Locally Global Planning:
Decision-Theoretic Planning in OSCAR

John L. Pollock

Department of Philosophy
University of Arizona
Tucson, Arizona 85721
(e-mail: pollock@arizona.edu)

Abstract

It is conjectured that MDP and POMDP planning will remain unfeasible for complex domains, so some form of "classical" decision-theoretic planning is sought. However, local plans cannot be properly compared in terms of their expected values, because those values will be affected by the other plans the agent has adopted. Plans must instead be merged into a single "master-plan", and new plans evaluated in terms of their contribution to the value of the master plan. To make both the construction and evaluation of plans feasible, it is proposed to evaluate plans and their interactions defeasibly.

1. Approaches to Decision-Theoretic Planning

Two assumptions made by classical planning are: (1) the agent has full knowledge about the consequences of its actions in any circumstances; and (2) all that is important about a plan is that it achieves its goal. For an applied planning problem pertaining to a narrowly circumscribed domain, we might be able to pretend that the first assumption holds, but not for planning in an autonomous agent operating in a complex and only partially predictable environment. This problem can be handled in part by reasoning defeasibly about the consequences of actions, as I described in my (1998), but that is at best a partial solution. Turning to the second assumption, it is generally recognized that in evaluating plans we have to weigh the value of the goal achieved against the cost of achieving it. That cost can consist of both normal execution costs and the costs (positive and negative) of possible side effects. The costs of side effects can include things like opportunity costs. Executing one plan can make it impossible (or merely more difficult or more expensive) to execute another plan. Furthermore, nothing is certain. We have to discount the value of the goal by the probability of achieving it, and discount execution costs similarly by their probabilities. In other words, we must evaluate plans decision-theoretically.

1.1 Local Plans

There are two general approaches to decision-theoretic planning. One is to construct plans with the same general structure as classical plans, but then to evaluate them in terms of their expected values, where the expected values are computed by assigning values to possible

outcomes, discounting the values by the probabilities of the outcomes actually occurring[1]. An important characteristic of the plans constructed in this way is that they are "local" plans. That is, they aim at restricted goals like transporting an object to a location, acquiring a certain bit of information, recharging one's battery, etc. As such, they concern a small slice of the world. This makes the construction of such plans manageable by goal regression techniques. The assumption is then that expected values can be computed for such local plans, and one plan is preferable to another iff it has a higher expected value. The aim of decision-theoretic planning is to produce optimal local plans.

This has been a popular approach, but as I pointed out in my (1992) and (1995), it represents too quick a generalization of classical decision theory to plans. Local plans cannot be compared and chosen for execution just by comparing their expected values. The basic difficulty is that classical decision theory concerned *acts*, which were taken to be unstructured entities that were logically independent of one another. Plans are not like that. Plans concern temporal slices of the world, and can aim at combinations of goals. As such, different plans can have different scopes. In particular, unlike the acts of classical decision theory, plans can embed one another as subplans, or they can have overlapping parts. This makes it impossible to compare them just by comparing their expected values. For example, if one plan embeds another as a subplan, its having a higher expected value than its subplan does not automatically make it a better plan. It may have a higher expected value just because it achieves additional goals. If there are other ways to achieve those additional goals, it might be better to adopt the subplan together with some other plan for the other goals rather than adopting the single plan that aims to achieve all the goals.

To make this difficulty more concrete, consider an example. One kind of case in which a single plan achieves multiple goals is a plan to run two errands on a single trip. A plan to run two errands on a single trip is often preferable to plans to run the errands separately, but it is not always preferable. Suppose I must pick up a ton of lead and a ton of gold from a single repository and deliver both to the same destination. Both will fit in my truck, and I could pick them up on a single trip, but doing so would risk damaging my springs. Then it might be better to deliver them on two separate trips. However, the plan to deliver them on a single trip, by virtue of achieving both goals, might have a higher expected value than any single plan with which it competes, e.g., the plan to deliver the gold without delivering the lead. What is better than adopting the plan to deliver them both on a single trip is adopting the two separate plans to deliver the gold on one trip and deliver the lead on another trip. So the plan with the higher expected value is not the best plan to choose. One should instead choose *two* other plans.

