounded resources of the agent.

Bratman did not have the instruments to consider the trade off between the cost of planning and the benefit produced by refining a solution.
Costs are necessary for producing *information*: which desire is the predominant one and the possibility of satisfying it can be discovered only after a planning phase: the knowledge produced by this planning phase can be exploited later by the agent to improve his performance.

Our view of planning as a source of information is supported also by (Zilberstein 1996): "the goal of planning is to provide the execution architecture with information that can improve the selection of actions" (p.33).

Some distinctions must be drawn also for what concerns desires and goals. Desires are modeled as preferences: they are expressed by a (multi-attribute) utility function[1] which guides the search process of the planner (see (Ha & Haddawy 1996)). Apart from the obedience to utility theory axioms, desires can be still irrational as prescribed by BDI models: an agent can prefer states which cannot be achieved by any action.

Goals in (Cohen & Levesque 1990) are defined as worlds chosen among the desired ones. Goals are usually proposed as the 'consistent' counterpart of desires: however, no clue is provided on how agents choose worlds from their desires. As in the classical planning tradition, for us, goals are just inputs to the planner.

Before the goals undergo the (even if partial) planning process, there is no guarantee at all that they will be chosen. Nor that they are consistent among each other. The consistency requirement for goals, posed by (Cohen & Levesque 1990), is costly and is satisfied by the fact that during the planning process, a goal is associated with a plan, which, if the planner is sound, will be consistent and will achieve its goal.

In contrast with Bratman's conception, even before the goals become intentions, an agent has to examine the means for achieving them.

Second, goals are not only linked to preferences. Obviously, some goals stem from the preferences of an agent, while other ones have different sources. As (Castelfranchi 1998) has noticed, goals can be *adopted* from other agents' ones when those agents want us to achieve some goals. In particular, we have considered two other sources of goals: cooperative settings (Boella, Damiano, & Lesmo 2000) and norms (Boella & Lesmo 2001).

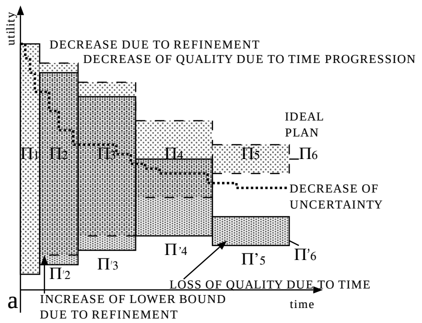
In our framework, the choice of the decision theoretic planner proposed by (Haddawy & Hanks 1998) is not only important for its ability to find optimal plans. The hierarchical organization of plans in DTP presents another fundamental advantage: abstract actions can be associated with approximate evaluations of the costs and benefits of executing them by means of the primitive plans they subsume.
More formally, (Ha & Haddawy 1996) define the semantic of an *abstract* action of the specification hierarchy in this way:[2] "an abstract action **A** is a function **A**: $S \rightarrow 2^S$ [where S is the set of states expressed as probability distributions on the values of the attributes describing the world] that maps a probability distribution $P \in S$ to a set of probability distributions $\{A(P) \mid A \in \mathbf{A}\}$ [since an action may have uncertain outcomes, its effects are expressed by a set of states without a probability distribution on them]".

Hence, given an initial world (i.e., a probability distribution) and an abstract action A, the set of outcomes resulting from its projection subsumes all the outcomes of all the possible primitive plans Pr(A) deriving from the refinement of that action.

The expected utility of a set of states whose probability distribution is not known is *uncertain* (in the technical sense of the term, see (Keeney & Raiffa 1976)): the relative probability of the outcomes is still not known, since none of the subsumed primitive plans has been already examined and chosen. For this reason, the expected utility is expressed as

[1] Multi-attribute utility functions require a number of assumptions on the structure of preferences of agents, as shown in (Haddawy & Hanks 1998) or in classical works as (Keeney & Raiffa 1976).

[2] Following (Ha & Haddawy 1996), we distinguish two kinds of abstract actions: *sequentially abstract* actions, which associate a complex operator with a decomposition formed by a sequence of steps. And *abstract* actions in the proper sense, which subsume more specific (possibly still abstract) actions. When there is no risk of misunderstanding, we will speak generically of abstract actions meaning both kinds of hierarchical relations.



a

Figure 1: Refinement under time boundedness and without it.

an *interval*, whose upper bound is the expected utility of the best outcome of the best subsumed action, and analogously for the lower bound: all the other (uncertain) outcomes are included within this interval; the subsumed actions will have a narrower utility interval.

The planning process is a form of search in the plan space: it starts from a set of candidate alternative abstract plans (plans containing at least one abstract action, *partial* plans, in Bratman's terms) and proceeds refining all the plans.

At each step of refinement, a plan where an abstract action occurs is substituted with the plans where the abstract action is replaced with all the subsumed less abstract actions or, in case of *sequential* abstraction, with a sequence of steps which constitutes the action decomposition. The new actions may be still abstract ones, which must be in turn refined. Even if the search space is large, the hierarchical organization of the plans allows *pruning* a large part of it: in fact, if the best outcome of an abstract action A is worse than the lower bound of another action B of the current set of alternatives, it is worthless to examine the actions A subsumes; so A is discarded without further refinement.

