Rational Handling of Multiple Goals for Mobile Robots

Richard Goodwin and **Reid Simmons**
School of Computer Science
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213-3890

Abstract

The mobile robot planning domain is dynamic, with goals becoming active asynchronously. In order to successfully operate in this environment, a robot must be able to interrupt and reformulate its plans of action on-the-fly. This paper investigates a method for incorporating the accomplishment of a new goal into a partially executed plan. A decision theoretic approach using net present value as the decision criterion serves as the basis for doing dynamic goal ordering. The appropriateness of net present value over benefit-cost ratio is argued. The approach has been implemented on a robot operating in an office setting. Examples from this domain are used to show the advantages of the approach with respect to fixed priority and heuristic based approaches.

1 Introduction

This paper examines a method for handling multiple active goals in mobile robots. Specifically, the focus is on asynchronous goal activation and on how to incorporate the accomplishment of a newly active goal into a partially executed plan. A utility based decision theoretic approach is adopted for investigating the tradeoffs that must be made.

The applicability of decision theory to problems in artificial intelligence and planning in particular has long been recognized [Horvitz *et al.*, 1988]. Feldman and Sproull, in their analysis of planning for the hungry monkey problem, use utility based decision theory to evaluate plans taking into account uncertainty and risk [Feldman and Sproull, 1977]. More recent work has focused on deciding when to refine plans [Boddy and Dean, 1989]. The work presented here differs from these in that it focuses on plan evaluation when all goals are not initially known and the plan must be reformulated as goals become active. In such cases, the time dependent utility of goal satisfaction, as well as the time distribution of utilities and resource use, must be taken into account.

Given a utility based framework, one must choose an appropriate decision criterion. In this paper two such criteria are analyzed: net present value and benefit-cost ratio. While both measures take into account the time distribution of utilities, net present value is shown to have some advantages over benefit-cost ratio when dealing with non-independent goals with discrete resource requirements. Plan evaluation based on net present value has been incorporated into a planning system that can interrupt an executing plan and dynamically order goals. The planning system has been applied to two mobile robot domains. The Ambler [Simmons and Krotkov, 1991], a prototype planetary exploration robot designed to carry out scientific missions, has been used as a model for a number of simulations. A Hero 2000 robot, used to perform a number of tasks in our lab, has been used as a vehicle for implementing the ideas. This paper examines a number of examples from the Hero domain to show the advantage of using a decision theoretic approach over heuristic based methods and fixed priority schemes.

2 Utility Based Rationality

Modern decision theory is concerned with making rational choices between alternatives [Raiffa, 1968] [Schoemaker, 1980] [Sinn, 1983] [Dawes, 1988]. Rational is taken to mean choosing the course of action that maximizes the expected value of some desired quantity, such as utility. Decision theory provides mechanisms for dealing with uncertainty and the cost of acquiring information. For this reason, it is being increasingly used for planning in real-world domains [Wellman, 1988] [Boddy and Dean, 1989] [Hansson *et al.*, 1990] [Chrisman and Simmons, 1991].

There are two requirements for formulating a planning problem in terms of decision theory. A method is needed to assign benefits or utilities to the accomplishment of each goal and costs or negative utilities to the consumption of each resource. Secondly, a decision criterion is needed to assess the relative merit of alternative plans.

The assignment of utilities is highly dependent on the set of tasks being considered and the desired behaviour. The exact magnitude of the utility values assigned is not as significant as the relative magnitudes which should reflect the relative priority of the goals.

A number of possible decision criteria have been suggested in the literature [Sassone and Schaffer, 1979]. Much of the body of work done on the development and analysis of the different criteria has focused on its applicability to economic domains [Simon, 1982]. The insights resulting from this work can be adapted to the mobile robot domain. Two commonly used criteria are examined below with comments on their appropriateness.

