Against Optimality: Logical Foundations for Decision-Theoretic Planning in Autonomous Agents

John L. Pollock
Department of Philosophy
University of Arizona
Tucson, Arizona 85721
pollock@arizona.edu
http://www.u.arizona.edu/~pollock

Abstract

This paper investigates decision-theoretic planning in sophisticated autonomous agents operating in environments of real-world complexity. An example might be a planetary rover exploring a largely unknown planet. It is argued that existing algorithms for decision-theoretic planning are based on a logically incorrect theory of rational decision making. Plans cannot be evaluated directly in terms of their expected values, because plans can be of different scopes, and they can interact with other previously adopted plans. Furthermore, in the real world, the search for optimal plans is completely intractable. An alternative theory of rational decision making is proposed, called "locally global planning".

Keywords: decision theory, planning, optimality.

1. Decision-Theoretic Planning

Planning by rational agents operating in environments of real-world complexity must be decision-theoretic. This general point is widely appreciated in the AI planning community, and there is a lot of current interest in constructing decision-theoretic planners (for example, see Blythe and Veloso 1997, Boutelier et al 1999, Haddawy and Hanks 1990, Ngo, Haddawy and Nguyen 1998, Majercik 1999, Majercik and Littman 1999, Onder and Pollack 1997, 1999, Onder, Pollack and Horty 1998, Peterson and Cook 2000, and Williamson and Hanks 1994). However, the rush to implementation has proceeded without careful consideration of the theoretical foundations of such planning.

The normal presumption in decision-theoretic planning has been that plans can be compared directly in terms of their expected values. It is supposed that the objective of decision-theoretic planning is to find an optimal plan, i.e., a plan such that no alternative plan has a higher expected value. This approach to decision-theoretic planning amounts to taking classical decision theory, which is about choosing between alternative actions, and applying it directly to plans, simply replacing "actions" by "plans" throughout the theory. I will refer to this as *simple plan-based decision theory*. It seems to be assumed that this is somehow conflict with classical decision theory by prescribing the performance of actions contained in plans when those actions would not be individually optimal (and so would not be chosen by classical decision theory). If this can happen, then rather than supporting plan-based decision theory, classical decision theory entails the falsity of plan-based decision theory. And as I will show below (see also my (1992) and (1995)), this can happen. It follows that proponents of plan-based decision theory cannot get away with just waving their hands vaguely at classical decision theory and saying "so we can evaluate plans decision-theoretically".

If plan-based decision theory is true, classical decision theory must be false. So a defense of plan-based decision theory must be based upon a criticism of classical decision theory. I will present an argument to the effect that the proper objects of rational choice are plans rather than actions. Actions become derivatively rational by being prescribed by rationally chosen plans. This is a defense of plan-based decision theory, but does not yet tell us how it is to work in detail. If our objective is to solve planning problems in a carefully crafted planning domain of limited scope, and care is taken to limit the class of plans that can be considered, then simple plan-based decision theory may work. For example, a successful mail-delivery robot might well be constructed in this way. But for autonomous agents of unlimited scope, e.g., an autonomous planetary rover exploring an environment in which much is initially unknown, there is just one planning domain — the world. I will argue that once we leave toy problems behind and focus on decision-theoretic planning in the real world, plan-based decision theory cannot be viewed as a search for optimal plans. I will argue that optimality is not even well-defined outside of toy problems, so plan adoption must be based on somewhat different considerations. I will propose a theory that I call *locally global planning*.

My ultimate objective is to implement a decision-theoretic planner that makes rationally correct decisions. But before we can do that, we must have an account of rational decision making that is applicable to this problem. The first main point of this paper is that classical decision theory is not a correct theory of rational decision making, and that simple plan-based decision theory is not either. There is a theoretical problem here that must be solved before implementation should even begin. This paper is aimed at the theoretical problem of giving a correct theory of rational decision making applicable to decision-theoretic planners. It follows from the conclusions of this paper that all existing planners will produce logically incorrect solutions when applied to the complex planning problems facing a sophisticated autonomous agent operating in an environment of real world complexity (if they are able to produce any solutions at all). That is the second main point of the paper. It means that planning theorists must look in new directions to construct a generally adequate planner.

I could stop there and claim to have shown something important. Those are the two main points of the paper. However, it may be useful to make some tentative remarks about how we might produce an alternative planning system that will produce logically correct solutions to these planning problems. That requires the theory to be worked out in further detail. I feel that it is important to lay the theory out as just one of several orthogonal difficulties for the theory of decision-theoretic planning. All of this must be combined into a single theory before a correct decision-theoretic planner can be constructed.

2. Classical Decision Theory

Classical decision theory is a theory of rational choice. It is a theory of how an agent should, rationally, go about deciding what actions to perform at any given time. It is assumed that these decisions must be made in the face of uncertainty regarding both the agent's initial situation and the consequences of his actions. By "classical decision theory" I mean the nexus of ideas stemming in part from Ramsey (1926), von Neumann and Morgenstern (1944), Savage (1954), Jeffrey (1965), and others who have generalized and expanded upon it. The different formulations look very different, but the basic prescription of classical decision theory can be stated simply. We assume that our task is to choose an action from a set $\mathbf{A}$ of *alternative actions*. The actions are to be evaluated in terms of their outcomes. We assume that the *possible outcomes* of performing these actions are partitioned into a set $\mathbf{O}$ of pairwise exclusive and jointly exhaustive outcomes. We further assume that we know the probability $\text{prob}(O/A)$ of each outcome conditional on the performance of each action. Finally, we assume a *utility measure* $\mathbf{U}(O)$ assigning a numerical utility value to each possible outcome. The *expected value* of an action is defined to be a weighted average of the values of the outcomes, discounting each by the probability of that outcome occurring if the action is performed:

$$\mathbf{EV}(A) = \sum_{O \in \mathbf{O}} \mathrm{U}(O) \cdot \text{prob}(O / A).$$

The crux of classical decision theory is that actions are to be compared in terms of their expected values, and rationality dictates choosing an action that is *optimal*, i.e., such that no alternative has a higher expected value.

3. Groups of Actions

Classical decision theory has us choosing actions one at a time, each on its own merit. In Savage's (1954) "small worlds", in which different decisions can be made wholly independently of one another, this may be satisfactory. But I will argue that in general individual actions cannot be the units of rational decision. This is for at least two reasons. The first reason turns upon the observation that in the real world we typically have a number of different decisions to make at more or less the same time. An autonomous actions one at a time and has us choose them individually on the basis of their being optimal. The problem is that decisions can interact. Carrying out one decision may alter the probabilities and utilities involved in another decision thereby changing what action is optimal. It could be that, prior to deciding where to recharge its battery, because the battery charge is fairly low, the optimal decision would be to postpone delivering the package until after recharging the battery. But if the delivery vehicle decides to recharge the battery at a recharge station far from the destination of the package (perhaps because it costs less), and the vehicle has other things to do in that part of town that could occupy it for the rest of the afternoon, this may make it better to deliver the package before recharging the battery. Alternatively, because the battery is fairly low and the vehicle wants to recharge it before making the delivery, it might be better to choose a different (more expensive) recharge station. The point is that actions can interfere with one another, with the result that if several actions are to be chosen, their being individually optimal does not guarantee that the group of them will be optimal. This strongly suggests that the object of decision-theoretic evaluation should be the entire group of actions rather than the individual actions.

This same conclusion can be defended in another way. Often, the best way to achieve a goal is to perform several actions that achieve it "cooperatively". Performing the actions in isolation may achieve little of value. In this case we must choose actions in groups rather than individually. To illustrate, suppose the agent's objective is to transport a ton of silver and a ton of gold from one location to another. The agent has a one-ton truck. It could fit both the gold and the silver into the truck at the same time and transport them on a single trip, but in doing so it would risk damaging the truck springs. The actions being considered are to transport the gold on a single trip, to transport the silver on a single trip, and to transport both on a single trip. We can imagine the probabilities and utilities to be such that the action with the highest expected value is that of transporting both on a single trip, even though that risks damaging the springs. However, if the agent has time to make two trips, that might be the better choice. That is, it should perform *two* actions, transporting the gold on one trip and the silver on another, rather than performing any of the single actions I am considering. This illustrates again that actions cannot always be considered in isolation. Sometimes decision-theoretic choices must be between groups of actions, and the performance of a single action becomes rational only because it is part of a group of actions whose choice is dictated by practical rationality.