Of course, it is always possible to construct a fourth plan that merges the plans for the two trips into a single plan prescribing both trips, and that plan will have a higher expected value than the plan to deliver both the gold and lead on a single trip. But do we want to require agents to always consider all the ways of combining their plans into larger plans before they decide which plans to adopt? That is computationally problematic. It is also subject to a

---

[1] Some examples of this approach can be found in Haddawy and Hanks (1990), Haddawy, Doan and Goodwin (1995), Ngo et al (1998),Williamson and Hanks (1994). There is also work on probabilistic planning that is not decision-theoretic. See Draper et al (1994), Kushmerick et al (1995).

logical difficulty. If we are allowed to merge plans arbitrarily into larger and more inclusive plans, then in a complex world there may be no optimal plans. For every plan, it may be possible to construct a preferable plan by merging the first plan with other plans for achieving other goals. If decision-theoretic planning requires the adoption of optimal plans, it may have the consequence that no plans will be adopted.

1.2 Global Plans

The natural response to the preceding is to require plans to be "global" rather than local. We force them to be comparable by considering only plans that aim to achieve all goals simultaneously (insofar as that is possible). This is, in effect, the course taken by the other major strand in decision-theoretic planning—Markov decision planning.[2] Markov decision planners (MDP's) and partially observable Markov decision planners (POMDP's) proceed by building a state space whose nodes represent all possible states of the world. This is a graph in which nodes represent possible states of the world and links between nodes correspond to actions that would move the world from one state to another with some specified probability. It is assumed that we have a valuation function assigning a value to each node, and then the objective is to construct an *optimal policy*, which is in effect a global plan prescribing the best action to perform in each possible state. MDP's and POMDP's proceed by building the entire state space and then searching for an optimal path through it.

A generally recognized problem for Markov decision planning is that it is computationally infeasible in any but the simplest environments. States of the real world are characterized by a huge number of variables. To estimate the complexity of the real world, it has been estimated that there are 10^{78} elementary particles. If we take the state of a particle to be determined by four quantum states each having two possible values (a gross underestimation), each particle can be in 16 states, and so there are $16^{10^{78}}$ states of the universe.[3] This is a bigger number than we can write in the form $10^{..}$. The length of the exponent would be greater than the number of elementary particles in the universe. Clearly, an optimal policy cannot prescribe actions for all of these different states. It must abstract from the true complexity of the universe, making the assumption that most differences between states do not make any difference to how the agent should behave. Suppose we could confine our attention to just 300 two-valued variables. That is pretty unrealistic—it seems clear that many more than 300 parameters can make a difference to optimal behavior, and many of them are continuous-valued rather than two-valued. But even if we could confine our attention to 300 two-valued variables, an optimal policy would have to distinguish between 2^{300} states and prescribe behavior for each. 2^{300} is approximately equal to 10^{90}, which is twelve orders of magnitude larger than the number of elementary particles in the universe. Clearly, a real agent cannot deal with policies that large, and even such policies would be woefully inadequate because in some cases they would fail to make

---

[2] See particularly Barto et al (1995), Boutelier, Grafman and Geib (1997), Boutelier and Dearden (1994, 1995), Dean and Given (1997), Dean, Given and Kim (1998), Dean, Kaelbling, Kirman, and Nicholson (1993), Simmons and Koenig (1995), Tash and Russell (1994). A good summary can be found in Boutelier, Dean, and Hanks (1999).

[3] In fact, some of the paramters, like position, are continuous-valued, so there are really infinitely many possible states.

crucial distinctions.

Most ongoing research on MDP and POMDP planning is aimed at getting around this feasibility problem. The general approach is one of abstraction in which states are collected together on the basis of different criteria. But I am betting that this approach will never be successful for even moderately complex environments. There is a computational reason why human beings engage in local planning rather than global planning, and my surmise is that other autonomous agents operating in complex environments will have to do so as well.

