

Copyright © 1981, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

A MODEL FOR PLANNING IN COMPLEX SITUATIONS

by

Robert Wilensky

Memorandum No. UCB/ERL M81/49

30 June 1981

Used

A MODEL FOR PLANNING IN COMPLEX SITUATIONS

by

Robert Wilensky

Memorandum No. UCB/ERL M81/49

30 June 1981

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

A Model for Planning in Complex Situations*

Robert Wilensky

Computer Science Division
Department of EECS
University of California, Berkeley
Berkeley, California 94720

ABSTRACT

A model of planning applicable to complex, commonplace activities is being developed. This model differs from previous approaches in that it is based on the following assumptions: (1) A planning agent must be able to infer its own goals in addition to being able to generate plans for these goals, (2) everyday planning is primarily concerned with reasoning about the interactions between plans and goals; (3) *meta-planning* (formulating knowledge about how to plan in abstract plans and goals, and having the planner use this knowledge to solve its own planning problems) is used as a driving principle; (4) *projection* (simulating hypothetical futures based on current plans and world knowledge) is used to infer goals and debug plans, and (5) planning knowledge should be equally available for understanding as well as for planning.

Coupled together, the mechanisms for implementing these features give rise to a system of considerable power. For example, the ability to infer one's own goals is needed in an autonomous planning agent since it must deal with unexpected situations. However, together with meta-planning knowledge and the ability to project hypothetical futures, this feature enables the planner to detect and reason about complicated goal interactions, anticipate problems with proposed plans, and make choices in the face of competing alternatives.

This model has been developed in detail for the detection and resolution of goal conflicts. In doing so, we postulate several goal conflict resolution strategies, or meta-plans, called RE-PLAN, CHANGE-CIRCUMSTANCE and SIMULATE-AND-SELECT. The structure and application of these meta-plans is explored in the context of both decision making and understanding the actions of other planners.

*Research sponsored in part by the National Science Foundation under grant MCS79-06543 and by the Office of Naval Research under contract N00014-80-C-0732.

A Model for Planning in Complex Situations

Robert Wilensky

Computer Science Division

Department of EECS

University of California, Berkeley

Berkeley, California 94720

1. Introduction - Some Tenets for a Theory of Planning

Much previous work on planning and problem solving has been concerned with either very specialized systems or with highly artificial domains (e. g., consider Fikes and Nilsson (1971), Newell and Simon (1972), Sussman (1975), Shortliffe (1976)). More recently, there has been an increase in attention given to planning in commonplace situations. For example, Rieger (1975b) has proposed a set of "common sense algorithms" for reasoning about everyday physical situations; Hayes-Roth and Hayes-Roth (1979) are concerned with how a person might schedule a day's activities; and Carbonell (1980) POLITIC's program reasons dogmatically about political decisions. On another front, Sacerdoti (1977) and McDermott (1978), while operating perhaps in the more traditional problem solving context, have proposed some powerful approaches to problem solving in general.

1.1. Everyday Planning is Reasoning about Interactions Between Goals

We have been developing a theory of planning that is concerned with reasoning about commonplace situations. A central tenet of this theory is that most of

the planning involved in everyday situations is primarily concerned with the interactions between goals. That is, planning for individual goals is assumed to be a fairly simple matter, consisting primarily of the straightforward application of rather large quantities of world knowledge. The complexity of planning is attributed to the fact that most situations involve numerous goals that interact in complicated ways.

Thus while traditional problem solving research has been concerned with finding the solution to a single, difficult problem (e. g., finding the winning chess move), most everyday problem solving consists of synthesizing solutions to fairly simple, interacting problems. For example, a typical everyday situation that involves the sort of planning we are interested in might be to obtain some nails, and also buy a hammer. The plan for each goal is straightforward: One simply goes to the hardware store, buys the desired item, and returns. The problem lies in recognizing that it is a terrible idea to execute these plans independently. Rather, the seemingly simple common sense plan is to combine the two individual plans, resulting in the plan of going to the hardware store, buying both items, and then returning.

Simple as this situation may be, most conventional planners are ill-equipped to handle it. Although some planning programs have mechanisms for removing redundancies from a plan, they generally lack a mechanism for even noticing this sort of interaction if these plans are derived from heretofore unrelated goals. Perhaps more importantly, the interaction between plans may have more complex ramifications. For example, if enough items are to be purchased at the hardware store, then a better plan might be to take one's car, while walking may do otherwise. Thus a good part of planning involves detecting the interactions between goals, figuring out their implications, and then deciding what to do about them.

1.2. Planners Must Infer Goals As Well As Act Upon Them

Before a planner can determine that there is an interaction between its goals, it must of course know what goals it possesses. In most planning systems, high-level goals are simply handed to the planner, often in the form of a problem to be solved. However, an autonomous planner must be able to infer its own goals based upon its overall mission together with the situation in which it finds itself.

For example, a planner may be given the job of maintaining a nuclear reactor. In addition to sustaining the generation of power, the robot may also be in charge of keeping the floors clean, preventing meltdowns, cleaning up dangerous spills, and maintaining itself. However, much of the time, most of these functions are inoperative. E. g., the robot need not be concerned with cleaning up a spill until one occurs. Thus it is desirable to build a planner that can recognize one of these situations when it occurs, and infer the goal it should have at that time. Thus the robot should infer it has the goal of cleaning up a spill when one occurs, of defending itself and the plant if they are attacked by terrorists, and of replenishing its resources if its power supply is low.

Moreover, goal inference interacts in important ways with other planning capabilities. For example, consider the situation above in which a planner is trying to get a hammer and nails from the hardware store. In accomplishing this task, the planner may end up at the hardware store with the hammer and nails. However, it now needs to realize that it should go home with these items. That is, if having the hammer and nails were instrumental to some other task, for example, fixing one's roof, the planner now needs to realize that its plan for obtaining these items has caused it to have the new goal of going home. While the plan for this goal is likely to be trivial, it is crucial that the goal is inferred in order for the entire plan to be effective.

The mechanism that infers goals is used in a similar fashion together with the next two features of this model to provide a facility for detecting and reasoning about the interactions between goals, for anticipating problems with proposed plans, and for making choices in the face of competing alternatives.