These two examples illustrate two different phenomena. In the first, actions interfere with each other, changing their execution costs and hence expected values from what they would be in isolation. In the second, actions collaborate to achieve goals cooperatively, thus changing the expected values by changing the probabilities of outcomes occurring. These examples might be viewed as cases in which it is unclear that actions even have well-defined expected values in isolation. To compute the expected value of an action we must take account of the context in which it occurs. If the expected values are not well-defined, then classical decision theory cannot be applied to these decision problems. Alternatively, if we suppose the members of the group would not be chosen individually on their own strength. Rather, a pairwise comparison of actions would result in the action of transporting both the gold and silver on a single trip being chosen, and that is the intuitively wrong choice. In these examples, it is the group itself that should be the object of rational choice, and the individual actions are only derivatively rational, by being contained in the rationally chosen group of actions.

Groups of actions, viewed as unified objects of rational deliberation, are *plans*. The actions in a plan may be good actions to perform only because they are part of a good plan. It appears that the only way to get decision theory to make the right prescription in the above example is to apply it to plans rather than individual actions. The reason the agent should transport the gold alone on a single trip is that doing so is part of the plan of making two trips, and that plan is better than the plan of transporting both the gold and silver on a single trip. The plan of making two trips has a higher expected value than the plan of transporting both the gold and silver on a single trip, and that is the basis upon which it is chosen.

The reaction of most planning theorists to these examples will probably be, "Of course. So what?" But this is because I am preaching to the converted. The important point is that classical decision theory mandates that choices are to be made between individual actions. I have been arguing, to the contrary, that rational choices must often be made between plans. This, I take it, is the justification for the entire enterprise of decision-theoretic planning. Contrary to what is generally supposed in the decision-theoretic planning community, it does not proceed by deriving decision-theoretic planning from classical decision theory, but rather by arguing that classical decision theory is wrong in ways that require its replacement by decision-theoretic planning.

Given that rational choice requires decision-theoretic planning, how do we do it? The obvious proposal is to simply replace actions by plans in classical decision theory. *Simple plan-based decision theory* would propose that we choose between competing plans in terms of their expected values. Savage (1954) seems to suggest that plans can be chosen in this way. And all of the work on decision-theoretic planning cited above takes this same form.

4. Choosing Between Plans

Simple plan-based decision theory proposes that we can characterize rational choice by reconstruing classical decision theory to apply to plans rather than actions. However, we need not choose between plans unless they are in some sense in competition. If two plans are not in competition, we can simply adopt both. So to construct a plan-based theory of rational choice, we need an account of when plans compete in such a way that a rational choice should be made between them.

Competing plans should be plans that we must choose between, rather than adopting both. A copy to one of two offices. It has to choose between the plan to copy the document and deliver it to office-1 and the plan to copy the document and deliver it to office-2. It *could* do both. The plans do not compete strongly. But nevertheless, they are in competition. We do not want the robot to adopt both plans because a copy of the document only needs to be delivered to one office. We might capture this with a notion of weak competition — let us say that two plans *compete weakly* iff the plan that results from merging the two plans into a single plan has a lower expected value than at least one of the original plans. Delivering copies of the document to both offices has a lower expected value than delivering a copy to just one because (1) the execution cost is essentially the sum of the execution costs of the two separate plans, but (2) only one delivery is required, so the payoff from delivering copies to both offices is no greater than the payoff from delivering a copy to just one. It might be proposed, then, that two plans are competitors iff they compete weakly.

An appeal to weak competition does not generate a theory with quite the same structure as classical decision theory. The problem is that classical decision theory assumes we have a set of alternative actions, and prescribes choosing an optimal member of the set. However, weak competition doesn't generate a set of alternatives that are pairwise competitors. This is because weak competition is not transitive. $Plan_1$ may compete weakly with $plan_2$, and $plan_2$ with $plan_3$, without $plan_1$ competing weakly with $plan_3$. Thus if we simply pick a plan and let the set of alternatives be the set of all plans competing weakly with the given plan, it does not follow that other members of the set of alternatives will be in competition. It may be desirable to execute several of those other plans together. For instance, suppose I am planning a wedding and want to select something borrowed and something blue, but it is undesirable to select two borrowed things or two blue things. If x and y are borrowed, and y and z are blue, then selecting x competes weakly with selecting y, and selecting y competes weakly with selecting z, but selecting x does not compete weakly with selecting z. Accordingly, it seems that what classical decision theory really ought to say is:

> It is rational to adopt (decide to execute) a plan iff it has no weak competitor with a higher expected value.

However, this proposal turns out to be incompatible with another fundamental observation about cognitively sophisticated autonomous agents operating in a complex environment. First, such an agent (e.g., a planetary rover) is not faced with a single fixed planning problem. One reason for this is that its beliefs will change as it acquires experience of its environment and as it has time for further reasoning. This will affect what solutions are available for its planning problems. For instance, imagine a planetary rover that discovers, during its sojourn on a strange planet, that sand of a certain color tends to be loose and powdery, causing it to get bogged down. In light of this newly discovered information, when it plans a route to a destination it must detour around such patches of sand. Plans that would have been acceptable

An agent that engages in goal-directed planning must be equipped with a disposition to adopt new goals of attainment under various circumstances. Note that this is not just a matter of building in a general goal like "Always avoid sandstorms", because if on one occasion the rover is unable to take cover and weathers the storm out in the open, that general goal is henceforth unattainable no matter what the rover does, but we do not want that to affect its trying to avoid future sandstorms. This point is discussed in more detail in section nine.

The planning problems the agent faces evolve over time. This is a fundamental observation about cognitively sophisticated agents operating in complex environments. It has an important corollary. We cannot expect the agent to redo all of its previous planning each time it acquires new knowledge or new goals, so planning must produce lots of *local plans*. These are small plans of limited scope aiming at disparate goals.

The difficulty is now that simple plan-based decision theory cannot be applied to local plans. Simple plan-based decision theory is formulated in terms of whether there exists (in logical space) a competing plan with a higher expected value. The problem is that there will almost always exist such a competing plan. To illustrate, think again about the office robot that is choosing between copying the document and delivering it to office-1 and copying the document and delivering it to office-2. Suppose the former has the higher expected value (perhaps because it has a lower execution cost). Simple plan-based decision theory has the consequence that the plan of delivering a copy to office-2 is not rationally adoptable, but that does not imply that the plan of delivering a copy to office-1 is adoptable. This is because some other plan with a higher expected value may compete with it. And we can generally construct such a competing plan by simply adding steps to the earlier competing plan. For this purpose, we select the new steps so that they constitute a subplan achieving some valuable unrelated goal. For instance, we can consider the plan of delivering a copy to office-2 and then recharging the battery. This plan still competes with the plan of delivering a copy to office-1, but it has a higher expected value. Thus the plan of delivering a copy to office-1 is not rationally adoptable. However, the competing plan is not rationally adoptable either, because it is trumped by the plan of delivering a copy to office-1 and then later recharging the battery.

It seems clear that given two competing plans P_1 and P_2, if the expected value of P_1 is greater than that of P_2, the comparison can generally be reversed by finding another plan P_3 that pursues unrelated goals and then merging P_2 and P_3 to form P_2+P_3. If P_3 is well chosen, this will have the result that P_2+P_3 still competes with P_1 and the expected value of P_2+P_3 is higher than the expected value of P_1. If this is always possible, then there are no optimal plans and simple plan-based decision theory implies that it is not rational to adopt any plan.

In an attempt to avoid this problem, it might be objected that P_2+P_3 is not an appropriate object of decision-theoretic choice, because it merges two unrelated plans. However, recall the second example used to motivate the application of decision theory to plans rather than actions — the example of transporting a prescription in this example.

The inescapable conclusion is that the rational adoptability of a local plan cannot require that it have a higher expected value than all its competitors. The problem is that local plans can have rich structures and can pursue multiple goals, and as such they are indefinitely extendable. We can almost always construct competing plans with higher expected values by adding subplans pursuing new goals. Thus there is no way to define optimality so that it is reasonable to expect there to be optimal plans. Hence simple plan-based decision-theory fails when applied to local plans.