Let me note in passing that there is another, purely logical, problem for MDP's and POMDP's. This is analogous to a problem that arose in connection with the decision-theoretic evaluation of local plans. It was observed that if we are allowed to merge plans arbitrarily into larger and more inclusive plans, there may be no optimal plans. The analogous problem for MDP's and POMDP's is that the value of a policy is only well defined for finite state spaces. There must be a "planning horizon"—a time beyond which we do not plan. But there is no non-arbitrary way to choose a planning horizon for a real agent. Presumably, for any time t, there is some nonzero probability that the agent will survive to that time and still be achieving goals and suffering execution costs, and if the values of those "late life" goals and costs are sufficiently great they can be decisive in the comparison of policies that extend to that time. The upshot is that the comparison of completely global policies in terms of their expected values may not even make sense.

2. A Compromise Proposal

MDP's and POMDP's get the logic of the planning problem right (except for the problem of the planning horizon), but they do so at the expense of being impractical for realistic planning in complex environments. Assigning values to classical plans seems like a more feasible alternative, but it does not get the logic right. This paper formulates a third approach, which can be regarded as a kind of compromise between POMDP planning and classical decision-theoretic planning. The approach should be logically correct but feasible. The key to the approach is to use classical planning techniques to produce plans (but base them upon probabilistic connections rather than exceptionless causal connections), and then reason defeasibly about the expected value, not of individual plans, but of the whole package of plans that the agent has adopted at any one time.

To motivate this approach, let me call attention to another problem for the decision-theoretic evaluation of local plans. This problem is distinct from the problem I considered earlier. This is a problem for even computing the expected values in terms of which local plans are supposed to be evaluated. Both the values of goals and the values of execution costs are typically a function of the circumstances under which a plan is executed, and that in turn will be strongly influenced by what other plans the agent adopts. To take a trivial example, if my goal is to eat a dish of vanilla ice cream, the value of that goal will be seriously diminished by my adopting a plan that calls for my eating a dill pickle first. And even more obviously, execution costs can be seriously affected by the agent's other plans. A plan to deliver a package will be much harder to execute if a prior plan first takes the agent to the other side of town. Most of the literature on decision-theoretic classical planning assumes that execution costs are constant values for each action in a plan, and goals have fixed values. But for an agent operating in a

realistic environment, those assumptions aren't even good approximations. To compute an expected value for a plan, we must know what other plans the agent has adopted, and the right response to the construction of a new plan may be to adopt it and withdraw an earlier plan (in response to the new plan interacting negatively with that earlier plan).

So there are two distinct problems for decision-theoretic classical planning. The first is that plans cannot be evaluated in isolation from one another, because they can affect each other's expected values. The second is that local plans cannot be chosen for adoption just because they have higher expected values than any competing plans the agent has constructed. Plans can have different scopes, aiming at different sets of goals, and comparing them in terms of their expected values can be like comparing apples and oranges. Sometimes, the appropriate comparison is between *sets* of plans rather than individual plans. In the earlier example of transporting the gold and lead, the appropriate comparison is between the single plan of transporting both on one trip and the set of two plans for transporting each on a separate trip. Of course, as remarked, the two plans can be merged into a fourth plan which can be compared with the plan to transport both the gold and lead on a single trip, and the fourth plan will have a higher expected value. But we don't want to require an agent to consider all possible ways of combining its local plans into larger plans. If the agent constructs n local plans, there will be at least 2^n ways of merging them into larger plans.

2.1 Master Plans

Although we do not want the agent to consider all the possible ways of merging its local plans into larger plans, we can consider the single plan that results from merging all of the agent's local plans into a *master plan*. I suggest that it is master plans that are the appropriate objects of decision-theoretic evaluation. The master plan represents the fruits of all of the agent's planning, and evaluating a plan in the context of the agent's other plans is just a matter of evaluating its contribution to the value of the master plan. As a first approximation, we can say that it is reasonable to adopt a local plan just in case adding it to the master plan will increase the value of the master plan. This is only an approximation however, for two reasons. First, adding a plan to the master plan may only increase the value of the master plan if we simultaneously delete other plans from the master plan. Second, plans may have to be added in groups rather than individually. Let us consider these two points separately.