2.1 Benefit-Cost Ratio

The benefit-cost ratio is the sum of the present value of the benefits divided by the sum of the present value of the costs. The present value of a cost or benefit is the actual value discounted, by a fixed discount rate (d), for how long in the future it will occur.

$$\frac{B}{C} = \frac{\sum_{t=0}^{n} \frac{B_t}{(1+d)^t}}{\sum_{t=0}^{n} \frac{C_t}{(1+d)^t}} \; or \; \frac{\int_0^n \frac{B_t}{(1+d)^t} dt}{\int_0^n \frac{C_t}{(1+d)^t} dt} \tag{1}$$

d discount rate
B_t Benefit at time t
C_t Cost at time t

Using the present value of the costs and benefits takes into account their time distribution. Discounting future utilities creates a preference for benefits that accrue sooner and costs that occur further in the future. For example, given the choice between two plans that achieve the same benefit for the same initial cost, the one that returns the benefit sooner is preferred.

Taking the ratio of the present values of the benefits and the costs gives a measure of the rate of return. Alternatives that incur less cost to produce the same net benefit will be preferred. In economics, investments are selected by ranking the investment opportunities in order of decreasing benefit-cost ratio and accepting investment opportunities until the available resources are exhausted or the rate of return falls below the cost of capital. This greedy optimization algorithm allows the opportunities to be considered independently and leads to a very efficient decision procedure that is linear in the number of opportunities. [Etzioni, 1989], in the design of an autonomous agent, uses the algorithm as a basis for the agent's decision control loop.

As [Etzioni, 1989] points out though, there are problems when the opportunities require a discrete amount of each resource and resources are limited. In such cases, the problem can be shown to be intractable by a reduction from the knapsack problem [Garey and Johnson, 1979]. In practice, use of the greedy algorithm does lead to problems. Imagine a situation in which an exploration robot has located two adjacent items of interest. One item has a higher value than the other, but also consumes proportionately more resources to extract. Further suppose that there were not enough resources to take both samples — exactly one must be chosen. The greedy algorithm can fail in this case by selecting the option that gives the higher rate of return, but a lower net return.

A modified version of the greedy algorithm can be used in which the greedy solution is compared to a solution consisting solely of the item with the maximum net return, and the better of the two solutions selected. This modified algorithm can be shown to be within a factor of two of optimal [Garey and Johnson, 1979]. In the above sample selection example, the modified greedy algorithm correctly chooses the option with the highest net return. Suppose, however, that the situation was changed so that there were enough resources to sample both but that the item with the lower rate of return was degrading over time. The modified greedy algorithm would not be able to generate the optimal plan to first sample the low rate of return item and then to sample the high rate of return item. The algorithm fails because the opportunities are not independent. In general, considering opportunities in isolation is insufficient and combinations must be evaluated when selecting a plan.

Another difficultly with using the benefit-cost ratio is that it is dependent on the exact definition of costs and benefits. Suppose there are two methods a robot can use to traverse a room: one that is fast and uses little energy, but is noisy, and a second method that is quiet, but takes longer and uses more energy. Should the negative utility of disturbing others in the room with the noisy traversal be counted as a cost or as a negative benefit? Clearly, the way in which such external effects are treated will affect the ratio, and hence the robot's decisions.

2.2 Net Present Value

The net present value of a sequence of costs and benefits is the net of the present value of each negative utility/cost or positive utility/benefit. This is the most widely used metric in cost-benefit analysis and is generally considered superior to other metrics [Sassone and Schaffer, 1979]. In contrast to the benefit-cost ratio criterion, net present value chooses to maximize the net return rather than the rate of return.

$$\begin{aligned} Utility:\ U_t &= B_t - C_t \\ NPV &= \sum_{t=0}^{n} \frac{U_t}{(1+d)^t} \; or \; \int_0^n \frac{U_t}{(1+d)^t} dt \end{aligned} \tag{2}$$

As with benefit-cost ratio, the use of present values creates a preference for benefits that accrue sooner and costs that occur later. Resource investments are chosen by generating all feasible combinations of opportunities and selecting the one with the highest net present value. Even in the case where the investment options are not independent (as in the second version of the sampling example above) this method prefers the option that maximizes the net return.