I have heard it suggested that this problem does not arise for state space planners, e.g., Markov decision process planning. However, there are two ways of viewing state space planning. We could think of the entire world as a single state space and try to produce a single "universal" plan governing the agent's actions for all time. But that problem is completely intractable — see the next section. Alternatively, we might use state space planning techniques to find local plans (policies) by separating out toy problems, but here the preceding difficulty recurs. You can compare the plans produced for a single toy problem in terms of their expected values, but if you expand the problem and consider more sources of value, more possible actions, etc., in effect looking at a bigger toy problem, an optimal policy for the larger problem may not contain the optimal policy for the subproblem. So the problem recurs. It arises independently of the kind of planning algorithms employed.

5. Universal Plans

There is a way of trying to save simple plan-based decision theory from the preceding objection. The argument that led to the conclusion that plans cannot be selected for adoption just by comparing their expected values turned upon its always being possible to extend a plan by merging it with a subplan for achieving an additional goal. For local plans, this assumption is unproblematic. But there is one way of avoiding the argument — consider only "universal plans". These are plans prescribing what the agent should do for all the rest of its existence. Universal plans cannot be extended by adding subplans for new goals. Because universal plans include complete prescriptions of what to do for the rest of the agent's existence, any two universal plans will make different prescriptions, and so will strongly compete. It seems initially quite plausible to suppose that universal plans can be compared in terms of their expected values, and that a universal plan is rationally adoptable iff it is optimal, i.e., iff no other universal plan has a higher expected value.

Savage (1954) toys with the idea that rational decisions should be between universal plans, but he rejects it for the obvious reason. The real world is too complex. No agent with realistic computational limitations could possibly construct a universal plan prescribing the optimal action for every possible state situations that it can distinguish between and must plan for. A universal plan must prescribe the optimal action for each of these 2^{300} possible situations. 2^{300} is an immense number — approximately equal to 10^{90}. To appreciate just how large this number is, it has been estimated that the number of elementary particles in the universe is 10^{78}. So this would require choosing optimal actions for 12 orders of magnitude more possible world-states than there are elementary particles in the universe. And this is for just 300 properties. Of course, if the world is particularly well behaved, then it may be possible to give general descriptions of optimal actions in large classes of world-states rather than making explicit prescriptions for each world-state (i.e., factor the universal plan). But it is preposterous to suppose that even that will enable a real agent to find a universal plan for dealing with every possible world-state it may encounter in the real world. Real agents will not be able to find universal plans.

It is to be emphasized that the preceding remarks are aimed at agents with fairly sophisticated aspirations. If one wants to build a very simple planning agent that is only able to perform a very narrow range of tasks, then one might solve the problem by being very selective about the properties considered, and the agent might be able to construct a universal plan. This might work for a mail delivery robot, but it cannot possibly work for an agent as complex as a planetary explorer — or at least, not a good one. And it is interesting to see how easily the problem can arise even in quite restrictive domains. Consider the following problem, which generalizes Kushmerick, Hanks and Weld's (1995) "slippery gripper" problem. We are presented with a table on which there are 300 numbered blocks, and a panel of correspondingly numbered buttons. Pushing a button activates a robot arm which attempts to pick up the corresponding block and remove it from the table. We get 100 dollars for each block that is removed. Pushing a button costs two dollars. The hitch is that some of the blocks are greasy. If a block is not greasy, pushing the button will result in its being removed from the table with probability 1.0, but if it is greasy the probability is only 0.1. The probability of any given block being greasy is 0.5. We are given 300 chances to either push a button or do nothing. In between, we are given the opportunity to look at the table, which costs one dollar. Looking will reveal what blocks are still on the table, but will not reveal directly whether a block is greasy. What should we do? Humans find this problem terribly easy. Everyone I have tried this upon has quickly produced the optimal plan: push each button once, and don't bother to look at the table. On the other hand, I surveyed existing decision-theoretic planners a few years ago and discovered that none of them could solve this problem. In fact, most of them could not even encode the problem ("An easy 'hard problem' for decision-theoretic planners", OSCAR Project technical report, available at http://oscarhome.soc-sci.arizona.edu/ftp/publications.html).

We can cast this as a 599 step POMDP (partially observable Markov decision problem). Odd-numbered steps consist of either pushing a button or performing the null action, and even-numbered-steps consist of either looking at the table or performing the null action. World-states are determined by which blocks are on the table (T_i) and which blocks are greasy (G_i). The actions available are *nil* (the null action), P_i (push number of elementary particles in the universe. The state-space gets even larger when we move to POMDP's. In a POMDP, the agent's knowledge is represented by a probability distribution over the space of world-states. The probability distributions constitute *epistemic states*, and actions lead to transitions with various probabilities from one epistemic state to another (Aström 1965). In general, POMDP's will have infinite spaces of epistemic states corresponding to all possible probability distributions over the underlying state-space, however a reachability analysis can often produce a smaller state-space with just finitely many possible probability distributions. In the slippery blocks problem, it can be shown that in reachable epistemic states prob(T_i) can taken any value in the set {1.0, $.5{\cdot}.9$, $.5{\cdot}.9^2$, ... , $.5{\cdot}.9^{300}$, 0} and prob(G_i) can take the values .5 and 1.0. Not all combinations of these values are possible, but the number of reachable epistemic states is greater than 301^{300}, which is approximately 10^{744}. A universal plan is equivalent to a policy for this state space. There is no way to even encode such a policy explicitly.

On the other hand, as remarked above, humans have no difficulty solving this planning problem. Why? Part of the answer is that to find an optimal plan, we do not have to consider all of the states that could be reached by executing other plans. We need only consider the states that our chosen plan might get us into. Most of those states are irrelevant, and so there is no need to find (and no possibility of finding) a universal plan. And this arises already for what is basically a toy problem. The real world can be expected to be much more complex, even for agents that ignore most of it. So autonomous agents operating in the real world must aim at constructing local plans, not universal plans.

To recapitulate, cognitively sophisticated agents operating in the real world are constrained to construct local plans. In constructing a plan-based decision theory, we want to define a relation of "better plan" that can drive rational decisions. It is natural to suppose that a plan should only be adopted if it has no competitor that is better than it. If we make the latter assumption and we identify "better plan" with "plan with a higher expected value", then it will normally be the case that no local plan is rationally adoptable. This result is absurd, so such an identification must be incorrect. Thus *simple* plan-based decision theory must be rejected. Decision-theoretic planning must be based on an alternative theory of rational decision making.

6. The Real World vs. Toy Problems

My main objection to classical decision theory is that actions cannot be chosen in isolation. My main objection to simple plan-based decision theory is that optimality must always be relative to a set of alternatives, and there is no way of selecting an appropriate set of decision-theoretic alternatives in such a way that there will always or even normally be an optimal choice. A secondary objection is that even in those rare cases where there might be an optimal plan, real agents cannot be expected to find them.

Furthermore, some of these toy problems can generate very useful practical applications. However, the logic of toy problems is different from the logic of decision making in the real world. Typically, the search for an optimal plan in a toy problem can be made by searching a manageably small (finite or finitely describable) set of alternatives. In this case optimality may be well-defined, and the prescription to choose an optimal plan would seem sensible. It would seem reasonable to require the agent to evaluate each of the alternative plans and choose an optimal one.

However, outside of toy problems, the supposition that we need only search a very limited range of plans is indefensible. Plans are mathematical or logical constructions of unbounded complexity. In the real world, if we can construct one plan for achieving a goal, we can typically construct infinitely many. For planning by sophisticated autonomous agents in the real world, talk of "planning domains" is inappropriate. The only planning domain is the whole world.

It should be emphasized that this only is a claim about cognitively sophisticated autonomous agents operating in the real world — not in artificially limited environments. Ultimately I want to know how to build Isaac Asimov's *I-Daneel* and Arthur C. Clarke's *HAL*. I take it that this is one of the seminal aspirations of AI, although at this time no one is in a position to actually try to do it. For now I will be satisfied with building an autonomous planetary rover that can learn about its environment and meet the challenges it is apt to encounter. Even in connection with such more limited agents I sometimes get the reaction that we are not currently able to build such agents, so why worry about them? The answer is that although we haven't built one yet, it is something people are actively trying to do. And I, for one, do not think we are that far away from being able to construct agents with the requisite cognitive sophistication. That has been the longstanding goal of the OSCAR project (Pollock 1965, 1998a, 2002a), and the result so far is an agent that is able to reason in a sophisticated way about perceived situations, change and persistence, causes, and probabilities, and is able to do so in real time against the background of a large database of beliefs (on the order of 100,000).