1.3. Meta-Planning is Used as the Driving Principle

The third salient feature of our theory is that it is based on *meta-planning*. By this I mean that the problems a planner encounters in producing a plan for a given situation may themselves be formulated as goals. These "meta-goals" can then be submitted to the planning mechanism, which treats them just like any other goals. That is, the planner attempts to find a "meta-plan" for this meta-goal; the result of successful application of this plan will be the solution to the original planning problem.

A typical example of a meta-goal is the goal RESOLVE-GOAL-CONFLICT. A planner would presumably have an instance of this goal whenever it detects that some of its "ordinary" goals are in conflict with one another. The meta-plans for this goal are the various goal conflict resolution strategies available to the planner.

Meta-planning is described in more detail in Wilensky (1980). Here, we give only a brief characterization of its main features and advantages.

Meta-goals are organized by *meta-themes*. These are very general principles of planning that describe situations in which meta-goals come into being. We summarize these briefly:

Meta-themes

- 1) DON'T WASTE RESOURCES
- 2) ACHIEVE AS MANY GOALS AS POSSIBLE
- 3) MAXIMIZE THE VALUE OF THE GOALS ACHIEVED
- 4) AVOID IMPOSSIBLE GOALS

As an example of how these function, the meta-theme "ACHIEVE AS MANY GOALS AS POSSIBLE" is responsible for detecting goal conflicts. That is, if the planner intends to perform a set of actions that will negatively interact with one another, this meta-theme causes the planner to have the goal of resolving the conflict. If this meta-goal fails, i. e., the planner could not find a way to resolve the conflict, then the meta-theme "MAXIMIZE THE VALUE OF THE GOALS ACHIEVED" springs into action. This meta-theme sets up the goal of arriving at a scenario in which the less valuable goals are abandoned in order to fulfill the most valuable ones. The details of the meta-plans involved in these processes are described in length in the last two sections of this paper.

Meta-planning has a number of advantages over other approaches to planning; these advantages are summarized below:

1.3.1. Meta-planning knowledge can be used for both planning and understanding

As meta-goals and meta-plans are declarative structures in the same sense as are ordinary goals and plans, they may be used to understand situations as well as plan in them. Thus an understander with access to this knowledge would be able to interpret someone's action as an attempt to resolve a goal conflict, for example. In contrast, planning programs that have the equivalent knowledge embedded procedurally would not be able to conveniently use it to explain someone else's actions.

1.3.2. The same planning mechanism can apply to more difficult tasks.

Meta-planning knowledge generally embodies a set of strategies for complicated plan interactions. By formulating this knowledge in terms of goals and plans, the same planning architecture that already exists for simpler planning can be used to implement more complicated planning involving multiple goals, etc.

1.3.3. More general resolution of traditional planning problems

Traditional planners usually treat problems such as goal conflicts by special purpose means - by the introduction of critics, for example (Sussman 1975, Sacerdoti 1977). This is equivalent to having the general problem solver consult an expert when it gets in trouble. The meta-planning allows the general problem solver to call a general problem solver (itself) instead. Thus all the power of such a system can be focussed on planning problems, rather than just relying on a few expert tactics. Of course, all the specific knowledge usually embodied in critics would still be available to the general problem solver. But the meta-planning model allows this knowledge to interact with all other knowledge as it now take take part in general reasoning processes.

1.3.4. Representational advantages

The meta-planning model also provides more flexibility when no solution can be found. Since a meta-goal represents the formulation of a problem, the existence of the problem may be dealt other than its being fully resolved. For example, the problem solver may simply decide to accept a flawed plan if the violation is viewed as not being too important, or decide to abandon one of the goals that it can't satisfy. By separating solving the problem from formulating the problem, the problem may be accessed as opposed to treated, an option that most other problem solving models do not allow.

1.4. Projection is Used to Infer Goals and Debug Plans

The fourth significant feature of our planning model is that it is based on *projection*. That is, as the planner formulates a plan for a goal, the execution of this plan is simulated in a hypothetical world model. Problems with proposed plans may be detected by examining these hypothetical worlds.

Projection not only enables the planner to find problems with its own plans, but it also enables it to determine that a situation merits having a new goal. For example, sensing an impending danger requires the planner to project from the current state of affairs into a hypothetical world which it finds less desirable. Having done this projection, the planner can infer that it should have the goal of preventing the undesirable state of affairs from coming into being.

Projecting hypothetical realities also allows a general "goal detection" mechanism to work for meta-goals as well as for "ordinary" goals. When proposed plans for goals are projected, interactions will appear in the hypothetical world. Since such interactions generally indicate that some important planning principle is not being adhered to, the occurrence of this hypothetical negative interaction is usually a signal to the planner to achieve some particular meta-goal.

Working with projected universes entails some liabilities as well, as does the notion of meta-planning and of using highly declarative representations. However, our claim is that the prices associated with these ideas are prices that must be paid anyway. By putting them together in the manner described here, a deal of power is obtained for no additional cost.

1.5. Planning Knowledge Should be Equally Available for Understanding

The last tenet of our theory of planning is that it should be equally usable by both a planner and an understander. This is essentially an extension of the

point made by Rieger (1975a). That is, while a planner uses its planning knowledge to bring about a desired state of affairs, an understander may need to use this same knowledge to comprehend the actions of a person it is watching or of a character about whom it is reading. For example, a planner with the goal of keeping fit might take up jogging; an understander might use the same knowledge to infer that someone who has taken up jogging may have done so because he had the goal of staying in shape. Planning and understanding are rather different processes, and this will of course be reflected in our planning and understanding mechanisms. However, our theory of planning specifies that knowledge should be represented in a fashion so that it is usable by either mechanism. In particular, we have formulated meta-planning knowledge in terms of explicit goals and plans so that it might be used to understand a situation as well as to act in one.

In the next section, I discuss the general structure of a planning mechanism based on these assumptions. This is the structure used in PANDORA (Plan ANalysis with Dynamic Organization, Revision, and Application), a planning system now under construction at Berkeley. The sections following show how these mechanisms function together in reasoning about *goal conflict situations*. As we have noted, we intend these ideas to be applicable to understanding as well as planning, and in fact, they are being used in a new implementation of PAM, a plan-based story understanding system. While we do not discuss the structure of PAM here, the analysis of goal conflicts is presented in a form in which its use in understanding as well as planning may be seen.