The emergence of intentions

First of all, we apply the resource bounded paradigm to our BDI agent setting. Due to the resource boundedness of the agent, it is not always possible to complete the planning process and to choose the best plan available to him. The meta-reasoning component has to stop the *anytime* planner when refining further the current plans seems a losing strategy: a partial, possibly non optimal, plan becomes the intention of the agent.

At each refinement step, the set of candidate partial plans Π_i increases as an effect of substitutions with less abstract plans and decreases as a result of the pruning of suboptimal plans.[3] Moreover, at each step the length of the candidate plans increases, since the decompositions of (sequentially) abstract actions are inserted into the plans. In turn, the duration of the projection process increases with the length of the plans, so, at each step, the time for processing the plans becomes longer.

In Figure 1, we sketch the problem in a graphical way. Six refinement steps are depicted in light grey: at each step a rectangle represents the set of current candidate plans Π_i (their utility intervals all overlap, so that the upper side of the rectangle is the utility of the plan with the best upper bound and the lower side represents the worse lower utility among all plans). The horizontal dimension of the rectangle represents the time needed for the refinement, which increases proportionally to the number and length of the plans.

At each step of refinement,[4] the upper side of the rectangle may decrease (subsumed plans give a more precise - i.e., narrow - estimate), and the lower one may proceed in the opposite direction (some suboptimal plans may be discarded): at each step, the quality of the results increases in the sense that a more precise (i.e., less uncertain) estimate is provided (see the category *certainty* of (Zilberstein & Russell 1996)'s classification) and some suboptimal plans are discarded. The subtree-independence assumption of (Russell & Wefald 1991) is satisfied since the change in the utility regards the nodes of the search tree in an independent manner.

At each step, as time goes by, the utility of the single plans decreases as an effect of two factors. First, if a plan is executed in a state with less resources than another state (and planning is a resource consuming task), the plan can consume more resources than if executed from the previous one (e.g., you have to drive quicker to be at destination). Second, (after the first moment the goal provides the agent with some utility) the utility which a plan has for the agent is discounted according to the time it takes to reach the goal. The same plan, before and after a refinement step, has a lower utility, since it is executed from a worse (and later) initial state.

If there were no discount related with time, the utility interval of the set of candidate plans would proceed as shown by the light grey rectangles. However, in reality, the refinement proceeds as shown by the dark rectangles: the intervals still become more narrow, while the utility decreases. At a certain time (see Π'_5), some plans become worse than at the previous step (e.g., some good plan is not executable anymore in the time remaining).

The effect is that, after a certain amount of time, it is better to stop refining and let the execution component to choose a plan without any other planning effort, even if there is no guarantee of choosing an optimal plan; the *meta-reasoning module*, knowing or estimating the parameters involved in this process, can predict in advance when to stop planning and to accept the risk of choosing a suboptimal plan from the set of primitive plans subsumed by the partial plans built so far.

[3]Note that even if the number of candidate plans increases, the number of primitive plans subsumed by them remains the same or decreases as an effect of pruning.

[4]For the sake of brevity, here, we identify a step of refinement with the expansion of a level of depth in the search space. The actual algorithm is interruptible after the substitution of a single abstract action of a single candidate plan. So the real reduction of the uncertainty of the plans proceeds as in the dotted line of the Figure 1.

The parameters considered by the meta-reasoning process are:

1. The average narrowing factor $n(i)$ of the utility interval as a function of the number of refinement steps (we assume it is the same for upper and lower bound for simplicity). It is a positive factor which decreases with time since the number of candidate plans increases. Therefore, the uncertainty of the plans decreases monotonically, and also (Boddy & Dean 1994)'s requirement of diminishing returns is satisfied.
2. The average increase d in the duration of a refinement step.
3. The decrease of utility as a function of time (after some point after which the goal becomes useful for the agent): the function c (assume time costs separability (Russell & Wefald 1991)).
4. Π_i, the set of plans (after the pruning of suboptimal ones) at step i.
5. a is the number of refinement steps necessary to reach primitive plans (depending on the height of the trees of actions in the plan library rooted by elements of Π_1). This is the upper bound on the duration of the refinement.
6. The functions u and l, the upper and lower bounds of the utility of a set of plans Π_i.
7. The increase, due to external events, of the uncertainty in the world, as a function of time (this factor is reflected in the state S_i, where the plans Π_i are executed after i refinement steps).

Obviously, the values of these parameters can be different according to the problem at issue and can be adapted according to the agent's experience (see (Boddy & Dean 1994)): a *performance profile* can be created by gathering statistics. As suggested by (Zilberstein & Russell 1993), it is possible to parameterize also with respect to the actual utility of the ith set of candidate plans, instead of relying on an approximation with respect to the initial utility.