For some concrete applications it is possible to constrain the planner's environment sufficiently to turn the problem into what is in effect a toy problem. This may be possible if the planning problem can be fixed from inception and either (1) local plans aiming at distinct goals do not interact, or (2) it is possible to construct universal plans with manageably short time horizons and manageably few actions and world-states. In such a world, optimization will be both well-defined and desirable. But the most important point of this paper is that the techniques that work for toy problems will, for purely logical reasons, not scale up to the problem of building a cognitively sophisticated autonomous agent operating in the real world. And as I have illustrated, for this problem to arise, the agents needn't be all that sophisticated. For such agents, there is no reason to expect there to be optimal plans. The difference between toy problems and the real world is not just a difference of degree. Planning must work entirely differently. This is a logical problem that must be solved before we can even begin building a decision-theoretic planner of use in autonomous extendable, it will usually turn out that none is. This cannot have the consequence that we should not adopt any plan, so it follows that finding optimal plans cannot be the proper aim of decision-theoretic planning. On the other hand, there has to be some way of making sense of one plan being better than another, and it seems this must have something to do with the probabilities and utilities of outcomes. Further, it seems that there must be some account of rational decision making that proceeds in terms of this relation of one plan being better than another. In the next section, I will propose an alternative to simple plan-based decision theory.

7. When is One Plan Better than Another?

We want to define a notion of "better plan" which is of use in deciding whether a plan should be adopted. I have argued that this cannot be cashed out as the first plan merely having a higher expected value than the second. To get a grip on this notion, let us think about plan adoption in rational agents. First consider the limiting case in which an agent has no background of adopted plans, and a new plan is constructed. Should the new plan be adopted? The basic insight of classical decision theory is that what makes a course of action (a plan) good is that it will, with various probabilities, bring about various valuable results, and the cost of doing this will be less than the value of what is achieved. This can be assessed by computing the expected value of the plan. In deciding whether to adopt the plan, all the agent can do is compare the new plan with the other options currently available to it. If this is the only plan the agent has constructed, there is only one other option — do nothing. So in this limiting case, we can evaluate the plan by simply comparing it with doing nothing.

Things become more complicated when the agent has already adopted a number of other plans. This is for two reasons. First, the new plan cannot be evaluated in isolation from the previously adopted plans. Trying to execute the previous plans may affect both the probabilities and the utilities employed in computing the expected value of the new plan. For example, if the new plan calls for the agent to perform an operation that requires large amounts of battery power, the probability of the agent being able to do that may normally be fairly high. But if other plans the agent has adopted will result in the agent having depleted its battery power, then the probability of being able to perform the operation may be lower. So the probabilities can be affected by the context provided by the agent's other plans. The same thing is true of the values of goals. Suppose the new plan is a plan for recharging the agent's battery. In the abstract, this may have a high value, but if it is performed in a context in which the agent's other plans include replacing the battery in a short while, the value of the recharge may be seriously diminished. Execution costs can be similarly affected. If the new plan prescribes transporting an object from one location to another in a truck, this will be more costly if a previous plan moves the truck to the other side of town.

normally have their results only probabilistically.

The second reason it becomes more complicated to evaluate a new plan when the agent already has a background of adopted plans is that the new plan can affect the value of the old plans. If an old plan has a high probability of achieving a very valuable goal but the new plan makes the old plan unworkable, then the new plan should not be adopted. Note that this is not something that is revealed by just computing the expected value of the new plan.

We have seen that normal planning processes produce local plans. How should the agent decide whether to adopt a new local plan? The decision must take account of both the effect of previously adopted plans on the new plan, and the effect of the new plan on previously adopted plans. We can capture these complexities in a precise and intuitively appealing way by defining the concept of the agent's *master plan*. This is the result of merging all of the plans the agent has adopted but not yet executed into a single plan.

Don't confuse the master plan with a universal plan. The master plan simply merges a number of local plans into a single plan. Each local plan talks about what to do under certain circumstances, so the resulting master plan talks about what to do under every circumstance mentioned by any of the individual local plans. But this is still a very small set of circumstances relative to the set of all possible world-states. If none of the local plans have anything to say about what to do in some new previously unconsidered situation, then the master plan doesn't either. But by definition, a universal plan must include a prescription for what to do in every situation. If we have n local plans each making m prescriptions of the form "If C is true then do A", the master plan will contain $m \cdot n$ prescriptions. But supposing the conditions C are all logically independent of each other, a universal plan for the state space generated by just this limited vocabulary will contain $2^{m \cdot n}$ prescriptions. Typically an agent will be capable of considering what to do in a much larger set of circumstances not yet addressed by any of the local plans it has adopted. If there are N such circumstances, a universal plan must include 2^N prescriptions. For example, if the agent has thus far adopted 30 ten-step plans, the master plan will include 300 prescriptions, but a universal plan would have to consider at least 2^{300} (i.e., 10^{90}) prescriptions, and probably many orders of magnitude more.

Although master plans are totally different beasts from universal plans, they share an important property — master plans can be meaningfully compared in terms of their expected values. We can think of the master plan as the agent's tool for making the world better. The expected value of the master plan is the agent's expectation of how much better the world will be if it adopts that as its master plan. Thus one master plan is better than another iff it has a higher expected value. Equivalently, rationality dictates that if an agent is choosing between two master plans, it should choose the one with the higher expected value.

Although master plans are vastly smaller than universal plans, they can still be big plans. This is obvious for human beings. At any given time there is a large number of plans that I have adopted and not must have plans for how to handle various emergency situations it might encounter, it must have plans for how to perform a number of routine operations that it will perform in the course of its trip, and so on.

New observations can require (or make desirable) new planning. They can do this either by presenting the agent with new opportunities to achieve valuable results, or by providing information that leads to a re-evaluation of existing plans. The latter may happen, for instance, if our planetary rover finds that its planned path will result in its falling into a chasm, or becoming mired in quicksand. A re-evaluation of existing plans constitutes a re-evaluation of the master plan. So the objective of new planning is to find a master plan with a higher expected value than its current one. If the only way an agent had of finding a master plan with a higher expected value was to plan all over again from scratch and produce a new master plan essentially unrelated to its present master plan, the task would be formidable. It would at the very least be slow and complex, making it difficult for the agent to respond to emergency situations. And if the agent's master plan is sufficiently complex, the agent's inherent computational limitations may make the task impossible. It does not take a very large problem to bog down a planner. Often, constructing the master plan will be bigger than the largest problems existing planners can solve. Furthermore, if every planning problem requires the construction of a new master plan, then every little planning problem becomes immensely difficult. To plan a route around the chasm, our planetary rover would have to reconsider its entire set of plans. Applying this to human beings, to plan how to make a sandwich for lunch, I would have to replan my entire life.

Obviously, humans don't do this, and artificial agents shouldn't either. Normal planning processes produce local plans, not entire master plans. The only way agents with realistic computational powers can efficiently construct and improve upon master plans reflecting the complexity of the real world is by constructing or modifying them incrementally. When trying to improve its master plan, rather than throwing it out and starting over from scratch, what an agent must do is try to improve it piecemeal, leaving the bulk of it intact at any given time. This is where local plans enter the picture. The significance of local plans is that they represent the building blocks for master plans. We construct master plans by constructing local plans and merging them together.

We must consider more carefully how we are able to do this. After all, there can be interactions between local plans. A newly constructed local plan has at least the potential to interact with *all* of my previously adopted plans. Why should constructing a local plan *in the context of a master plan* be any easier than simply constructing a master plan? I will take this question up in a preliminary way in section nine. But before addressing this question, let us consider a simpler one. Earlier, we encountered the purely logical problem of how to evaluate a newly constructed local plan, given that we must take account both of its effect on the agent's other plans and the effect of the agent's other plans on the new plan. We are now in a position to answer that question. The only significance of local plans is as constituents of the master plan. When a new local plan is constructed, what we want to know is whether the master plan can

$$\mathbf{MEV}(P,M) = \mathbf{EV}(M+P) - \mathbf{EV}(M).$$

If the marginal expected value is positive, adding the local plan to the master plan improves the master plan, and so in that context the local plan is a good plan. Furthermore, if we are deciding which of two local plans to add to the master plan, the better one is the one that adds more value to the master plan. So viewed as potential additions to the master plan, local plans should be evaluated in terms of their marginal expected values, not in terms of their expected values simpliciter.