2. The Design of a Planner Based on Meta-planning

This section describes the overall architecture of a planner based on the tenets just considered. The planner is composed of the following major com-

ponents:

1) The Goal Detector

This mechanism is responsible for determining that the planner has a goal. The goal detector has access to the planner's likes and dislikes, to the state of the world and any changes that may befall it, and to the planner's own internal planning structures and hypothetical world models. The goal detector can therefore establish a new goal because of some change in the environment, because such a goal is instrumental to another goal, or in order to resolve a problem in a planning structure that arises in a hypothetical world model.

2) The Plan Generator

The plan generator proposes plans for the goals already detected. It may dredge up stereotyped solutions, it may edit previously known plans to fit the current situation, or it may create fairly novel solutions. The plan generator is also responsible for expanding high-level plans into their primitive components to allow execution.

3) The Executor

The executor simply carries out the plan steps as proposed by the plan generator. It is responsible for the detection of errors, although not with their correction.

The importance of the goal detector should be emphasized. As was mentioned above, most planner merely have their goals stated to them, but a planning system needs to infer its own goals for a number of reasons: An autonomous planner needs to know when it should go into action; it should be able to take advantage of opportunities that present themselves, even if it doesn't have a particular goal in mind at the time; and it should be able to protect itself from

dangers from its environment, from other planners, or from consequences of its own plans.

The goal detector operates through the use of a mechanism called the *Noticer*. The Noticer is a general facility in charge of recognizing that something has occurred that is of interest to some part of the system. The Noticer monitors changes in the external environment and in the internal states of the system. When it detects the presence of something that it was previously instructed to monitor, it reports this occurrence to the source that originally told it to look for that thing. The Noticer can be thought of as a collection of IF-ADDED demons whose only action is to report some occurrence to some other mechanism.

Goals are detected by having themes and meta-themes asserted into the Noticer with orders to report to the goal detector. *Theme* is a term used by Schank and Abelson (1975) to mean something that gives rise to a goal; a meta-theme, similarly, is responsible for generating meta-goals. For example, we can assert to the noticer that when it gets hungry (i. e., when the value of some internal state reaches a certain point), the planner should have the goal of being not hungry (i. e., of changing this value), or that if someone is threatening to kill the planner, that the planner should have the goal of protecting its life. On the meta-level, we might assert that if a goal conflict comes into existence, then the planner should have the meta-goal of resolving this conflict.

2.1. Using the Goal Detector to Spot Goal Interactions

The presumption of a goal detector coupled with meta-planning creates a system of considerable power. For example, no separate mechanism is required for detecting goal conflicts or for noticing that a set of proposed plans will squander a resource. The need to resolve conflicts or conserve resources is expressed by formulating descriptions of the various situations in which this

may occur, and the appropriate meta-goal to have when it does. By asserting these descriptions into the Noticer to detect meta-goals, goal conflicts and other important goal interactions are handled automatically.

The planner component of our model itself consists of three components:

1) The Proposer,

which suggests plausible plans to try

2) The Projector,

which tests plans by building hypothetical world models of what it would be like to execute these plans

3) The Revisor,

which can edit and remove parts of a proposed planning structure

The Proposer begins by suggesting the most specific plan it knows of that is applicable to the goal. If this plan is rejected or fails, the Proposer will propose successively more general and "creative" solutions. Once the Proposer has suggested a plan, the Projector starts computing what will happen to the world as the plan is executed. The difficult problems in conducting a simulation involve reasoning about "possible world" type situations which are not amenable to standard temporal logic (McCarthy and Hayes, 1969). However, we finesse this issue by defining hypothetical states in terms of what the planner thinks of in the course of plan construction. In other words, our solution is to let the system assert the changes that would be made into a hypothetical data base, in the meantime letting the goal detector have access to these states. Thus if the plan being simulated would result in the planner dying, say, this would constitute a hypothetical undesirable state, which might trigger further goals, etc.

As the Projector hypothetically carries out the plan, and other goals and meta-goals are detected by the goal detector, the original plan may have to be modified. This is done by explicit calls to the Revisor, which knows the plan

structure and can make edits or deletions upon request. The modified plan structure is simulated again until it is either found satisfactory or the entire plan is given up and a new one suggested by the Proposer.

Actually, the function of the Projector is somewhat more pervasive than has so far been described. The Projector must be capable of projecting current events into future possibilities based both on the intentions of the planner and on its analysis of those events themselves. For example, if the planner sees a boulder rolling down the mountain, it is the job of the Projector to project the future path that the boulder will traverse. If the projected path crosses that of the planner, for example, a preservation goal should be detected. Thus the Projector is a quite powerful and general device that is capable of predicting plausible futures.

3. Reasoning about Goal Conflicts

In the next two sections we give a more detailed analysis of one particular part of our planning model, namely, the resolution of goal conflicts. The problem is important in its own right; however, the presentation that follows is aimed at demonstrating the kind of "strategy architecture" to which the model is conducive. In particular, the section illustrates a number of important meta-goals and the meta-plans for them, and describes how they would be invoked and utilized by the model. The section also emphasizes the utility of meta-planning for the application of planning knowledge to understanding goal conflicts as well as to planning for them.

Since it is desirable to achieve all of one's goals, a planner faced with a goal conflict will probably attempt to resolve that conflict. We express this by saying that the state of having a goal conflict is a situation that causes the meta-theme "ACHIEVE AS MANY GOALS AS POSSIBLE" to become active. In such a situation, this meta-theme creates the meta-goal RESOLVE-GOAL-CONFLICT. This is a

meta-goal because resolving the conflict can be viewed as a planning problem that needs to be solved by the creation of a better plan. In this formulation, the resolution of the goal conflict is performed by the execution of a meta-plan, the result of which will be a set of altered plans whose execution will not interfere with one another.