Set b to:[5]

(1) b = min($j \leq i \leq a$) such that α*u($\Pi_i(S_i)$) + (1 - α)*l($\Pi_i(S_i)$) $>$ α*u($\Pi_{i+1}(S_{i+1})$) + (1 - α)*l(Π_{i+1}(S_{i+1}))

(j is 1 if the agent has not started refining the original plans, otherwise it is equal to the number of steps performed). u($\Pi_i(S_i)$) (and respectively l($\Pi_i(S_i)$)), if c and d are constants, can be approximated as a function of u($\Pi_1(S_1)$):

u($\Pi_i(S_i)$) = (u($\Pi_1(S_1)$)-($n(i) * i$))-$c(d * (i * (i+1)/2))$

If $b < a$, then the agent will have to choose one abstract candidate plan among the alternative ones in Π_b using the formula:

p^{best}=argmax$_{p \in \Pi_b}$(α*u($p(S_b)$) + (1 - α)*l($p(S_b)$))

and, on the basis of the information offered by p^{best}, the execution module of the agent architecture will choose randomly a primitive plan among the ones p^{best} subsumes[6] (it is guaranteed that those plans have a utility interval which is included in the interval of the subsuming plan p^{best}). This process does not ensure that an optimal plan is chosen, but providing the decision process with more time would have led to a possibly suboptimal choice, anyway.

The solution proposed, however, must be still refined a bit.

First, the quality of a plan depends not only on the expected utility interval but also on a parameter α which expresses the attitude towards uncertainty.

Why does the agent prefer a (more) certain outcome with respect to adopting an uncertain better option? As recent developments in decision theory have found, an agent proceeds refining his current plans because humans prefer less ambiguous alternatives with respect to more ambiguous and uncertain ones (see the Ellsberg's paradox). In order to compare the utility intervals with different degrees of uncertainty, it is necessary to scale the upper and lower bounds with respect to the number of possible primitive plans Pr(Π_i) whose respective probabilities are not known. In fact, the uncertainty decreases at each step (note that the primitive plans subsumed decreases:

Pr(Π_{i+1})$\subseteq$Pr(Π_i)).

If α is a constant, the formula for deciding between uncertain alternatives is reduced to Hurwitcz pessimism/optimism criterion.

Decision theory scholars have not yet agreed on a single technique for dealing with uncertainty. We adopt here (Ghirardato & Marinacci 2001)'s proposal to weight the possible outcomes according to a non increasing function α expressing the averse attitude of the decision maker towards the uncertainty of a set of alternatives Π_i:

α(Π_i)*u($\Pi_i(S_i)$) + (1-α(Π_i))*l($\Pi_i(S_i)$).

Uncertainty aversion can sometime lead an agent to suboptimal solutions, but humans adopt this attitude since, in the spirit of the so called *rule utilitarianism*, it is rational for them from the point of view of general policies for acting, even if, in some single cases, it is not the optimal policy.

Second, depending only on the parameters n, d and α, it seems possible that an agent could prefer to choose a plan from a more uncertain set of plans Π_i: in fact, the upper bound of the utility intervals decreases at each step so the refined solution seems to lose some good opportunities. Abstracting away from the problem of resource boundedness for a moment, we would like that the agent prefers always to choose an action from a more certain set of options. A tentative modification could be to impose a constraint on the relative values of n, d and α.

Instead, there is no need to do so. We are simply misinterpreting the role of the refinement process. We have said that at each refinement step, the set of candidate plans results in a narrower utility interval: the upper bound decreases and the lower bound increases. The decrement and the increment correspond to two partially different phenomena.

[5] As in (Horvitz 2001) we assume that if an uncertain solution is preferred to a refined one, further refinements would not change this decision.

[6] Other decision strategies are possible, see for example (Helwig & Haddawy 1996).

The upper bound decreases as an effect of the better estimate of the resulting outcomes, which is offered by the less abstract plans produced in the refinement step. So the upper bound of Π_{i+1} expresses the real (upper) utility of the same primitive plans included in Pr(Π_i): no optimal plan is discarded, while their utility estimate has been revised. The same happens for the lower bound: the estimates are revised since an abstract plan has been replaced by a set of more primitive plans.

However, the refinement process involves also a *pruning* step which effectively discards some of the plans: the suboptimal ones. In this case, we have that the lower bound increases as a result of the fact that some of the plans are thrown away (without any loss for the agent).

This means that the formula above must be changed: the comparison between the utility interval of a set of plans Π_i and the utility of the following one Π_{i+1} must be made with respect to the same upper bound: the one of the subsequent set of plans Π_{i+1}, minus time costs. It is an estimate, but the upper bound of Π_i is not certain anyway. Thus, formula (1) becomes:

(2) b = min($1 \leq i \leq a$) such that $\alpha(\Pi_i)$*u($\Pi_i(S_i)$-n(i)) + (1 - $\alpha(\Pi_i)$)*l($\Pi_i(S_i)$) $>$ $\alpha(\Pi_{i+1})$*u($\Pi_{i+1}(S_{i+1})$) + (1 - $\alpha(\Pi_{i+1})$)*l(Π_{i+1}(S_{i+1}))

In fact, u(Π_i(S_i)) is only a wrong estimate of u(Π_{i+1}(S_1)): the best plan of both sets is the same one. On the other hand, the lower bound is not (only) a wrong estimate, rather, it is the result of a greater set of alternatives. If the agent stopped and chose one primitive plan randomly from Π_i, he could select the worst plan in Pr(Π_i). But this plan may be a different (and also worse) one with respect to the worse case of Pr(Π_{i+1}), where, in fact, some plan may have been thrown away.