8. Locally Global Planning

Stating it more precisely, my proposal is that the "better plan" relation is a three-place relation, comparing the marginal expected values of plans relative to master plans. But how exactly do we use this relation to choose between plans? It appears that the aim of plan search is to construct local plans and use them to improve the master plan. It may at first occur to one that the objective should be to find an optimal master plan. But that cannot be right, for two reasons. First, it is unlikely that there will ever be optimal master plans that are smaller than universal plans. If a master plan leaves some choices undetermined, it is likely that we can improve upon it by adding decisions regarding those choices. But as we have seen, it is not possible for real agents to construct universal plans, so that cannot be required for plan adoption.

The idea that rationality requires choosing optimal master plans is a holdover from classical decision theory. Classical decision theory envisages a kind of "ideal rationality" where an agent can survey all possible courses of action and choose an optimal one. But that is a computationally impossible ideal. Real rationality — the rules governing rational cognition in real agents operating in complex environments — must set different standards. Most work in AI has assumed that an agent can complete all relevant reasoning before deciding how to act. But outside of toy problems, that will never be the case. Assuming that the agent's reasoning about the world involves at least full first-order logic, and more likely some defeasible (nonmonotonic) reasoning about its environment, that reasoning will not produce a recursive set of conclusions and so will in general be non-terminating.[1] Even if the agent is only engaging in classical planning, if it has to reason about its environment to detect threats to causal links then the set of threats will not generally be recursive, and I showed in my (1998) that this makes the set of ⟨*problem,solution*⟩ pairs not even recursively enumerable. In general, reasoning will be non-terminating. There will be no point at which an agent has exhausted all possibilities in searching for plans. Despite this, agents must take action. They cannot wait for the end of a non-terminating search before deciding what to do, so their decisions about how to act must be directed by the best plans found to date — not by the best possible plans that *could* be found. The upshot is that plan adoption must be defeasible. Agents must work with the best knowledge currently available to them, and as new knowledge becomes available they may have to change some of their earlier decisions. If we only test agent designs on toy problems, even big toy problems, we are apt to be led to architectures that cannot handle this rather fundamental observation.

This point is fairly obvious, and yet it completely changes the face of decision-theoretic planning. The objective cannot be to find optimal master plans. First, non-terminating reasoning may produce better and better master plans without limit, so there may be no optimal master plans. Second, even if there were optimal master plans the agent would have no way of knowing it has found one until all of the non-terminating reasoning is completed. Planning and plan adoption must be done defeasibly, and actions must be chosen by reference to the current state of the agent's reasoning at the time it has to act rather than by appealing to the idealized but unreachable state that would result from the agent completing all possible reasoning and planning. Agents begin by finding good plans. The good plans are "good enough" to act upon, but given more time to reason, good plans might be supplanted by better plans.[2] The agent's master plan evolves over time, getting better and better, and the rules for rationality are rules directing that evolution, not rules for finding a mythical endpoint. Accordingly, a decision-theoretic planner should implement rules for continually improving the master plan rather than implementing a search for the endpoint, i.e., a search for optimal master plans. We might put this by saying that a decision-theoretic planner should be an *evolutionary planner*, not an optimizing planner. An evolutionary planner will be implemented as an infinite loop rather than a terminating search program. A program for evolutionary planning will systematically direct an agents "entire life" rather than just discrete segments aimed at local goals.

The upshot of all this is that rational decision making becomes a theory of how to construct local plans and use them to systematically improve the master plan. I call this *locally global planning*. As a first approximation, we might try to formulate locally global planning as follows:

> It is rational for an agent to adopt a plan iff its marginal expected value is positive, i.e., iff adding it to the master plan increases the expected value of the master plan.

However, for two reasons, this formulation is inadequate. The first reason is that adding a new plan may only increase the expected value of the master plan if we simultaneously delete conflicting plans. For example, suppose a robot truck has adopted one plan for getting gas, and then discovers a better plan (gas is cheaper at another station). The master plan cannot be improved by simply adding the new plan to it.

---

[2] This is reminiscent of Herbert Simon's (1955) concept of "satisficing", but it is not the same. Satisficing consists of setting a

That would result in the agent's going to two gas stations when in fact it can only fill its tank once. To improve the master plan it must adopt the new plan and simultaneously delete the earlier plan.

The second reason the preceding formulation is inadequate is that plans may have to be added in groups rather than individually. Recall again the example of transporting the gold and silver to a common destination in a truck. The plan to deliver the gold and silver on a single trip, by virtue of achieving both goals (and taking account of the possible damage to the truck), had a higher expected value than any single plan with which it competes, e.g., the plan to deliver the gold without delivering the silver. 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 silver on another trip.[3] So suppose the agent first adopts the plan to deliver both the gold and the silver on a single trip. Then it occurs to the agent that it could make two trips. The change it should make to the master plan at that point involves deleting the plan to deliver the gold and silver on a single trip, and adding two other plans — the plan to deliver the gold on one trip and the plan to deliver the silver on another trip.

In general, a change to the master plan may consist of deleting several local plans and adding several others. This includes as a special case that of modifying an a previously adopted local plan, because this can be identified with deleting the unmodified plan and replacing it with the modified plan. Where M is a master plan and C a change, let $M\Delta C$ be the result of making the change to M. We can define the *marginal expected value of a change* C to be the difference it makes to the expected value of the master plan:

$$\mathbf{MEV}(C,M) = \mathbf{EV}(M\Delta C) - \mathbf{EV}(M).$$

The *principle of locally global planning* can then be formulated as follows:

> It is rational for an agent to make a change C to the master plan M iff the marginal expected value of C is positive, i.e., iff $\mathbf{EV}(M\Delta C) > \mathbf{EV}(M)$.

This is my proposal for a theory of rational decision making. It answers the purely logical question of how plans should be evaluated in deciding what plans to adopt. I propose this principle as a replacement for classical decision theory and as a replacement for "simple" plan-based decision theory. It captures the basic insight that rational agents should guide their activities by considering the probabilities and utilities of the results of their actions, and it accommodates the observation that actions must often be selected as parts of plans, and the observation that optimality makes no sense for plans outside of toy examples. A decision-theoretic planner should be an evolutionary planner, not an optimizing planner. The principle of locally global planning addresses the logic of evolutionary planning. This proposal provides the starting point for building a decision-theoretic planner, because it tells us what we should want our implementations to rather than a terminating search for an optimal solution.

AI planning theorists are often more interested in writing programs than in carefully working out the theory of what the programs should do. The preceding remarks do not tell us how to actually build an evolutionary planner faithful to the principle of locally global planning. But what they do establish is that existing planning technologies are based on logically incorrect theories of decision making. When applied to the real planning problems faced by sophisticated autonomous agents in the real world, existing planning technologies will either be unable to provide solutions, or they will be prone to providing incorrect solutions.

I too think it is important to implement. I doubt that one can ever get the theory wholly right without testing it by implementing it. But there is such a thing as premature implementation. We need to work out the basics of the theory before we know what to implement. Unfortunately, the theory presented here is not yet a complete theory of decision-theoretic planning. For one thing, the criticisms I have leveled at simple plan-based decision theory are not the only things that I believe to be wrong with it. Elsewhere (Pollock 2002, 2003, 2006) I have raised other difficulties which must also be handled by a correct theory.

I am not yet ready to implement, because there is more theoretical work to be done. However, it may be useful to briefly sketch how I think an implementation might go, and indicate what further theoretical issues need to be addressed before such an implementation can be constructed. That will be undertaken in the next section. These remarks are, however, quite tentative, and are not the main point of the paper.