The knowledge needed to replan around a goal conflict is quite diverse, and may depend upon the particular goals in question and on the nature of the conflict. However, the meta-plans with which this knowledge is applied are rather general. To see why, it is necessary to ask how it is possible for goal conflicts to be resolved at all. There appear to be two ways in which goal conflicts can come about that determine how they may be resolved:

- 1) The conflict detected is based on the plans for one's goals, rather than on the goals themselves. In this case, it may be possible to achieve the goals by other, non-conflicting plans.
- 2) The conflict depends upon some additional circumstance or condition beyond the stated goals or plans. The conflict might therefore be resolved if this circumstance is changed.

We therefore define two very general meta-plans, RE-PLAN and CHANGE-CIRCUMSTANCE. Of course, to be effective, we need to supply these meta-plans with more information; if we use RE-PLAN blindly, for example, we might end up enumerating all possible plans for each conflicting goal, although many of these plan combinations will present the same problem that caused the original goal conflict.

3.1. RE-PLAN

There are a number of different re-planning strategies applicable to goal conflict situations. They are given here in order of decreasing specificity. This

is in accordance with our belief about the order in which such plans would actually be used, i. e, the most specific one first, then progressively more general ones, until a satisfactory set of plans is found. In this respect, meta-plans are entirely analogous to ordinary plans insofar as the planning process is concerned.

The order of plan application is just a corollary of the First Law of Knowledge Application - "Always use the most specific piece of knowledge applicable".

3.1.1. USE-NORMAL-PLAN applied to resolving goal conflicts

The most specific re-planning strategy is likewise analogous to the planning strategy for ordinary goals, namely, find a normal plan. A normal plan in the case of goal conflict is to find a stored plan specifically designed for use in a goal conflict between the kinds of goals found in the current situation. For example, consider the following situation:

- (1) Mary was very hungry, but she was trying to lose some weight. She decided to take a diet pill.

In (1), there is a conflict between the goal of losing weight and satisfying hunger, as the normal plan for the latter goal involves eating. The RE-PLAN meta-plan is used, and the USE-NORMAL-PLAN strategy applied. The normal plan found that is applicable to both goals is to take a diet pill.

Just as many objects are functionally defined by the role they play in ordinary plans, so some objects are functionally defined by the role they play in plans aimed at resolving specific goal conflicts. Thus a diet pill is an object functionally defined by its ability to resolve the conflict between hunger and weight loss; a raincoat is defined by the role it plays in preventing wetness when one must go outside. In fact, a great deal of mundane planning knowledge appears to consist of plans for resolving specific types of goal conflicts.

3.1.2. Intelligent use of TRY-ALTERNATE-PLAN to find non-conflicting set

A general planning strategy that is applicable when a plan cannot be made to work is to try another plan for that goal. In the case of resolving goal conflicts, this means that alternative plans for each conflicting goal can be proposed until a set is found that are not in conflict. As noted above, this may be costly, but it will only be tried when no canned conflict resolution plan has been found. Moreover, the plan can provide some intelligent ways of proposing new alternatives that may help keep costs down.

For example, consider the following situation:

- (2) John was going outside to pick up the paper when he noticed it was raining. He looked for his raincoat, but he couldn't find it. He decided to get Fido to fetch the paper for him.

Here, John first thought to walk outside, but then found that this would cause a conflict. As his normal plan for resolving this conflict failed, John tried proposing other plans, looking for ones that wouldn't entail his getting wet. Since getting the dog to fetch the paper is such a plan, and since John presumably doesn't care if Fido gets wet, this plan is adopted.

The meta-planning strategy used here is called TRY-ALTERNATIVE-PLAN. The difference between using this meta-plan and blind generate and test strategies is that some control can be exerted here over exactly what is undone and what is looked for as a replacement. For example, the backtracking here need not be chronological or dependency-directed, but can be *knowledge-directed*. That is, rather than undo the last planning decision, a planning decision related to either goal can be undone, possibly based on an informed guess.

In addition, when fetching a new plan, it may be possible to specify in the fetch some conditions that the fetched plan may have to meet without actually testing that plan for a conflict. For example, in the case of getting the newspaper when it is raining, we can ask for a plan for getting something that doesn't

involve going outside. That is, we can look for a plan for one goal that does not contain an action that led to the original conflict. If our memory mechanism can handle such requests, then we can retrieve only those plans that do not conflict in the same way that the original plan does.

In order for this to work, TRY-ALTERNATIVE-PLAN needs to know what part of a plan contributed to the goal conflict so it can look for a plan without this action. This generally depends upon the kind of conflict. We can formulate this within the meta-planning framework by defining a meta-plan called MAKE-ATTRIBUTION. Here, MAKE-ATTRIBUTION is used as a subplan of the TRY-ALTERNATIVE-PLAN meta-plan, although we shall make other uses of it below. TRY-ALTERNATIVE-PLAN first asks MAKE-ATTRIBUTION to specify a cause of the problem, and then fetches a new plan without the objectional element in it.

TRY-ALTERNATIVE-PLAN can also control how far up the proposed goal-subgoal structure it should go to undo a decision. Thus, if no alternative plan for a goal can be found, the goal itself can be questioned if it is a subgoal of some other plan. For example, consider the following scenario:

- (3) John was going to get the newspaper when he noticed it was raining. He decided to listen to the radio instead.

Here the entire subgoal of getting the newspaper was eliminated. Since this was apparently a subgoal of finding out the news, the alternative plan of listening to the radio can be substituted. Once again, MAKE-ATTRIBUTION is used to propose a plan that doesn't involve an unwanted step. The difference between this and the last case is that here a plan lying above the conflicting goal is re-planned.

3.2. CHANGE-CIRCUMSTANCE

In addition to the RE-PLAN meta-plan, the other general goal conflict resolution strategy is to change the circumstance that contributes to the conflict. This is actually more general than RE-PLAN, because it may be applicable to

conflicts where the goals themselves exclude one another, whereas RE-PLAN requires the conflict to be plan-based.

CHANGE-CIRCUMSTANCE can resolve a goal conflict by altering a state of the world that is responsible for the goals conflicting with one another. Once this has been achieved, the original set of plans may be used without encountering the original problem.