There are several different reasons the addition of a plan to the master plan may necessitate deleting other plans. The simplest is that the new plan may be incompatible with another previously adopted plan, in the sense that executing the new plan will make it impossible to successfully execute the other plan. In the technical sense common to causal-link planning, a plan step *threatens* a causal-link if it *could* make its subgoal false between the time it is produced and the time it is used by a later step. A subplan *undermines* a causal link if it constitutes a plan for the negation of the subgoal between the time it is produced and the time it is used. So the way in which a new plan may make it impossible to successfully execute another plan is that the new plan may introduce steps into the master plan which jointly comprise a subplan that undermines the earlier plan. Then attempting to execute the earlier plan would still exact execution costs but would not achieve the goal of the plan. Hence the master plan will have a higher expected value if the earlier plan is deleted when the new plan is adopted.

The technical notion of a subplan undermining a plan is derived from non-decision-theoretic deterministic planning, where it constitutes the only way in which plans can destructively

interfere with each other.[4] However, in decision-theoretic non-deterministic planning, other kinds of destructive interference are possible as well. A generalization of undermining arises when a subplan lowers the probability that a causal link will work. A different kind of destructive interference consists of a subplan lowering the values of the goals of another plan, or increasing the disvalue of possible side-effects and execution costs.

Apparently, in deciding whether to adopt a plan the agent cannot just compare the expected-value of the master plan before and after adding the new plan to it. Simply adding the new plan may lower the expected-value of the master plan unless interfering plans are simultaneously removed. Notice that the interference can go both ways. The new plan may interfere with previously adopted plans, and previously adopted plans may interfere with the new plan. In either case, the best decision may be to adopt the new plan and retract the previously adopted plans.

A different kind of complication for the evaluation of plans in terms of their contributions to the value of the master plan is that it may be necessary to add plans in sets rather than individually. Recall the example of transporting the gold and lead. If the agent begins by adopting the plan to transport both on a single trip, and then considers the separate plans to transport the gold on one trip and the lead on another, adding either of the latter plans by itself (while deleting the plan for transporting both on a single trip) will lower the value of the master plan rather than raising it. The master plan will only receive a higher value if we add *both* of the new plans at the same time.

The upshot of this is that plan evaluation should always proceed at the level of the master plan. Local plans should be adopted or rejected because of their effect on the master plan. Furthermore, changes to the master plan will in general consist of simultaneously adding and deleting sets of local plans, not just individual local plans.

2.2 Constructing and Evaluating the Master Plan

The master plan becomes a data structure of supreme importance in decision-theoretic planning. We can think of it provisionally as a composite plan that results from merging into a single plan all of the local plans the agent has adopted. It can be regarded as the agent's current best approximation to a globally optimal policy. It is "as global" as the agent's current planning has been able to manage, but it makes no attempt to discriminate between all possible states of affairs, and as such is small enough not to overwhelm the agent's memory capacity.

Although the master plan is of manageable size from the perspective of information storage, it will still be a very large plan by the standards of current AI planning technology. It is not unreasonable to expect that at some given time a sophisticated autonomous agent will have adopted 1000 local plans with an average length of 10 steps each. That translates into a master plan of 10,000 steps. Furthermore, the local plans will typically deal with a very wide variety of goals and circumstances, and may draw collectively from 1000 different possible actions. The production of such a plan is several orders of magnitude beyond the capabilities of current automated planning algorithms. Weld (1999) observes that the current state of the art is represented by BLACKBOX (Kautz and Selman 1998), which can find a plan with 105 steps in

---

[4] The latter was proven for plans without conditional effects by McAllester and Rosenblitt (1991), and for plans with conditional effects by Penberthy and Weld (1992). I proved it for a more inclusive set of plans in my (1998).

a world with 10^{16} possible states. That is impressive when compared with previous planning technologies, but calculation reveals that $10^{16} = 2^{53}$, so this is still a world characterized by only 53 fluents. Computing an expected value for a 10,000 step plan is also an immense undertaking. It may be impossible to compute it exactly.