Stability vs opportunities

The main issue of Bratman's definition of intention is the *commitment* of the agent, i.e., its stability over time: "we simply are not capable of constantly redetermining without inordinate costs, what would be the best thing to do in the present, given an updated assessment of the likelihoods [...] Rather, we settle in advance on prior, partial plans and tend to *reconsider* them only when faced with a problem" ((Bratman 1987), p.29).

In (Boella 2002) we addressed the question of why agents commit to partial plans and they use these plans for guiding further planning. Here we focus on the stability of intentions.

Our framework suggests a solution to the question why an agent maintains an intention, and, in particular, why he maintains it even if he does not know whether it is still the most useful alternative for him.

The meta-reasoning process, when the agent has to decide whether to *reconsider* the intention, has to take into account the fact that he is partway the execution of his plans, while a new plan would need to be executed from scratch (possibly facing more risks) from a possibly worse initial state with respect to the current one.

These factors surely increase the persistence of the chosen plan with respect to the other alternative candidate plans which the agent has produced during the previous planning phase, before choosing and executing the currently intended plan.

What happens when the world changes? When the world changes, the agent does not know whether the change provides him with some new (better) opportunity. Some planning is needed. When it is rational that the meta-reasoning module starts reconsidering the intention anyway?

According to Bratman, intentions resist reconsideration since reconsidering them would take too much time with respect to the limited resources at disposal of the agent. Reconsideration is costly, if we interpret it in the sense of making some object level planning. In a classical planning framework, the meta-deliberation process cannot be but a full deliberation phase, where the agent replans from his prior goals. In contrast, in a hierarchical decision theoretic framework, the planner can be used to make some low cost, even if uncertain, predictions: in fact, even the abstract actions at the top of the plan hierarchy provide an estimate of the utility they have for the agent.

The problem of the stability of intentions is reduced to the problem of whether to refine the uncertain estimates of abstract plans or to continue the execution of the current plan.

If we exploit the bounded rationality principle, we can limit dynamically the reconsideration if it is too costly, without barring it by assuming commitment as a primitive principle. Also (Bratman, Israel, & Pollack 1988) consider the possibility of an opportunity analyzer which overrule the intention maintenance principle. In an experimental setting (Kinny & Georgeff 1991) notice that the boldness and cautiousness of agents in reconsidering intentions should be tailored to the dynamicity of the environment. Finally, (Wooldridge & Schut 2001) propose a framework inspired to (Russell & Wefald 1991) for dealing with intention reconsideration policies. Differently from our proposal, they base the decision of reconsidering the current intention or not on an estimate of the utility of (re-)deliberating "based on distributions which determine how the environment changes".

In contrast, in our model, we propose that agent always starts redeliberating, but we apply the resource boundedness model in order to limit this redeliberation so that the agent does not spend too much time reconsidering his intentions. Moreover, we try to show why the redeliberation results in a more limited reasoning with respect to a complete deliberation which starts from prior goals.

In Figure 2 we represent the situation where an agent has a plan p_1 and, at t_1 he decides, given a new assessment of the situation at hand, to meta-deliberate about whether to change his intention or not to (the grey boxes in the background represent three possible subsequent phases of reconsideration). At t_2 the agent has some new elements to consider: the abstract plans Π_2 and Π'_2 with their uncertain outcomes.

Compare the current situation with that of Figure 1: in that case, for each goal, the agent has only partial plans at his disposal. In this case, the agent already committed to one de-

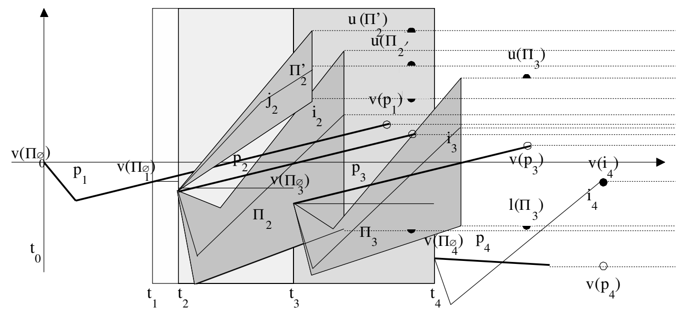


Figure 2: Deciding whether to abandon the current plan p_1 or not.

tailed plan p_1 (in fact, he started executing it) and projected some abstract actions Π_2 and Π'_2. In that case, after some refinement steps, the agent stops the planning phase and commits to a (possible suboptimal) primitive plan, since he can predict that going on with the refinement process would lead him to a situation where a more certain solution would be dispreferred with respect to a more partial solution.