9. Incremental Decision-Theoretic Planning

This final section will make some first, tentative, steps towards constructing an implementation. At this point, they are highly speculative. In this I will be guided by my speculations about how planning works in human beings, but it is not part of my enterprise to defend the view that humans work in accordance with the principles I will describe. I think they do, but that is not really relevant to the AI enterprise. These remarks will not culminate in an implementation, because as remarked above, there is too much theoretical work remaining to be done.

The principle of locally global planning forms the basis for an "evolutionary" theory of rational decision making. It is a very general principle. An agent that directs its activities in accordance with this principle will implement the algorithm scheme diagrammed in figure 1. However, a complete theory of rational decision making must include more details. In particular, it must include an account of how candidate changes *C* are selected for consideration. This is an essential part of any theory of how rational agents should go about deciding what actions to perform.

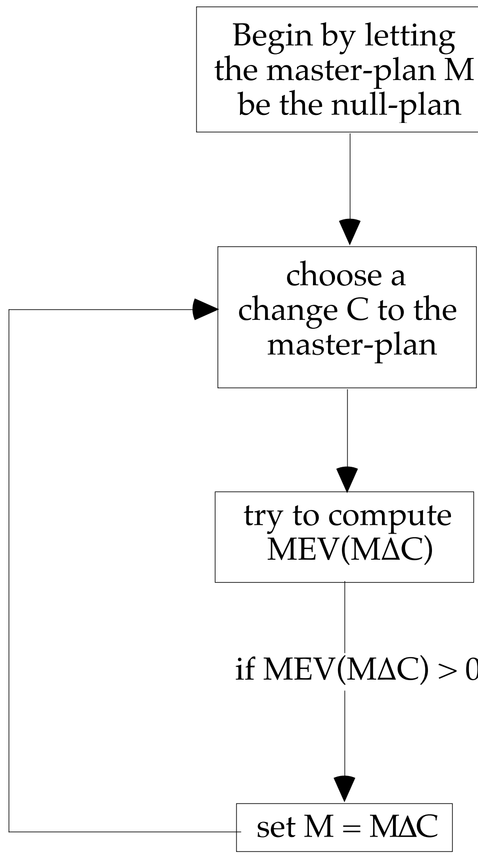

Figure 1. An algorithm schema for locally global planning

Algorithms that direct the search for possible changes to the master plan are *planning algorithms*. Let it be acknowledged immediately that there may be more than one good way of expanding the algorithm schema of figure 1. In deterministic planning, AI planning theory has produced a number of quite different planning algorithms, and the same thing may be true in decision-theoretic planning. The planning context and the cognitive resource bounds of the agent may determine which algorithms are best, with the result that there may be no single algorithm that is best for all agents and all times. So the remarks I will make here should be viewed as suggestions for one way in which decision-theoretic planning can work. This may not be the only way.

If we reflect upon human planning, it seems to have the general structure of what is called "refinement planning". In refinement planning, the planning agent begins by constructing a crude local plan, then searches for "flaws" in the plan, and refines the plan to remove the flaws. The flaws can be either internal to the plan or external to the plan. The external flaws derive from relations of "destructive interference" between the plan and the agent's master plan. Normally, refinement planning just refines the plan being constructed. In locally global planning, refinement planning must be generalized, allowing both the local plan and the master plan to be refined in light of the discovery of destructive interference between the two plans. This can produce multi-plan changes to the master plan. It is crucial to this general approach that the construction of multi-plan changes is driven by the construction of individual local plans. The and ask why we should expect this general approach to planning to lead to the incremental improvement of the master plan. This can be justified if we make four defeasible assumptions. I will refer to these as the *pivotal planning assumptions*:

Assumption 1: The process of constructing "crude local plans" produces plans that will normally have positive expected values.

Assumption 2: Ordinarily, the expected value of the result of merging two plans will be the sum of the expected values of the two plans.

Assumption 3: Computationally feasible reasoning procedures will reveal those cases in which the second assumption fails.

Assumption 4: There will be "repair techniques" that can often be used to modify either the local plans or the master plan in such a way as to remove the destructive interference leading to the failure of the second assumption without having to replan from scratch.

The justification of these assumptions will unfold below. Given the pivotal planning assumptions, the planning agent can begin the construction of the master plan by constructing a single local plan having a positive expected value, and take that to be the master plan. Then the agent can systematically construct further local plans with positive expected values, and on the basis of the second assumption it can be assumed defeasibly that each time one of them is merged with the existing master plan, the result will be a master plan with a higher expected value. On the basis of the third assumption, rational investigation will enable the agent to discover those cases in which the defeasible assumptions fail. This amounts to discovering destructive interference. The fourth assumption tells us that it will often be possible to refine the local plan and/or the master plan so as to avoid the destructive interference, thus leading to a modification of the original plans which, when merged, produce a master plan with a higher expected value than the original master plan. By proceeding in this way, a rational agent can systematically evolve progressively better master plans.

Now let us turn to the justification of the pivotal planning assumptions. This will simultaneously generate a better understanding of the way in which refinement planning might work in order to make the four assumptions true.

9.1 Goal-Directed Planning

The evolution of the master plan begins with the construction of what I have called "crude local plans".

directed planning, the agent adopts goals, and then searches for ways of achieving them by using goal-regression. In goal-regression, we observe that a goal could probably be achieved in a certain way if a certain condition were satisfied. If the condition is not currently satisfied, we adopt the condition as a new subgoal. In this way, we work backwards from goals to subgoals until we arrive at subgoals that are already satisfied. The details of goal-regression planning are complex, but see my (1998) for a general theory of deterministic goal-regression planning.

Goal-regression is not the only game in town. See Weld (1999) for a (now somewhat dated) survey of some alternative approaches to goal-directed planning in AI. I am skeptical, however, of being able to apply the current "high-performance" planning algorithms in domains of real-world complexity. This is because of the algorithms' stringent knowledge requirements — they make essential use of the closed world assumption, and that is totally unreasonable outside of toy problems. For example, if I want to adopt a kitten, in order to plan how to do that using any of these planners I would have to have explicit knowledge of every kitten in the world and all of its properties that are relevant to adoptability. Such an assumption is simply a non-starter for autonomous agents working in the real world. But it is absolutely essential to planners like GRAPHPLAN (Blum and Furst 1995, 1997) and BLACKBOX (Kautz and Selman 1996, 1998). There is no way to make them work without it.

Goals are simply valuable states of affairs — states of affairs that, were they achieved, would add to the overall value of the world (from the agent's point of view). Any valuable state of affairs can be chosen as a goal. In the cognitive architecture of a rational agent, that has the effect of initiating the search for plans for achieving the goal. The agent will probably not be in a position to construct plans for achieving many (even most) valuable states of affairs, so although they may technically be regarded as goals, they will have no effect on the agent's reasoning. They will simply be recorded as desirable, and then left alone unless the agent later encounters a way of achieving them.

The significance of goal-directed planning is that it produces local plans for the achievement of valuable goals. Assuming that the planning process will monitor the cumulative execution costs of the plan steps and not produce plans with high execution costs for lower-valued goals, goal-directed planning will produce plans that it is defeasibly reasonable to expect to have positive expected values. Then in accordance with the second pivotal planning assumption, it is defeasibly reasonable for the agent to expect to be able to simply add the new plans to the master plan and thereby improve the master plan.

Much work in AI planning theory has focused on goal-directed planning. However, this is not adequate as a general account of planning. This is most easily seen by thinking more carefully about goals. As I have described them, goals are "goals of attainment". That is, they are valuable states of affairs that can, in principle be brought about, thereby attaining the goal. Rational agents operating in an environment of real world complexity may continually acquire new goals in response to the acquisition of new knowledge about their environment. In section four I used the example of the planetary rover who, attainment that the agent had all along. For example, the robot truck's desire not to run out of fuel when it realizes that it is in danger of doing so derives from a general desire not to run out of fuel. But that general desire cannot be represented as a goal of attainment. In particular, it is not the goal of *never running out of fuel*. If that were its goal, then if it ever did run out of fuel that goal would become henceforth unachievable regardless of anything it might do in the future, and so there would be no reason in the future for trying to avoid running out of fuel. General desires are more like *dispositions* to form goals. As such, they are not themselves the target of planning.[4]

In different circumstances, an agent may be presented with different opportunities or exposed to different dangers. An opportunity is an opportunity to achieve an outcome having positive utility. This is accomplished by adopting the achievement of that outcome as a goal and then planning accordingly. Similarly, a danger is a danger that some outcome of negative utility will occur. That should inspire the agent to adopt the goal of preventing that outcome, and planning accordingly.