For example, consider the following situations:

- (4) John had a meeting with his boss in the morning, but he was feeling ill and wanted to stay in bed. He decided to call his boss and try to postpone the meeting until he felt better.
- (5) John wanted to live in San Francisco, but he also wanted to live near Mary, and she lived in New York. John tried to persuade Mary to move to San Francisco with him.

In (4), John's conflict is caused by his plan to attend the meeting and his plan to stay home and rest. These plans conflict because of the time constraints on John's meeting, which force the plans to overlap; the plans require John to be in two places at once, so they cannot be executed simultaneously. If the time constraint on attending the meeting were relaxed, however, then the conflict would cease to exist. Thus rather than alter his plans, John can seek to change the circumstances that cause his plans to conflict by attempting to remove the time constraint that are a cause of the difficulty.

In (5), the conflict is between living in San Francisco and being near Mary, who is in New York. The basis for this exclusion involves the location of San Francisco and of Mary. However, if one of these locations were changed so that the distance between them were reduced, then the state would no longer exclude one another. Thus John can attempt to change Mary's location, while still maintaining his original goals.

To decide what circumstance to change, a planner once again needs to

analyze the cause of the conflict. Thus CHANGE-CIRCUMSTANCE first requires the use of MAKE-ATTRIBUTION to propose a candidate for alteration. As was the case for RE-PLAN, MAKE-ATTRIBUTION requires access to detailed knowledge about the nature of negative goal interactions in order to find a particular circumstance with which to meddle. An analysis of such interactions appears in Wilensky (1978).

4. Goal Abandonment

When attempts to resolve a goal conflict are unsuccessful, a planner must make a decision about what should be salvaged from this situation. The choices here involve abandoning some the goals completely while pursuing some select subset, or modifying some of the goals so as to avoid the conflict. As was mentioned above, we consider the latter alternative, goal modification, to be a form of goal abandonment rather than conflict resolution because the strategy does not result in the fulfillment of the original set of goals.

In terms of meta-planning, we can describe goal abandonment situations as follows. The inability to achieve a RESOLVE-GOAL-CONFLICT meta-goal results in the planner having this failed meta-goal on his hands (or more precisely, in his representational space). Having a failed Resolve-Goal-Conflict meta-goal is a condition that triggers the meta-theme Maximize-the-Value-of-the-Goals-Achieved. This triggering condition causes this meta-theme to invoke a new meta-goal, called CHOOSE-MOST-VALUABLE-SCENARIO. This goal is satisfied when the relative worth of various achievable subsets of the conflicting goals is assessed, and the subset offering the greatest potential yield determined.

To achieve this meta-goal, we postulate a SIMULATE-AND-SELECT meta-plan. This plan proposes various combinations of goals to try, and computes the worth of each combination. The most valuable set of goals is returned as the scenario most worth pursuing.

4.1. The SIMULATE-AND-SELECT meta-plan

The SIMULATE-AND-SELECT meta-plan has a rich structure. To begin with, it makes a number of presumptions about evaluating the cost and worth of goals and of comparing them to one another. For the purposes at hand, we shall not dwell on exactly how the evaluation is done. Partly this is because the details of how to do this are not completely clear; moreover, they are not crucial for the upcoming discussion. We presume that values can be attributed to individual states in isolation *ceteris paribus*, and that the value of a set can be computed from its parts. This does not presume that the computation is simple; indeed, it may involve the consultation of large amounts of world knowledge. However, we do assume that all values can be made commensurable.

This being said, the SIMULATE-AND-SELECT meta-plan has in effect two distinct options. The first is quite straightforward. It simply involves construction maximal achievable (i. e., non-conflicting) subsets from among the conflicting goals, and evaluating the net worth of each one. Since we are generally dealing with two goals in a conflict, this means just evaluating the worth of one goal and comparing it to the value of the other. Thus if having the newspaper is deemed more valuable than getting wet, then the planner walks outside to get the newspaper and allows himself to get soaked. Alternatively, a reader trying to understand someone else's behavior would use knowledge about this meta-plan to make inferences about their value system. If we observe John risking getting wet into to get his morning paper, then we conclude that having his paper is worth more to him than staying dry.

4.1.1. Goal Abandonment versus Goal Substitution

Sometimes, goal abandonment is operating in a sort of disguise. For example, consider the following stories:

- (6) John wanted to see the football game, but his wife said she would divorce him if he watched one more game. John settled for watching the tennis match together with his wife.
- (7) John wanted to live in San Francisco, but he also wanted to live near Mary, and she lived in New York. John decided he could probably find another girlfriend in San Francisco.

In (6), John must resolve a conflict between his goal of watching the football game and preserving his marriage. On the surface, it appears as if John modifies the goal of watching the football game to watching tennis instead, an activity that John's wife apparently does not object to. Similarly, in (7), John abandons his goal of maintaining his relationship with Mary, and appears to substitute it with the goal of having a similar relationship with somebody else.

Schank and Abelson (1977) describe such situations as instances of *goal substitution*. According to them, goal substitution is a strategy applicable whenever a goal is blocked. They state the rule that "following blockage, the maximum feasible number of valued characteristics of the goal object should be preserved." As they point out, goals can be "stacked" within a hierarchy of many levels, so for example, wanting to watch the Giants can be an instance of wanting to watch a football game, which in turn is an instance of wanting to watch a sporting event, and so on. Thus if a goal is to be substituted, it is desirable to replace it with a goal occupying as similar a place in the hierarchy as the goal it is replacing. So watching the Jets is likely to be a better substitution for watching the Giants than watching tennis, which is in turn a better substitute than operating a hydraulic ram.

However, we offer an analysis that is somewhat at odds with theirs. The main problem with Schank and Abelson's formulation is based on a subtle, but important distinction. To understand this problem, consider that the question of *why* one goal may be substituted for another. That is, how is it possible that a "similar" goal can be substituted for the original goal? If a goal is truly

something that a planner wants to achieve, then why should some other, different state do as well? The answer that Schank and Abelson provide is that the original goal is really just an instantiation of a more general goal; therefore another instantiation will do.