Here the agent has to decide whether to continue the reconsideration process with the new information produced at t_2: he has at his disposal the projection of his current intention v($p_1(t_2)$) (t_2 is the time at which the first step of the meta-deliberation process ends) and some uncertain alternatives. Is it necessary a further deliberation step? If there is a plan such as Π'_2, which dominates p_2, the agent should revise his intentions and abandon p_2. Otherwise, if there is no such plan, consider the other alternative provided by the partial plans Π_2. Since it is partial, its utility is an interval [l($\Pi_2(t_2)$), u($\Pi_2(t_2)$)]. The interval must be compared with the expected utility value of $p_1(t_2)$. Some more refinement would be desirable, but the time employed in the decision leads to choosing among a worse set of alternatives.

The refinement mechanism works well if all the plans contain uncertainty: in fact, only for those plans there is an advantage in devoting time to planning. In contrast, for non-partial plans, as the current plan p_2 is, refining the remaining plans results only in a loss of utility due to time costs (compare u($p_1(t_0)$), u($p_2(t_2)$) and u($p_3(t_3)$)). Moreover, a primitive plan without uncertainty is not affected by the risk aversion of the agent.

As a consequence, formula (2) leads to an early stop of the refinement process, i.e. to maintaining the commitment to the current plan, even if the agent has to take a decision from a more uncertain set of alternatives.

Remind that when in Figure 1 we stop refining the candidate plans is not because we do not have a criterium for deciding among partial (i.e. uncertain) plans. Rather we proceeds refining the partial plans since we gather more information about the solutions to have more precise estimates about the plans outcomes and to discharge suboptimal plans from the set of possible solutions. The wider the utility intervals involved in the decision, the greater the possibility that a bad suboptimal plan is chosen due to the resource boundedness of the agent.

From the performance profile of the planner it is possible to predict whether the current estimate v($p_2(t_2)$) is more promising than a partial plan Π_n which is the product of the refinement of Π_2.

In summary, in a traditional planning framework, at time t_2 it is not possible to have an evaluation of the partial plans devised so far: a plan is comparable only when refined down to its primitive actions. But this refinement is too costly, so the definition of a primitive notion of *commitment* becomes necessary.

Some problems have still to be solved. (Castelfranchi & Conte 1997) have noticed a situation where linking decision theory to intentionality leads to an undesired behavior: if the agent is repeatedly presented with a new better alternative while he is performing his plan, he would choose the new opportunities every time. E.g., while he is executing p_1 at t_2, he commits to a new plan q_2, and while executing q_2, at t_3, he revises his intention in favor of the better plan r_3, etc. Such a lucky person would not ever reach any goal!
Our reduction of intentions in decision theoretic terms is subject to the same criticism. We, however, advance some considerations about how to face this anomaly.

According to normative theories of economics, given a better opportunity, the agent should abandon the previous intention, even if the agent has invested resources on it (the "sunk costs" problem). So the consideration of how much the agent spent so far is not a reasonable factor to include in the decision of non reconsidering intentions.

Escalating commitment in the presence of "sunk costs", however, is not always irrational, as (Camerer & Weber 1998) notice: "For example, the film 'The titanic' had an initial budget of $150m and was expected to gross, say $200m. [...] Now partway through, the studio has actually spent $200m and thinks it will cost $85m to finish. Assuming that the half-filmed movie is worthless, should the studio 'escalate'? The answer is 'Yes'."

Agents' plans are usually more similar to this case than to the examples used in the literature about sunk costs: plans have a duration and costs, and cannot be reduced to one-shot decisions as accepting a bet or selling actions: when resources are consumed the agent gets no immediate benefit, but, at the same time, these resource consumption (e.g., fuel consumption) forward the agent towards the end of the plan, from which the benefit is produced. As an example consider Figure 2 where, when at t_1 the intention is reconsid-

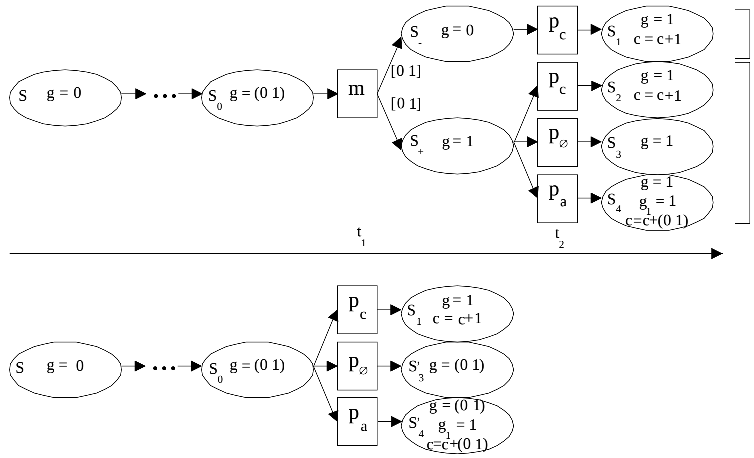


Figure 3: The agent compares a conditional plan preceded by a sensing action M with the alternatives at his disposal in the current state S_0.

ered, the agent has already spent some resources in achieving p_1, while he has not gained almost anything from the execution of the plan p_1.