It is probably impossible to apply conventional plan construction and evaluation techniques to the master plan. Fortunately, that is not necessary. I represented the master plan as a composite of all the local plans the agent has adopted, and that description should be taken seriously. The master plan cannot be produced by combining all of the agent's goals into a single conjunctive goal and planning for that from scratch, but it can be produced by planning separately for the individual goals as they arise and then merging the resulting local plans into the composite master plan. When they are merged, the agent must be on the lookout for both destructive and constructive interference, but what makes the planning manageable is that it can be assumed defeasibly that there is no interference until some is found. It is suggestive that human planners seem to proceed in this way. When solving local planning problems we do not look continuously at the big picture. Rather, we find a plan that seems to work subject to local constraints, and then we worry later about how it fits in with the rest of our plans. If we don't see a problem, we assume there is none, although we remain vigilant in case a problem later emerges.

I suggest that plan evaluation can work similarly. It is really the master plan that we want to evaluate, but we can do that defeasibly by evaluating local plans. If the local plans are independent then the expected value of the master plan will be the sum of the expected values of the local plans it comprises. So the proposal is that the agent can assume defeasibly that the local plans it produces are independent, and evaluate the master plan accordingly. It can then look for failures of independence, and when they are found, the local plans entering into the dependence can be merged together and evaluated directly, and then it can be assumed defeasibly that that is the value contributed by that set of local plans to the master plan. If subsequent investigation turns up a larger set of dependencies, then the larger set can be merged and evaluated and it can again be assumed defeasibly that that evaluation constitutes the value contributed by that larger set of local plans. In this way the agent never has to evaluate the entire master plan directly. It just evaluates relatively small plans, either local plans or composites of several local plans, and then sums the resulting values to get a defeasible estimate of the value of the master plan.

The proposal is then that direct plan construction and plan evaluation be confined to local plans or small composites of local plans. The master plan is constructed by merging the local plans and it is evaluated defeasibly by summing the values of the small plans. This makes the computational task of constructing plans and evaluating changes to the master plan relatively simple. It does assume, however, that the agent has the computational tools required for detecting interference between the local plans, because it is the presence of such interference that defeats the defeasible assumption. Let us turn then to the topic of plan interference.

2.3 Decision-Theoretic Interference
If two plans are truly independent, they can be merged into a single plan and the expected value of that composite plan will be the sum of the expected values of the constituent plans. When the expected value of that composite plan is equal to the sum of the expected values of the constituent plans, let us say that they exhibit *decision-theoretic independence*. *Decision-theoretic interference* is the failure of decision-theoretic independence. The literature on classical deter-

ministic planning also recognizes failures of independence in its treatment of "threats" or "underminings". This is regarded as an obstacle to merging the plans at all. But it is important to recognize that this is just the limiting case of the failure of decision-theoretic independence. What is wrong with merging two plans when one undermines the other is that the undermined plan is then prevented from achieving its goal and hence from contributing its expected value to the expected value of the composite. So interference at the level of plan construction can be subsumed under decision-theoretic interference.

In adding local plans to the master plan, our defeasible assumption is one of decision-theoretic independence, so what is needed is tools for detecting decision-theoretic interference. The standard tools for detecting undermining will be a subspecies of these tools, but they must be generalized to handle the cases in which merging two plans raises or lowers the expected value of the composite rather than preventing one of the constituent plans from contributing anything at all to the value of the composite. For this purpose, let us survey the ways in which interference can arise.

Undermining arises when executing the steps of one plan prevents a causal link of the second plan from working. It does this by making the subgoal of the link false between the time it is produced and the time it is used. When the causal link records a merely probabilistic connection, a weaker variety of undermining can arise from the steps of the one plan lowering the probability of the subgoal being true when it is to be used. This can happen in two ways. (1) It could change the probability of the subgoal being made true in the first place, or (2) it could change the probability of its remaining true until the causal link target is executed. The latter is analogous to classical undermining, and includes it as the limiting case. The former has no non-probabilistic analogue, because in classical planning it is assumed that actions are *guaranteed* to have their effects, regardless of circumstances, providing only that the preconditions are satisfied.