The problem here is that we are forced to conclude that the original goal was not a goal at all, but merely a plan for achieving some real goal. This is clearly the case some of the time, as the following example illustrates:

- (8) John needed a pen to sign a letter. When he couldn't find his own, he asked Bill if he could borrow his.

In this example, John at one point seems to have the goal of possessing his pen. When this goal fails, he replaces it with the goal of possessing Bill's pen. On the surface, one goal appears to be substituted by another, similar goal. However, strictly speaking, it is unlikely that John really had the goal of possessing his pen; most likely, his goal was "possess some pen" or possibly, "possess a writing implement". But such a goal is too abstract to be achieved directly. While this goal is not specific to the particular writing implement selected, the real world has only individual objects. At some point, one such object must be designated in order for a plan to be executed. That is, in order to have any pen, a planner eventually plan to obtain one particular object.

Because this sort of narrowing down is really part of the planning process rather than part of a planner's goal structure, we refer to it as *plan specification*. Specifically, we postulate a PLAN-SPECIFICATION meta-plan that makes a plan work by supplanting generic descriptions with descriptions of more specific entities (This is in contrast to *goal specification*, a term used by Schank and Abelson to make the conditions of a goal more elaborate. The effects of these two processes is quite different. Plan specification enables the fulfillment of a goal by producing a plan more capable of execution. Goal

specification makes fulfillment of a goal more difficult by posing additional constraints that a state must meet to achieve the desired goal).

We can see at once that John wanting his pen in story (8) is almost certainly an instance of plan specification. His goal is to have a writing implement, and as a way of achieving this goal, he used the PLAN-SPECIFICATION meta-plan to produce the plan of obtaining his particular pen. As he failed to execute this plan, he re-specifies by selecting another instance of the generic class of object which he is seeking.

Thus the situation in story (8) obeys Schank and Abelson's hierarchy rules, but it is unrelated to goal substitution; it merely supplants one plan to another and it is therefore simply an instance of re-planning for a goal. In a real goal substitution situation, an actual goal rather than a plan-specified instance of a goal must be substituted. Otherwise this situation is no different than substituting the plan of asking with the plan of begging when the attempt to execute the plan successfully fails.

Most of Schank and Abelson's candidates for goal substitution are actually simple re-planning instances. For example, consider that goal of "enjoy eating at a Chinese restaurant". Schank and Abelson would suggest that this is "dominated" by the goal "enjoy eating good food", and therefore if it fails, then eating at a French restaurant may be substituted as it preserves these salient features. But in fact, these features are not a property of this goal at all. To see this, note that the goal dominating this one may just have easily been "partake in the Chinese New Year celebration". Now if eating at a Chinese restaurant were blocked, then a reasonable substitute might be "watch the Chinese New Year parade". However, "enjoy eating at a French restaurant" would fail utterly. The explanation for this difference is of course that in both cases, eating at a Chinese restaurant is not a goal, but rather a plan for another goal.

Of course, "enjoy eating at a Chinese restaurant" may very well be a goal, but in that case, it would be unsubstitutable. Since this being a real goal would be most unusual, this may be difficult to see. However, in general such goals are not uncommon. For example, the goal "treat my wife to dinner for her birthday" will hardly be amenable to substitution by replacing my wife with my son, or with her twin sister. And if I really wanted to take her out to dinner, buying her a ring will just not do.

According to this analysis, goal substitution is just not possible. The semantics of what it means to be a goal are such that if it were substitutable, then our original specification must have been incorrect in the first place. If I said my goal were "treat my wife to dinner for her birthday", but treating her to a fancy lunch were an acceptable substitute, then I must have meant that my original goal were really "treat wife to fancy meal", and the "treat her to dinner" merely a plan.

However, there is still a problem. For if I really desired to treat my wife to dinner, how can it be that if this goal fails, then buying her a ring still seems more appropriate than going to a football game with my best friend? The answer is that when a planner truly has a goal such as taking one's spouse to dinner, he often has other related goals as well. For example, he is likely to have the goal "enjoy a good meal" and "buy wife a present" in addition to the particular desire to take her to dinner. While these goals are in principle unrelated, it happens to be the case that the plan for the latter one is in fact equivalent to the achievement of the other two. That is, in addition to being a goal in its own right, taking one's spouse to dinner is a plan for enjoying a good meal and for buying her a present. Rather than an instance of goal substitution, replacing "take wife to dinner" with "buy her a ring" is actually an abandonment of this goal and of "enjoy good meal", and a re-specification of the "buy wife a present" goal with

the plan "buy wife a ring" instead of "take wife to dinner".

In sum, goal substitution does not exist. Either we have a case of re-specification, in which case the original "goal" was actually a plan, or we have an instance of goal abandonment in which one of the goals abandoned also happened to be equivalent to a plan for another goal, which was subsequently re-specified into some new plan.

As an example, in story (6) above, suppose we assume John's goal were really to watch the football game. In this case, he actually abandons this goal completely. However, he might have another goal, "enjoy leisure activity" for which watching the football game is a plan. Thus watching tennis is a new plan specification for this goal, and the original goal that conflicted with preserving his marriage is given up entirely.

Thus notion of goal abandonment is sufficient to handle situations like (6) and (7), once we realize that there are likely to be a number of goals active here at once. Of course, making such a realization requires knowledge about the normal goal structures people are likely to have. In the case of story understanding, for example, if we are told that a planner desires such and such a state, we use our knowledge of normalcy to infer what the goal really is. Unusually specific or general goals must be described as such explicitly to overcome this default inference. Thus "John wanted to go to New York" is interpreted as the goal "John wanted to be in New York", as going in itself is not normally desirable; if we hear "John wanted a hamburger", we interpret this as "John wanted some generic hamburger", as individual hamburgers are rarely distinguished by one's goals.

This tendency toward normality is why example (7) seems somewhat unusual - one is supposed to highly value the particular individual involved in a relationship, thus realizing that the goal here is actually more abstract comes

as somewhat of a surprise.