If he changes his intentions before achieving anything in return for the resources spent so far, he will re-deliberate always in a state which is worse that the one where he took the last redeliberation. Hence at each redeliberation, for deciding to revise his intention he has to find an action which is much better than the first intention he committed to. Before deciding to reconsider his current action the agent should see also at this factor: how much the current state is worse than the one he could have reached by staying committed to previous plans.

After a first revision (that is, a new opportunity has been searched for and chosen despite the sunk costs plus future costs), an even greater opportunity is needed to make the agent change his idea. Hence, the commitment after a sequence of revisions increases, as a matter of fact and not because of subjective considerations.

What we have said in this Section concerns situations where the agent knows how the world has changed, while he has still to compute whether the new state of the world offers him some more opportunities. But agents' life is not always so simple. In the next Section, we will address the problem of reconsidering intentions in situations where the agent is not aware of which is the real situation and he has also to decide whether to gather new information about how the world changed.

Monitoring intentions

Some more words must be devoted to the question: when to abandon intentions? "By making explicit the conditions under which an agent drops his goals, i.e., by specifying how the agent is committed to his goals, the formalism captures a number of important properties of intentions" (Cohen & Levesque 1990). As (Cohen & Levesque 1990) prescribe in their definition, the intention must be released when it has been already achieved or it is impossible or irrelevant.

As it is obvious, in a DTP framework all these conditions can be reduced to the utility of a plan: in case the goal is achieved, the plan is impossible to execute or it does not contribute to the goals it was relevant for, going on in the execution of the plan would be just a waste of resources without any return in utility.

However, also this more general condition in terms of utility is subjected to the criticism of (Singh 1992): it is tautological to say that an intention is maintained as long as it is useful. Moreover, from this generalization we do not gain anything at all: in a realistic situation we cannot assume that the agent has at his disposal any information regarding the fact that the plan succeeded, is still executable or relevant. The real problem is how much resources devoting to the monitoring of the conditions which allow saying that a goal has been reached or impossible to achieve.

Fortunately, in a DTP setting, as (Boutilier 2000) showed, it is possible to compute the utility of monitoring such conditions. The costs and benefits of monitoring those conditions (and acting accordingly to the resulting outcomes) must be traded off against the benefits and costs of going on in the execution with possibly wrong beliefs about the world.

Information guide the agent, but they are costly in the same way as planning. We have to *meta-deliberate* about whether to gather such information or to continue with the estimate we have about the current state.

Let's focus only on the case of the satisfaction of the goal: it is possible that some event or agent have made the goal already true before the execution or the completion of the current plan.

There are however, two different situations to take into account. The simplest one is where the agent believes that the goal *may* have been already achieved. In the second one the agent has no clue at all about how the current situation has changed. The former case can be represented by an agent who has at his disposal a probability distribution on events that change the state as a function of time. In this case the decision about what to do can be summarized in the following way:

1. for each possible state of affairs in S_0, project the sensing action M and from the resulting situations:

   - if the goal has not been already achieved, project the current plan p_c.

   - else, project the abstract actions p_a for alternative goals[7] and for doing nothing $p_\emptyset$.

   - (in principle, you could project also the current plan, but its goal has been achieved; anyway, it may have an indirect utility from some goal it is relevant to).

2. compute the resulting utility of the outcomes of p_c, $p_\emptyset$ and p_a and combine it with the probability of the different initial states and the probability of the different outcomes of M.

This is similar to what happens in (Boutilier 2000), with the difference that we do not assume to have at our disposal without any cost the best alternative plan for dealing with failures; rather, we exploit the estimate provided by abstract actions.

However, this is not the only possible case. The agent does not have always at his disposal the exact probability distribution on the current state of affairs.

This situation of uncertainty is represented in the DTP model of (Ha & Haddawy 1996) by states which are composed of a set of probability distributions.
The representation of uncertainty as a set of probability distributions is directly related to (Shafer 1976)'s Theory of Evidence:[8] an uncertain world is not represented as a world where all values of an attribute have an equal probability; rather, it is represented by the so-called Basic Probability Assignment (BPA), where the value 1 (the total mass of probability) is associated with the universe of possible values.

For example, in Figure 3, in state S_0, the goal g is uncertain between the two possible values 0 and 1 (i.e., true and false). However, this is a particular distribution of BPA, which, in general, could associate a probability value to any subset of the universe of possible outcomes, thus representing a situation of *partial ignorance* about the a priori probability of the different situations. Since what is of interest is just the utility of the outcomes, it is possible to focus the attention on the worst and best cases, instead of considering all the possible outcomes. So, the utility of the different states is evaluated, and the situation of ignorance is represented as an interval whose lower bound is the state for which the utility is worst and the upper one is the state for which the utility is best.