The first kind of probabilistic undermining can be described more precisely as follows. The causal link is based upon a probability prob(*subgoal*/*action* & *C*), where *C* is a context established (or made probable) by the earlier parts of the plan. Undermining occurs when the second plan makes a context *C** probable and prob(*subgoal*/*action* & *C* & *C**) ≠ prob(*subgoal*/*action* & *C*).

The second kind of probabilistic undermining is somewhat different. The plan relies upon a probability prob(*subgoal-at-t* / *subgoal-at-t_0*). This may just be based upon a defeasible temporal projection, or it may be based upon more concrete probabilistic knowledge. Undermining occurs when there is a context *C** made probable by the undermining plan and an act *A** such that prob(*subgoal-at-t* / *A** & *C** & *subgoal-at-t_0*) ≠ prob(*subgoal-at-t* / *subgoal-at-t_0*).

A third kind of undermining is *value-undermining*. Computing the expected value of a plan depends upon assumptions about the values of goals and side-effects. These are measured by conditional utilities U(*G*/*C*) where *C* is a circumstance made probable by the plan. Value undermining occurs when the second plan makes *C** probable, where U(*G*/*C* & *C**) ≠ U(*G*/*C*).

There is a fourth kind of decision-theoretic interference. Recall that decision-theoretic interference occurs when the value of merging two plans is not the sum of the values of the plans. One way in which this can happen is when the plans share steps, because in that case the execution costs of the merged plan will typically be less than the sum of the execution costs of the individual plans, due to something's having to be done only once. This is the "good kind" of interference that improves the merged plan. I suggest that the decision when to share steps between plans should usually be done at the level of the master plan, not at the level of the local plans, because it is a decision that must be based upon the expected values of the

results. The only exception to this occurs when resource constraints dictate that something cannot be done more than once, in which case the plan construction algorithm can determine that steps must be shared.

To incorporate decision-theoretic planning of the sort I am describing into an autonomous agent, we need procedures for finding decision-theoretic interference, and also procedures for fixing plans that exhibit it, i.e., making changes to them to eliminate undesirable interference. Of course, not all interference is undesirable, so we also need procedures for adding desirable interference. In general, whenever we find a plan with a positive expected value, we should adopt interest in finding ways to modify it to increase the expected value. These procedures can be modeled on conventional threat detection and plan repair techniques used in classical planning.

2.4 Hierarchical Planning

This approach to decision-theoretic planning interacts in illuminating ways with some familiar ideas from classical planning. One is that planning should be hierarchical. This is usually defended on the grounds that it makes planning more efficient (Barrett and Weld 1994), but it is noteworthy that it can also produce plans with higher expected values. This is because we can be more confident of being able to perform a high level action (e.g., *drive across town*) in *some way or other* than we can in being able to perform it in any particular way. If we build a detailed route into our plan, the probability of plan failure may be high, but if we just plan to drive across town some way or other, the probability of plan failure may be low. We make our plans more secure by planning hierarchically, where the high-level tasks are such that we can construct alternate plans when one plan fails or is apt to fail. We rely heavily upon being able to fix plans on the fly, and it is largely our decomposition plans that we are fixing or replacing. To accommodate this, the master plan must consist in part of high level plans, with links to lower level plans. To evaluate the adoption of decomposition plans for the performance of high level actions, it looks like we will have to evaluate the master plan on different levels—both by ignoring decomposition plans, and by adding them into the plan. This is a matter calling for further investigation.

2.5 Conditional Planning

Another way to improve expected values is to adopt *conditional plans* that prescribe different actions under different circumstances. If the success of a plan depends upon something (a contingency) being true that has only a certain probability of being true, we may be able to increase the expected value of the plan by adding a plan for what to do if the contingency fails.

In the literature on conditional (or contingency) planning, it is generally assumed that a conditional plan must contain subplans for an exhaustive set of contingencies (see Peot and Smith 1992, and Pryor and Collins1996). In fact, though, it can be useful to plan for a single contingency, or a small list of non-exhaustive contingencies. We need not assume that our plans are bullet-proof. We need only ensure that they have a reasonable expected-value, and then if they fail we shrug our shoulders and go on. On the other hand, we may also be able to increase the expected-value by adding additional subplans for further contingencies.