Net present value treats negative benefits and costs equivalently. There is no need to make arbitrary distinctions.

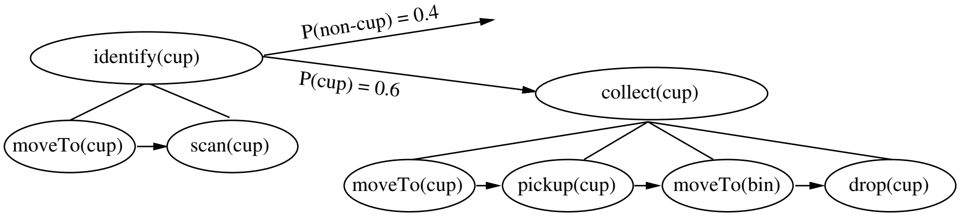

Figure 1: **Cup Collection Plan.**

Summing the costs and benefits does mean, however, that they must be normalized to the same scale. In economics this is done by expressing quantities in equivalent dollar values. For the robot domain, quantities can be normalized to their equivalent value in terms of a specific resource or benefit such as time, battery charge or samples taken.

The net present value method does not exhibit a preference for options that conserve resources. No distinction is made between two options that have the same net return but incur different costs. Conserving resources is desirable to the extent that it allows the unused portion to be used for other purposes. The benefit-cost ratio naturally prefers plans that incur less cost to produce the same net benefit since it is a measure of the rate of return. When using net present value, the opportunity cost of consuming a resource has to be taken into account. In a situation where net present value does not distinguish between two options, the one with the lower cost should be preferred.

Adopting the use of net present value leads directly to the intractable problem of having to generate and evaluate a combinatorial number of alternatives. Some method must be used to reduce the number of alternatives that have to be considered. The approach taken in this work has been to generate only a subset of the possible combinations and to do this incrementally as new opportunities become available. Details of the method used are given in the following section on the planning framework.

2.3 Discount Rate

While discounting future values accounts for the time preferences of costs and benefits, the choice of a discount rate is highly problematic. The discount rate reflects a willingness to trade present benefits for future costs. A low discount rate results in decisions that focus on long term impacts; a high discount rate results in greedy decisions. The discount rate incorporates assumptions about risk aversion and the predictability of the environment. For example, using a higher discount rate decreases risk by reducing dependence on the accuracy of predictions about the future since shorter plans with more current benefits are preferred. It would be desirable to adjust the discount rate as the model of the environment was refined and confidence in its accuracy increased.

3 Planning Framework

We have developed a planning framework that is geared towards handling asynchronous activation of goals involving robot motion and manipulation. A set of abstract actions is used to construct linear, conditional plans which are refined for execution by means of hierarchical decomposition of the abstract actions. Associated with each abstract action is the information needed to determine if and how the action can be interrupted. This information is used by a plan generator to create a set of plans that merge the handling of a newly active goal into the existing plan. The plan with the highest estimated net present value is selected for execution.

The conditional plan to achieve a set of active goals is represented as a tree of abstract actions. Figure 1 shows a simplified version of the plan for putting a potential cup in the bin. The plan consists of two abstract actions: one to determine if the object is a cup and the second to put it in the bin if it is. There is a branch in the plan for each possible outcome of an abstract action and associated with each branch is the probability of the corresponding outcome. These probabilities are used to weight the value of each branch when calculating the net present value of a plan.