The execution of a sensing action M modifies the agent's beliefs as a function of the external world state and not according to what the agent believes.

[7] Whether to refine further these actions is a problem which can be again solved by the meta-reasoning proposed above.

[8] Representing uncertain beliefs is a far from trivial problem. Here we simply adopt the framework proposed by (Ha & Haddawy 1996) in relation with the foundation of DTP. See (Walley 1991) for a comprehensive survey on the problem.

The decision about whether to execute the sensing action will be done by projecting the abstract plans from the states resulting (in the agent's mind) from the projection of the sensing action M. But even if the projection is made starting from an uncertain state S_0 the resulting states $S_1, \ldots, S_4$ will not be uncertain in themselves; rather, the agent is uncertain about which state he is in.

In Figure 3, we depict the alternative of sensing the world and then proceeding accordingly to the outcomes of the sensing action, and we compare it with the alternative of deciding on the basis of the beliefs represented by state S_0, where the agent does not know whether the goal g has been satisfied. Note that even after the sensing action M, the relative probability of S_- and S_+ is not known (it is expressed by the interval of probabilities [0 1]). While the two states are not uncertain in themselves.

However, after the sensing action M the agent can devise a conditional plan: the current plan p_c is executed in case the goal is still not satisfied, else, the other alternatives p_a and $p_\emptyset$ are taken into account in S_+. The utility of the conditional plan is bounded by the best (and worse) outcomes in all the branches of the conditional plan:

[min(u(S_1), max(u(S_2),u(S_3),u(S_4))), max(u(S_1), max(u(S_2),u(S_3),u(S_4)))].

In contrast, in case the agent does not sense the world, he has to decide among three different alternatives which maintain the uncertainty in the respective outcomes: for example, in state S'_4 resulting from the alternative plan p_a, the agent does not know whether the goal has been reached.

By comparing the (uncertain) utilities of this two alternative solutions, the agent can decide whether to sense the world (and possibly to discard the current intention) or to continue without any new information.

Up to know, we have discussed the problem of reconsidering the whole plan which is the intention of the agent. In the next Section, we will have a brief look at the problem of revising only in a partial way the current intentions.

The preference for replanning

"Not only agents care whether their attempts succeed, but they are disposed to replan to achieve the intended effects if earlier attempts fail" ((Cohen & Levesque 1990), p.217).
A plan can be not executable anymore if a precondition of some future action in the plan does not hold anymore. If the agent re-evaluated the topmost actions for achieving his (original and new) goals Γ in the current state S_e, he would get a set of plans Π_1 whose expected utility is:

I($\Pi_1(S_1)$)=[l($\Pi_1(S_1)$),u($\Pi_1(S_1)$)]

where, usually u($\Pi_1(S_1)$) - l($\Pi_1(S_1)$) is large. Refining Π_1, on the other hand, requires time.
Since the current plan p_A is included in the refinements of Π_1, its utility I($p_A(S_1)$) is included in I($\Pi_1(S_1)$). If it is still the best alternative, it could be reached by the planner. But this would require too much time.

Given the limited amount of time at disposal, the strategy described above risks to end in a set of plans Π_b whose utility interval is:

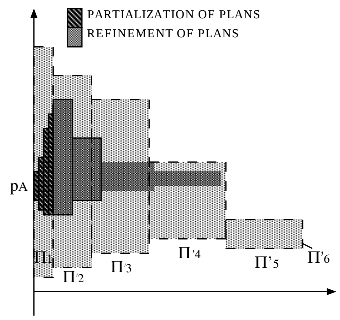


Figure 4: Partializing and then refining.

I($\Pi_b(S_b)$)=[l($\Pi_b(S_b)$),u($\Pi_b(S_b)$)].

The larger uncertainty of Π_b with respect to p_A does not worth the cost of replanning from zero.

But there is still some room for finding another solution: the current intention provides information which can become the starting point of the replanning. The idea is that the current solution is closer to a non suboptimal plan in the plan space: it takes less time to search for it than searching for an optimal one from scratch.

Replanning a solution from the current plan is an idea which has independent support. E.g., Adversarial planner (Elsaesser & MacMillan 1991) justifies a similar strategy with the motivation that it is better to save those parts of the plan which are still feasible and to *maintain momentum*, instead of planning from scratch. In our framework, we can justify this choice with considerations regarding resource boundedness, as in the cases discussed above.

The replanning process proceeds by producing more abstract plans from the current one in a sequence of partialization moves starting from the step of the plan which cannot be executed anymore. At a certain point, from the current partial plan, the refinement process is restarted and a primitive plan is searched for. Since the refinement, in general, restarts from plans composed of actions occurring in the middle of the abstraction hierarchy and not from the topmost actions, the search space is greatly reduced (the height of the tree leading to primitive plans is lower than the height of the trees rooted in the topmost actions). A lowly uncertain plan can be found in a quicker way.

An important property of the algorithm is that, if the nearest partial plan does not satisfy the goal the plan aims to, the replanner proceeds towards more and more partial plans until the topmost actions, where this strategy will converge with the former one.