Planning for such newly acquired goals is not different in kind from planning for fixed goals, so this creates no problem for a theory of goal-directed planning. What is harder to handle, however, is the observation that we often plan ahead for "possible dangers" or "possible opportunities", without actually knowing that they will occur. For example, while trekking through tiger country, I may note that I might encounter a tiger, observe that if I do I will acquire the goal of not being eaten by him, and with that in mind I will plan to carry a gun. The important point here is that I am planning for a goal that I do not yet have, and I may begin the execution of the plan before acquiring the goal (that is, I will set out on my trek with a gun).

These are examples of what, outside of AI, is usually called "contingency planning". We plan for something that *might* happen, without knowing whether it actually will happen. If it does happen, then there will be a goal of the sort goal-directed planning normally aims at achieving. What makes it reasonable to begin planning before we actually acquire the information that generates the goals — just on the promise that we might acquire the information — is that such contingency plans can have positive expected values even before we acquire the information. So in decision-theoretic planning, such contingency planning is desirable, but the problem is to fit it into the framework of our planning theory. The goals are hypothetical, but the plans are real in the sense that we may not only adopt them but also start executing them before acquiring the goals at which they aim. Boutelier *et al* (1999) make a similar observation, taking it to be a criticism of goal-directed planning.

The best way to understand contingency planning is to think more generally about the function that goals play in planning. The rational agent is trying to find plans with positive expected values. Goals are of only instrumental value in this pursuit. On pain of computational infeasibility, we cannot use the British Museum algorithm and randomly survey plans until we find plans with high expected values. We must direct our planning efforts in some more intelligent fashion, and goals provide a mechanism for doing will have positive expected value. Hence goal-directed planning is a mechanism for finding plans with positive expected values. But what these examples show is that goal-directed planning by itself will not always suffice for finding such plans. It is too restrictive to require that planning must always be initiated by actual goals. It must be possible to plan for hypothetical goals as well. Hypothetical goals will become goals *if* something is the case. Escaping from the tiger will become a goal if I learn there is a tiger, and acquiring more fuel will become a goal if I learn that I am running low on fuel. To initiate planning for these hypothetical goals, it should suffice to know that the antecedent is sufficiently probable to make a plan for achieving the goal have a positive expected value. Of course, the latter is not something we can be certain about until we actually have the plan, so the *initiation* of planning cannot require knowing that. What we can do is take the value of the goal in the context of the planning to be the value the goal would have were the antecedent true, discounted by the probability the antecedent will be true. Then we can engage in planning "as if" the goal were a real goal, and evaluate the plan in terms of the attenuated value of the goal. If the resulting plan has a sufficiently high marginal expected value, we can add it to the master plan and begin its execution before the goal becomes real (e.g., we can carry a gun or top off the fuel tank).

The preceding remarks constitute only the briefest sketch of some aspects of goal-directed planning, but hopefully they will point the way to a general theory of goal-directed planning that can be incorporated into a more mature theory of locally global planning. These remarks are intended as a sketchy defense of the first pivotal planning assumption. I do not suppose for a moment that it is going to be an easy task to work out a theory of goal-directed planning for a decision-theoretic planner. All I suggest is that it may be useful to try. One thing is worth noting. Classical goal-directed planners have difficulty solving hard problems that can often be solved by high-performance satisfaction planners. As remarked above, the latter are, unfortunately, inapplicable to the planning of sophisticated autonomous agents in the real world, because of their reliance on the closed world assumption. So we may be stuck with some form of goal-directed planning. If planning problems are modeled as systematic and blind search through a tree of actions, the problem quickly becomes computationally too difficult. However, the problem may actually be easier in decision-theoretic contexts, because the search need not be blind. We have important information at our disposal for guiding search that we do not have in classical planning, namely probabilities and utilities. How much this helps remains to be seen, but I think it is at least plausible to suppose it will help a lot.

9.2 Presumptively Additive expected values

The second pivotal planning assumption was that, ordinarily, the expected value of the result of merging two plans will be the sum of the expected values of the two plans. This is equivalent to assuming defeasibly that the marginal expected value of a plan is equal to its expected value. In other words, when try to improve the master plan incrementally, by constructing local plans and adding them to it. This principle can be derived from principles I have defended elsewhere.

Let M be the master plan, and P be a local plan we are considering merging with M to produce the merged plan M+P. Both M and M+P will normally be nonlinear plans. The expected value of a nonlinear plan ought to be determined by the expected values of its linearizations. If a planning agent discovers that some linearizations have lower expected values than others, then he should add ordering constraints that preclude those linearizations. So it is reasonable to define the expected value of a nonlinear plan to be the minimum of the expected values of its linearizations. Ideally, all the linearizations should have the same expected value. A linearization L of M+P will result from inserting the steps of a linearization P^* of P into a linearization M^* of M. The steps of P^* will not be dependent on any of the steps of M^*, so the expected value of L will be the expected value of M^* plus the expected value of P^* *in the context of* L. By the latter, I mean that the probabilities and utilities employed in computing the expected value of P^* in that context will all be conditional on the previous steps of L having been attempted.

If it is defeasibly reasonable to expect the utilities and probabilities of elements of P^* to remain unchanged when they are made conditional on the previous steps of L having been attempted, it will follow that it is defeasibly reasonable to expect that $\mathbf{EV}(L) = \mathbf{EV}(M^*) + \mathbf{EV}(P^*)$, and hence $\mathbf{EV}(M+P) = \mathbf{EV}(M) + \mathbf{EV}(P)$. This divides into two separate expectations — that the utilities remain unchanged, and that the probabilities remain unchanged. I have argued elsewhere that both of these expectations are defeasibly reasonable. I will briefly sketch the arguments for these claims.

Direct inference is the kind of inference involved in deriving single-case probabilities from general probabilities. For example, it governs the kind of inference involved in inferring the probability that it will rain today from general probabilities of rain in different meteorological conditions. I proposed a general theory of direct inference in my (1990). In my (2002) and (2006) I showed that the general principles underlying direct inference imply a defeasible presumption of statistical irrelevance for single-case probabilities:

(IR) For any P,Q,R, it is defeasibly reasonable to expect that $\mathrm{prob}(P/Q\&R) = \mathrm{prob}(P/Q)$.

This is exactly the principle we need to justify the defeasible assumption that the probabilities relevant to the computation of the expected value of a plan do not change when the plan is merged with the master plan. Of course, the identity in (IR) will often fail in concrete cases, but the point of (IR) is that it is reasonable to expect it to hold unless we have a definite reason for thinking otherwise.

Turning to utilities, we can understand the conditional utility of a state of affairs S as the amount of value it adds to the circumstance C:

$$\mathbf{U}(S/C) = \mathbf{U}(S\&C) - \mathbf{U}(C)$$

parameters on the basis of the utility-measures of the individual parameters or small combinations of them, and that in turn would make decision-theoretic reasoning impossible. There must be at least a defeasible presumption that for any state of affairs S and circumstances C, $\mathbf{U}(S) = \mathbf{U}(S/C)$. One way to defend this is to think of the utility $\mathbf{U}(S)$ associated with a state of affairs S as the value caused by the agent's being in those circumstances. On this conception of value, being in circumstances S literally causes the associated quantity of value to exist. This conception of value is discussed in much more detail in my (2001) and (2006). It has the consequence that general principles for reasoning about values can be derived from analogous principles for reasoning about causation. I proposed the following as a general principle regarding causes:

Causal Irrelevance

If the agent has no reason to think otherwise, it is defeasibly reasonable to think that P does not cause Q.

In other words, we expect disparate events to be causally independent unless we have some concrete reason for thinking otherwise. Given the causal conception of value, this has the consequence that it is defeasibly reasonable to expect states of affairs to be value-neutral, i.e., to cause no change in the quantity of value produced by a situation. In other words, $\mathbf{U}(S/C) = \mathbf{U}(S)$. This is the assumption that we need regarding the utilities that play a role in computing $\mathbf{EV}(M+P)$. These remarks are admittedly very sketchy, but for a more sustained defense of this reasoning, see my (2006).