A fundamental issue in conditional planning is when to do it. That is, for which contingencies should we plan? Existing conditional planners do not address this issue, taking it as part of the specification of the planning problem by the user. However, an autonomous planning agent must be able to figure this out for itself. The solution to this problem must appeal to

decision-theoretic considerations. Roughly, we should plan for a contingency when we have reason to think that doing so may produce a plan with a reasonable expected-value. This will turn in part on the likelihood of the contingency coming true, and prior knowledge of what is apt to happen if it is does come true. But working this out in implementable detail may be difficult.

Another issue that arises in conditional planning is how to treat information access. To execute a conditional plan, we must know whether a contingency is true. It is initially tempting to think that the antecedents of conditionals in conditional plans are the contingencies themselves. This accords with the way we think about cases where it is obvious to us whether the contingencies hold (e.g., it is raining). But as Pryor and Collins (1996) observe, sometimes information gathering will be complex, and we must make sure that the steps involved in gathering information don't conflict with the plan itself. This suggests that the plan should incorporate the information gathering steps as part of it. This can be accommodated by making the antecedents of the conditionals epistemic: "If you know that *P*, do *A*", "If you know that ~*P*, do *B*". We could also have a condition "If you don't know that *P*, do *C*", or "If you are uncertain whether *P*, do *C*". The latter could not be captured by taking the antecedents to be non-epistemic. Notice that the cases in which it will be obvious whether the contingencies hold can be accommodated by employing high level operators like "Observe whether *P*". Then we do not have to plan explicitly for information gathering, leaving that planning until the time of execution. So there is an important interaction between hierarchical planning and conditional planning here.

An independent argument for including knowledge acquisition steps is that at the time when you want to know whether *P*, it may be too late to take the steps that are required for getting that knowledge. For instance, you might have to ask a question of a friend before he leaves, or use the phone before you do something that makes it inaccessible. So you will need to incorporate these steps into the plan. However, in cases in which you expect it to be easy to acquire the knowledge when you need it, you can leave the details undetermined.

The inclusion of information acquisition steps interacts in important ways with plan execution. It is tempting to suppose that in executing a plan the agent should monitor the course of execution and verify that contingencies hold and subgoals are achieved before continuing the execution. However, that will not always be possible. For example, one step of a plan might involve calling a friend and asking him to do something. I may have no way of verifying that he does it. The best I can do is continue plan execution on the assumption that the subgoal has been achieved. If it hasn't, then I incur execution costs that do not actually contribute to the achievement of the goal. What this illustrates is that in computing the expected value of a plan, one must take account of the possibility that plan execution will continue, incurring execution costs, even though the plan has already failed. This suggests that monitoring steps should be explicitly included in the plan when they are to be employed, because whether we monitor plan execution can affect the expected value of the plan.

2.6 Realistic Planning

I can summarize the suggestions of the previous subsections by saying that realistic planning in autonomous agents operating in complex and partially unpredictable environments must be continuous, decision-theoretic, hierarchical, conditional, and classical. These are all familiar ideas in contemporary planning, but I doubt that any one of them can be worked out correctly in isolation. They have important interactions with each other, and what is needed is a single

theory that accommodates them all. I have only mentioned plan execution in passing here, but I think that the theory must address that issue as well. The key to my proposed solution to this nexus of problems is that plan search should proceed classically, but using probabilistic connections rather than exceptionless causal connections. The kind of classical planning involved should be both hierarchical and conditional. The plans produced must be evaluated decision-theoretically, but that is done by incorporating them into the master plan and using defeasible rules for evaluating it. The master plan will also provide the data structure for keeping track of decomposition plans, and updating the planning task as new goals are adopted and plan steps are performed in the course of executing plans.