In Figure 4, the replanning process is compared with planning a solution from scratch. The difference with Figure 1 is that before the first refinement step the current plan (in dark grey) is made more partial, until more promising solutions Π'_1 (in medium grey) are predictable (the partialization process requires less time than the refinement).

The replanning algorithm explains the persistence of the current intention without considering it as a reason for insisting on the current way of action: rather, intentions play again a role in an indirect way. Since intentions are predominant achievable desires and the predominance means that more resources have been allocated for determining if they are the predominant ones and achievable, the information produced must be exploited fruitfully in the decision process.

References

Boddy, M., and Dean, T. 1994. Deliberation scheduling from problem solving in time constrained environments. *Artificial Intelligence* 67:245–285.

Boella, G., and Lesmo, L. 2001. Deliberate normative agents. In Conte, R., and Dellarocas, C., eds., *Social order in MAS*. Kluwer.

Boella, G.; Damiano, R.; and Lesmo, L. 2000. Cooperation and group utility. In Jennings, N., and Lespérance, Y., eds., *Intelligent Agents VI — Proceedings of the Sixth International Workshop on Agent Theories, Architectures, and Languages (ATAL-99, Orlando FL)*, 319–333. Springer-Verlag, Berlin.

Boella, G. 2002. Intentions: choice first, commitment follows. In *Proc. of AAMAS Conference*.

Boutilier, C. 2000. Approximately optimal monitoring of plan preconditions. In *UAI-00*.

Bratman, M.; Israel, D.; and Pollack, M. 1988. Plans and resource-bounded practical reasoning. *Computational Intelligence* 4:349–355.

Bratman, M. E. 1987. *Intention, Plans, and Practical Reason*. Cambridge (MA): Harvard University Press.

Camerer, C., and Weber, R. 1998. The econometrics adn behavioral economics of escalation of commitment. *Social Science Working Paper* 1043.

Castelfranchi, C., and Conte, R. 1997. Limits of strategic rationality for agents and m-a systems. In *Proc. of 4th Modelage Workshop on Formal Models of Agents*, 59–70.

Castelfranchi, C. 1998. Modeling social action for AI agents. *Artificial Intelligence* 103:157–182.

Cohen, P., and Levesque, H. 1990. Intention is choice with commitment. *Artificial Intelligence* 42:213–261.

Doyle, J. 1987. *Artificial Intelligence and Rational Self-government*. CMU-CS-88-124. CMU Computer Science.

Elsaesser, C., and MacMillan, R. 1991. Representation and algorithms for multiagent adversarial planning. *Technical Report MITRE* 91W000207.

Ghirardato, P., and Marinacci, M. 2001. Risk, ambiguity, and the separation of utility and beliefs. *Mathematics of Operations Research*.

Ha, V., and Haddawy, P. 1996. Theoretical foundations for abstraction-based probabilistic planning. In *Proc. of 12th Conference on Uncertainty and Artificial Intelligence*. Portland.

Haddawy, P., and Hanks, S. 1998. Utility models for goal-directed, decision-theoretic planners. *Computational Intelligence* 14:392–429.

Helwig, J., and Haddawy, P. 1996. An abstraction-based approach to interleaving planning and execution in partially-observable domains. In *Plan Execution: Problems and Issues: Papers from the 1996 AAAI Fall Symposium*, 81–88. AAAI Press, Menlo Park, California.

Horvitz, E., and Zilberstein, S. 2001. Computational tradeoffs under bounded resources. *Artificial Intelligence* 126(1-2):1–4.

Horvitz, E. 2001. Principles and applications of continual computation. *Artificial Intelligence* 126(1-2).

Keeney, R. L., and Raiffa, H. 1976. *Decision with multiple Objectives: preferences and value tradeoff*. Cambridge (UK): Cambrige University Press.

Kinny, D., and Georgeff, M. 1991. Commitment and effectiveness of situated agents. In *Proc. 12th IJCAI*, 82–88.

Russell, S., and Wefald, E. 1991. Principles of metareasoning. *Artificial Intelligence* 49:361–395.

Shafer, G. 1976. *A Mathematical Theory of Evidence*. Princeton University Press.

Singh, M. 1992. A critical examination of the Cohen-Levesque theory of intentions. In *Proc. of 10th ECAI*, 364–368.

Walley, P. 1991. *Statistical reasoning with imprecise probabilities*. London: Chapman and Hall.

Wooldridge, M., and Schut, M. 2001. The control of reasoning in resource-bounded agents. *Knowledge Engineering Review*.

Zilberstein, S., and Russell, S. J. 1993. Anytime sensing, planning and action: A practical model for robot control. In *Proc. of the Int. Joint Conf. on Artificial Intelligence*, 1402–1407.

Zilberstein, S., and Russell, S. 1996. Optimal compositionof real-time systems. *Artificial Intelligence* 82:181–213.

Zilberstein, S. 1996. Resource-bounded sensing and planning in autonomous systems. *Autonomous Robots* 3:31–48.