The second pivotal planning assumption follows from these two principles for reasoning defeasibly about probabilities and utilities.

9.3 Finding Decision-Theoretic Interference

The third pivotal planning assumption is that computationally feasible reasoning will reveal those cases in which the second assumption fails. The second assumption was that when two plans are merged into a single plan, the expected value of that composite plan will be the sum of the expected values of the constituent plans. When this assumption holds, let us say that the plans exhibit *decision-theoretic independence*. *Decision-theoretic interference* is the failure of decision-theoretic independence. 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.

Consider the ways in which decision-theoretic interference can arise. The expected value of a plan is determined by various probabilities and utilities. The probabilities are (1) the probabilities that goals will be achieved, side effects will occur, or execution costs will be incurred if certain combinations of plan steps are attempted, and (2) the probabilities that the agent can attempt to perform a plan step if certain change. Without going into detail, it is clear that algorithms can be designed for searching for decision-theoretic interference by looking for constituents of the master plan that will change these probabilities and utilities. For example, if a plan relies upon the probability prob(*goal*/*subgoal* & action A is performed), we can search for "underminers" P such that prob(*goal*/*subgoal* & action A is performed & P) $\neq$ prob(*goal*/*subgoal* & action A is performed) and such that the master plan contains a subplan that achieves P with some probability. This will be analogous to threat detection in deterministic goal-regression planning. However, it remains to be seen whether it can be done efficiently.

9.4 Repairing Decision-Theoretic Interference

The fourth and final pivotal planning assumption is that we will often be able to make relatively small changes to local plans or the master plan to avoid any decision-theoretic interference that is detected. This is analogous to what occurs in classical refinement planning, and many of the same repair techniques will be applicable. The simplest will consist of adding ordering constraints to $M+P$ to avoid interference. Another way of repairing the decision-theoretic interference is to add further steps that prevent the lowering of the probability. This is analogous to what is called *confrontation* in classical planning (Penberthy and Weld 1992). The upshot is that familiar ideas taken from classical planning will be applicable here as well. There may also be other ways of resolving decision-theoretic interference that do not correspond to techniques used in classical planning. That is a matter for further research. But this much is clear. The discovery of decision-theoretic interference need not cause us to reject our local plans altogether. We may instead be able to modify them in small ways to resolve the interference. Of course, this will not *always* be possible. Sometimes the interference will be irresolvable, and then we must reject one or more local plans and look for other ways of achieving some of our goals. However, there is no reason to expect this to make decision-theoretic planning generally intractable.

10. Conclusions

This paper investigates decision-theoretic planning for cognitively sophisticated autonomous agents operating in environments of real world complexity. The most important conclusion of the paper is that existing algorithms for decision-theoretic planning are based on a logically incorrect theory of rational decision making for autonomous agents. As such, they cannot possibly provide the basis for decision-theoretic planning in such agents. What is needed is an evolutionary planner that implements the principle of locally global planning. Existing planning algorithms may work fine for practical applications of limited scope that can be modeled as toy problems, but it is in principle impossible for them to scale up to the kind of planning a sophisticated autonomous agent must do in the real world.

decision theory, which selects actions one at a time rather than in packages (plans). To defend decision-theoretic planning, we must instead argue that classical decision theory is wrong. I did this by arguing that actions cannot be evaluated in isolation. We often have to choose plans rather than individual actions. This led to simple plan-based decision theory according to which it is rational to adopt a plan iff it is an optimal plan from a set of alternatives.

Simple plan-based decision theory fails because outside of toy problems there is no way to define optimality in such a way that it is reasonable to expect there to be optimal plans. The problems for optimality stem from two sources. First, in domains of real-world complexity, reasoning is non-terminating, producing potentially infinitely many plans and making it impossible to know one has found an optimal plan even if optimal plans exist. Second, because local plans can be extended indefinitely by merging them with other plans, there is no appropriate set of competitors to use in defining optimality.

A decision-theoretic planner should be an evolutionary planner rather than an optimizing planner. An evolutionary planner finds good plans, and replaces them by better plans as they are found. I have proposed the principle of locally global planning as the structure for an evolutionary planner. Locally global planning is based on two observations: (1) planning produces local plans; (2) local plans cannot be evaluated in isolation. They must be evaluated in the context of all the agent's other plans. This is accomplished by considering the contribution a local plan makes to the master plan. The result is that the objective of rational deliberation should be to find a good master plan and to be on the continual lookout for ways of improving the master plan. Planning becomes a non-terminating process rather than a terminating search for an optimal solution.

That is the main conclusion of the paper. To draw that conclusion, the paper could have stopped at the end of section eight. However, I thought it useful to make some preliminary remarks about how one might go about building an evolutionary planner implementing locally global planning. That is the topic of section nine. But those remarks are only intended as suggestions.

Acknowledgment

This work was supported by NSF grant no. IRI-9634106.

References

Aström, K. J.
1965 "Optimal control of Markov decision processes with incomplete state estimation", *J. Math. Anal. Appl.* **10**, 174-205.
Blum and Furst

Boutelier, Craig, Thomas Dean, and Steve Hanks

1999 "Decision theoretic planning: structural assumptions and computational leverage", *Journal of AI Research* **11**, 1-94.

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.

Israel, David

1980 "What's wrong with non-monotonic logic?" *Proceedings of the First Annual National Conference on Artificial Intelligence.* 99–101.

Jeffrey, Richard

1965 *The Logic of Decision*, McGraw-Hill, New York.

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.

Majercik, S.,

1999 "C-MAXPLAN: Contingent planning in the MAXPLAN framework", AAAI Spring Symposium on Search Techniques for Problem Solving Under Uncertainty and Incomplete Information, 83-88.

Majercik, S. and M. Littman,

1999 "Contingent planning under uncertainty via stochastic satisfiability", AAAI99, 549-556

Ngo, L., P. Haddawy, H. Nguyen

1998 "A modular structured approach to conditional decision-theoretic planning", *AIPS'98*, 111-118.

Onder, Nilufer, and Martha Pollack

1997 "Contingency selection in plan generation", *ECP97*.

1999 "Conditional, probabilistic planning: a unifying algorithm and effective search control mechanisms", AAAI 99.

Onder, Nilufer, Martha Pollack, and John Horty

1998 "A unifying algorithm for conditional, probabilistic planning", *AIPS1998 Workshop on Integrating Planning, Scheduling, and Execution in Dynamic and Uncertain Environments*.

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.

Peterson, Gilbert, and Diane J. Cook,

2000 "Decision-theoretic planning in the Graphplan framework", AIPS2000, Proceedings of the Workshop on Decision-Theoretic Planning 69-76.

1995 *Cognitive Carpentry*, MIT Press.

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

1998a "Perceiving and reasoning about a changing world", *Computational Intelligence* **14**, 498-562.

2001 "Evaluative cognition", *Nous* **35**, 325-364.

2002 "Causal probability", *Synthese* **132**, 143-185.

2002a "Defeasible reasoning with variable degrees of justification", *Artificial Intelligence* **133**, 233-282.

2003 "Rational choice and action omnipotence", *Philosophical Review* **111**, 1-23.

2003a "An easy 'hard problem' for decision-theoretic planners", OSCAR Project technical report, available at http://oscarhome.soc-sci.arizona.edu/ftp/publications.html.

2006 *Thinking about Acting: Logical Foundations for Rational Decision Making.* New York: Oxford University Press.

Ramsey, Frank

1926 "Truth and probability", in *The Foundations of Mathematics*, ed. R. B. Braithwaite. Paterson, NJ: Littlefield, Adams.

Reiter, Raymond

1980 "A logic for default reasoning". *Artificial Intelligence* **13**, 81–132, 1980.

Savage, Leonard

1954 *The Foundations of Statistics*, Dover, New York.

Simon, Herbert

1955 "A behavioral model of rational choice", *Qart. J. Economic*s, **59**, 99-118.

von Neumann, J., and Morgenstern, O.

1944 *Theory of Games and Economic Behavior*. New York: Wiley.

Weld, Daniel

1999 "Recent advances in AI planning", *AI Magazine* **20**, 55-68.

Williamson, Mike and Steve Hanks

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