References

Barrett, A., and Weld, D.
1994 "Partial order planning: evaluating possible efficiency gains", *Artificial Intelligence* **67**, 71-112.
Barto, A. G., S. J. Bradtke, and S. P. Singh
1995 "Learning to act using real-time dynamic programming". *IEEE Transactions on Automatic Control* 34, 589-598.
Boutelier, Craig, Ronan Brafman, and Chistopher Geib
1997 "Prioritized goal decomposition and Markov decision processes: toward a synthesis of classical and decision theoretic planning", *IJCAI-97*.
Boutelier, Craig, Thomas Dean, and Steve Hanks
1999 "Decision theoretic planning: structural assumptions and computational leverage", forthcoming in *Journal of AI Research.*
Boutelier, Craig and Richard Dearden
1994 "Using abstractions for decision-theoretic planning with time constraints", *IJCAI'94*, 24-32.
Boutelier, Craig Richard Dearden, and Moises Goldszmidt
1995 "Exploiting structure in policy construction", *IJCAI'95*, 1104-1111.
Dean, Thomas, and Robert Givan
1997 "Model minimalization in Markov decision processes", *AAAI-97*.
Dean, Thomas, Robert Givan, and Kee-Eung Kim
1998 "Solving stochastic planning problems with large state and action spaces", *AIPS'98*, 102-110.
Dean, Thomas, Leslie Kaelbling, Jak Kirman, and Ann Nicholson
1993 "Planning with deadlines in stochastic domains", *AAAI'93*, 574-579.
Draper, Denise, Steve Hanks, and Daniel Weld
1994 "Probabilistic planning with information gathering and contingent execution", *AAAI Spring Symposium on Decision-Theoretic Planning,* 76-82.
Haddawy, Peter, Doan, A., and Goodwin, R.
1995 "Efficient decision-theoretic planning: techniques and empirical analysis", *UAI*, 229-236.
Haddawy, Peter, and Steve Hanks
1990 "Issues in decision-theoretic planning: symbolic goals and numeric utilities", *Proceedings of the DARPA Workshop on Innovative Approaches to Planning, Scheduling, and Control*, 48-58.
Kautz, H. A., and B. Selman

1996 "Pushing the envelope: planning, propositional logic, and stochastic search", *Proceedings of AAAI-96*, 1194-1201.

1998 "Blackbox: a new approach to the application of theorem proving to problem solving", in *AIPS98 Workshop on Planning as Combinatorial Search*, 58-60.

Kushmerick, N., Hanks, S., and Weld, D.

1995 "An algorithm for probabilistic planning". *Artificial Intelligence* **76**, 239-286.

McAllester, David, and Rosenblitt, David

1991 "Systematic nonlinear planning", *Proceedings of AAAI-91*, 634-639.

Ngo, L., P. Haddawy, H. Nguyen, "A modular structured approach to conditional decision-theoretic planning", *AIPS'98*, 111-118.

Penberthy, J. Scott, and Weld, Daniel

1992 "UCPOP: a sound, complete, partial order planner for ADL". *Proceedings 3rd International Conference on Principles of Knowledge Representation and Reasoning*, 103-114.

Peot, M. A., and D. Smith

1992 "Conditional nonlinear planning", *Proceedings of the First International Conference on Artificial Intelligence Planning Systems*, 189-197.

Pollock, John

1992 "New foundations for practical reasoning", *Minds and Machines*, **2**, 113-144.

1995 *Cognitive Carpentry*, MIT Press.

1998 "The logical foundations of goal-regression planning in autonomous agents", *Artificial Intelligence* 106, 267-335.

Pryor, Louise, and Gregg Collins

1996 "Planning for contingencies: a decision-based approach", *Journal of Artificial Intelligence Research* **4**, 287-339.

Simmons, Reid and Sven Koenig

1995 "Probabilistic robot navigation in partially observable environments", *IJCAI'95*, 1080-1087.

Tash, Jonathan, and Stuart Russell

1994 "Control strategies for a stochastic planner", *AAAI'94*, 1079-1085.

Weld, Daniel

1999 "Recent advances in AI planning", *AI Magazine*, forthcoming.

Williamson, Mike and Steve Hanks

1994 "Beyond symbolic goals: a step toward utility-directed planning", *AAAI Spring Symposium on Decision-Theoretic Planning*, 272-278.