4.1.2. Partial Goal Fulfillment

Goal abandonment can therefore account for a diverse number of complicated situations. However, there is another set of situations that we cannot yet cope with. Consider once again the example of fetching the newspaper in from the rain, in which the original goals are to get the newspaper and to remain dry. Rather than abandon either goal complete, a reasonable alternative is to try to reduce the degree to which one gets wet as much as possible. A plan for remaining as dry as possible while moving through the rain is to run as fast as one can. This plan satisfies one goal entirely, and another to a degree. The total value of this scenario is likely to be greater than the value of staying dry but not getting the paper, and since the other abandonment possibility (getting the paper but getting soaked) is clearly worse than this (i. e., getting the paper but getting less soaked), the scenario involving partial fulfillment is likely to be adopted.

Partial goal fulfillment is a general principle that is applicable to all goals that involve scalar values. If one wants not to be hungry, being less hungry is not as good as not being hungry at all, but it is better than not being less hungry. That is, it is implicit in what it means to be a goal that states closer to the goal have a value less than the goal but greater than those further from the goal.

In the cases where this does not appear to hold, it is usually because the actual goal structure has been incorrectly identified. For example, if it were raining acid rain whose very touch were deadly, getting just a little wet would not be better than getting very wet. But this is because the real goal in this situation is not "don't get wet", but "don't get killed". Not getting wet is simply a plan for this goal when it is raining acid rain. Achieving the plan to a degree does not fulfill the goal at all, and is thus not a very good plan. In the previous

case, the goal was not such a binary one, so partial fulfillment is possible. In other words, the goal must refer to a degree-like scale. As the story goes, a father will not be comforted to learn that his unmarried daughter is just a little pregnant.

Partial goal fulfillment provides the SIMULATE-AND-SELECT meta-plan with a complicated option. In addition to strict abandonment scenarios, it now also has the possibility of proposing options in which the partial fulfillment of one goal enables the (possibly partial) fulfillment of the other. The process that proposes this partial fulfillment scenario is as follows: MAKE-ATTRIBUTION determines that the problem with the "stay dry" goal is that it requires not going outside. Thus a partial version of this goal is sought that doesn't involve this condition. Thus in the case of not getting wet above, the "stay as dry as possible" alternative is postulated because this doesn't require not going outside. This scenario is therefore hypothesized and evaluated along with the strict abandonment options, and the one with the highest value chosen.

4.1.3. Plausibility Evaluation

So far, we have been assuming that the planner has absolute certainty about each scenario that is envisioned. However, the success of a plan is more apt to be a probabilistic quantity. That is, a planner may believe that an envisioned scenario will fulfill his various goals with different degrees of certainty. Since the degree of expectation of success will affect the overall value of the scenario, the planner must take these factors into account in selecting an option.

For example, if the planner believes that some scenario offers a very small chance of fulfilling a very important goal, but a competing scenario offers a larger chance of fulfilling a less important goal, then these likelihoods must be considered to determine which scenario is to be preferred. Thus crossing the

street may be in conflict with not getting hit by a car, and the value of preserving one's life is apt to be greater than the ultimate value of the goal underlying crossing the street. However, the perceived likelihood of such an occurrence is usually so small as to undermine the superior value of that goal from influencing one's plan.

Once the probability of success enters into consideration, we realize that situations we have been calling goal conflicts are actually often only plausible conflicts. That is, all the goal conflict detection and reasoning apparatus that has been discussed so far may be called into play when the planner expects that pursuing one goal may impede the fulfillment of another, but in which the interference may itself only be probabilistic.

Fortunately, these considerations can be accommodated without serious alterations to the planning structures I have proposed if we assume that the planner can attach plausibility judgments to anticipated events and propagate these judgments to the derived scenarios. Then the descriptions of situations used by the goal detector to spot goal conflicts can contain the degree of plausibility of the goal as well. For example, the planner might have a description denoting that walking down a dark alley alone at night conflicts to some particular degree with preserving one's money; upon considering such an action, this conflict will be spotted as would an "absolute" conflict. Then the relative values and likelihoods of fulfillments of the goals could be considered to determine the appropriate course of action.

To propagate along plausibility judgments, the planner needs to be able to access the probability that a plan will succeed if all its preconditions are fulfilled; then the interaction with the other plans amounts to accessing the probability that those preconditions will be met. A goal conflict then is a situation in which pursuing one plan lowers the probability of the preconditions for some

other plan being met.

In situations in which the likelihood and values of the goals involved are in concord, selecting the most productive scenario is unproblematic. However, in those situations in which the values and expected achievement of goals are inversely related, a way of comparing these quantities needs is needed. However, the computation of the expected value of plan of action does not appear to be something that people carry out in a principled way. That is, these situations are those in which it is notoriously difficult for people to come to a decision. Possibly, rather than a principled means of computing expected value, a set of heuristics is used. Our theory may be agnostic on this point, however. We need only assume that the value of a scenario computed by SIMULATE-AND-SELECT takes into account the relative likelihood of success in addition to the value of one's goal. In those cases where it cannot, the SIMULATE-AND-SELECT plan may simply fail, leaving the planner in a quandary as to which action to take.

4.1.4. An Example

As many of the situations just described are complicated by the presence of a number of goals all involving the same plan, it is instructive to look at one such situation in some detail. One plan I might have for eating lunch is to go to the Three C's with Lotfi and order a crepe. The goal structure behind this a plan is actually quite rich. First, there are several goals that this plan is seeking to satisfy. They include

- 1) Satisfying my hunger, as it is lunch-time and eating most anything will fulfill this goal,
- 2) Enjoying the experience of eating, as I like good food and the crepes at this restaurant are tasty,

- 3) Getting back to the office quickly, as I need to return to my office soon, and this restaurant is nearby,
- 4) Enjoying company, as I prefer not to eat alone, and enjoy Lotfi's company,
- 5) Not spending more than four dollars for lunch.

If in fact these are precisely my goals, then the plan I chose is only one a number of alternatives that might work just as well. For example, I could conceivably go to another restaurant that is equidistant and which serves just as good food, or I might try another companion if the one specified in this plan were not available. These variations would fulfill the actual goals just as well, and thus constitute instances of re-specification rather than goal substitution.

On the other hand, if no restaurant near my office were opened, then I might have to settle for the partial fulfillment of a goal, namely, I would attempt to get back as quickly as possible. Or if the only good restaurant nearly served very small portions, I have have a goal conflict between "satisfy-hunger" and "enjoy good food". This conflict might be resolved once again by settling for partial fulfillment, achieving "satisfy-hunger as much as possible" rather than "satisfy-hunger".

Alternatively, my original goal structure could have been somewhat different. For example, I might have had a urge for a Three C's crepe, in which case going to another good restaurant would have to be viewed as goal abandonment. Or I might have had some particular reason to talk with Lotfi, so choosing another co-diner would not be acceptable. It is unlikely, however, that I would have had only very general original goals like "enjoy any activity" and "satisfy some urge", so that when I could find no restaurant opened nearby, I would be just as happy to say, go want a movie with a friend. Of course, I might actually end up doing so in this situation. But it must be interpreted as abandoning a number of my original goals rather than as finding a different way of achieving

them.

5. Summary and Projections

We have proposed a theory of planning based on the following tenets: (1) Common-sense planning is essentially the consideration of interactions of otherwise simple plans, (2) a planner must be able to deduce its goals as well as produce plans for them, (3) planning problems should be formulated as meta-goals, and solved by the same planning mechanism responsible for the fulfillment of ordinary goals, (4) to accomplish much of its mandate, the planner makes projections of the future based on its current knowledge of the world and its own tentative plans, and (5) knowledge about planning should be usable both by a planner and an understander.

These tenets form the basis for a model of planning whose most salient features are a goal detector and a projector. The goal detector is used to infer goals, including meta-goals, based on the situations in which the planner finds itself; the projector is used to guess what the future will bring based on the planner's current beliefs and plans. As the goal detector has access to the hypothetical situations simulated by the projector, it can detect problems with currently intended plans by noticing their consequences in hypothetical realities. These problems are dealt with by setting up meta-goals to try to assure a more desirable future state of affairs.

We examined this model of planning in the particular domain of goal conflict resolution. Here we found use for the meta-plans RE-PLAN (consisting of USE-NORMAL-PLAN and USE-ALTERNATE-PLAN) and CHANGE-CIRCUMSTANCE for the meta-goal RESOLVE-GOAL-CONFLICT. Both meta-plans make use of the powerful sub-plan MAKE-ATTRIBUTION. For the related goal of CHOOSE-MOST-VALUABLE-SCENARIO, the SIMULATE-AND-SELECT meta-plan is used to create alternatives

involving goal abandonment and partial goal fulfillment. MAKE-ATTRIBUTION was found to be useful here as well.

We are currently attempting to test these ideas in two programs. PAM, a story understanding system, uses knowledge about goal interactions to understand stories involving multiple goals. That is, PAM can detect situations like goal conflict and goal competition, and, realizing that these threaten certain meta-goals, PAM will interpret a character's subsequent behavior as a meta-plan to address the negative consequences of these interactions. Thus PAM makes use of the knowledge structures described above, but of course, it does not test the model of planning per se.

Both the model of planning knowledge and of planning is being used in the development of PANDORA (Plan ANalyzer with Dynamic Organization, Revision and Application). PANDORA is given a description of a situation and determines if it has any goals it should act upon. It then creates plans for these goals, using projection to test them. New goals, including meta-goals, may be inferred in the process, possibly causing PANDORA to revise its previous plans.

The following is an example of the kind of planning situation that PANDORA can handle. PANDORA is presented with a task that requires it to get some nails and to get a hammer. PANDORA proposes the normal plans for these goals, which require it to go to the store, buy the desired item, and return. As the plans involve some common preconditions, the meta-theme "DON'T WASTE RESOURCES" causes PANDORA to have the meta-goal COMBINE-PLANS. A meta-plan associated with this goal synthesizes a new plan that involves going to the store, buying both objects, and returning.

PANDORA can also detect and resolve a number of goal conflict-base situations. In addition, PANDORA is being used to model the planning processes of a human who needs to cook dinner during a power failure, in which most of the

normal plans for one's goals will not be effective.

References

- 1) Fikes, R. and Nilsson, N. J. (1971). STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence* 2, 189-208.
- 2) Hayes-Roth, B. and Hayes-Roth, F. (1978). Cognitive Processes in Planning. RAND Report R-2366-ONR.
- 3) McDermott, Drew (1978). Planning and Acting. In *Cognitive Science* vol. 2, no. 2.
- 4) McCarthy, J. and Hayes, P. J. (1969) Some philosophical problems from the standpoint of artificial intelligence. In Meltzer and D. Michie (eds.) *Machine intelligence*, vol. 4. New York: American Elsevier, pp. 463-502.
- 5) Newell, A., and Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, N. J.: Prentice Hall.
- 6) Rieger, C. (1975a). One System for Two Tasks: A Common-sense Algorithm Memory that Solves Problems and Comprehends Language. MIT AI Lab working paper 114.
- 7) Rieger, C. (1975b). The Common-sense Algorithm as a Basis for Computer Models of Human Memory, Inference, Belief, and Contextual Language Comprehension. In *Theoretical Issues in Natural Language Processing*, R. Schank and B. L. Nash-Webber, (eds.). Cambridge, Mass.
- 8) Sacerdoti, E. D. (1977). *A Structure for Plans and Behavior*. Elsevier North-Holland, Amsterdam.
- 9) Schank, R. C. and Abelson, R. P. (1977). *Scripts, Plans, Goals, and Understanding*. Lawrence Erlbaum Press, Hillsdale, N.J.
- 10) Shortliffe, E. H. (1976). *MYCIN: Computer-based Medical Consultations*. American Elsevier.
- 11) Stefik, Mark J. (1980). Planning and Meta-Planning - MOLGEN: Part 2. Stanford Heuristic Programming Project HPP-80-13 (working paper), Computer Science Department, Stanford University.
- 12) Sussman, G. J. (1975). *A Computer Model of Skill Acquisition*. American Elsevier, New York.
- 13) Wilensky, R. (1978). Understanding goal-based stories. Yale University Research Report No. 140.

- 14) Wilensky, R. (1980). Meta-planning: Representing and Using Knowledge about Planning in Problem Solving and Natural Language Understanding. Berkeley Electronic Research Laboratory Memorandum No. UCB/ERL/M80/33.