Executing actions requires use of the robot's resources, such as grippers. As in the O-PLAN plan representation [Currie and Tate, 1985] each abstract operation specifies the resources that it requires. Resource information allows the planner to efficiently interrupt an action. For example, if the *collect(cup)* action in Figure 1 were interrupted to handle a recharge goal, the robot would not need to put the cup down since recharging does not require the gripper resource. Abstract actions may need to include "phantom" sub-actions [Secerdoti, 1977] to establish any required state since conditions may be clobbered between abstract actions. In the cup collection plan of Figure 1, both abstract actions include the *moveTo(cup)* sub-action. The second moveTo is a "phantom" action since it will not actually be executed if the plan is not interrupted.

Generating a plan for a set of goals is, in general, an intractable problem [Chapman, 1987]. The use of the linearity assumption that goals can be satisfied one at a time allows the planner to decompose the problem and generate a plan efficiently. Even with this linearity assumption, there is still a combinatorial number of possible goal orderings that could be considered when trying to optimize the plan. In order to avoid considering all possible orderings, the plan generator creates only a subset of the alternatives that is linear in the size of the original plan. New plans are created by inserting the plan for the new goal immediately, if the current action can be interrupted, or after that action otherwise and after each subsequent abstract action.

The decision not to consider goal reordering or interleaving is based on the assumption of a benign world and a near optimal original ordering for the actions. It is similar to the strategy used in intention-based planning where the planner makes a commitment to its existing plan and filters out options that are inconsistent with this commitment [Bratman *et al.*, 1988]. Unlike Bratman et al.'s IRMA architecture, our current planner does not have a mechanism to override its commitment to its current plan. Whether limiting the planner to examining only a subset of possible goal orderings is rational depends on whether the opportunity cost of not considering other possible orderings is offset by the savings in computation time [Doyle, 1988].

The plan generator can also include domain-specific methods for generating plans. For the Hero domain, a method was added for inserting a new goal when the currently executing action involves carrying an object from one location to another. An *on-the-way* plan is created in which the robot immediately starts achieving the new goal, but drops any objects it is carrying at the point on the new path that is closest to its intended destination.

4 Hero Robot Domain

The Hero 2000 robot operates in an office setting performing a number of tasks [Simmons *et al.*, 1990]. These include delivering printer output, taking objects from one workstation to another, and finding cups on the floor and putting them in a bin. The robot must also maintain its battery charge in order to be able to perform these tasks. The robot has a single manipulator and can carry only one object at a time.

Plan interruption, generation and selection using a net present value decision criterion has been incorporated into the software used to control the Hero 2000 robot. The Task Control Architecture [Simmons *et al.*, 1990], an operating system for robot's, is used as a basis for the implementation. It provides the mechanisms needed to schedule and control multiple goals and actions and to monitor the environment.

Direct experimentation with the Hero robot is time consuming. In order to investigate a larger variety of examples and a larger range of parameter values, a system for simulating the robot, using the planning framework described above, was created with the Maple symbolic math system [Char, 1987]. The simulation software is domain independent. It is targeted to a particular domain by specifying action models, expected time and outcome probabilities as well as the utility of accomplishing each goal.

The characteristics of the Hero domain were determined empirically. Euclidean distance and average speed are used to estimate travel times. A discount rate of 0.2% per second was chosen which results in discounting utilities six minutes in the future by 1/2. The six minute time frame is sufficient for the robot to complete one or two tasks, reflecting the robots effective planning horizon.

The utility of having the robot accomplish one of its goals depends on its value to the people in the office in terms of the amount of time it saves them. Time saved, or not saved, was used as a basis for normalizing costs and benefits. The utility values used are the normalized sums of the costs and benefits for satisfying each goal. The time dependent nature of the goals was also taken into account. Delivering printer output and carrying objects from one workstation to another must be done in a timely fashion since people are waiting. The utility of both these activities is represented by a function that is initially almost flat, but that decreases to near zero after a delay. The height of the function represents the intrinsic value of accomplishing the goal and the cut off represents the acceptable delay (Eqs 3 and 4). Cup collection is of general benefit, but since no one is waiting for it, it is time insensitive and of relatively low importance (Eq 5). Charging after a low battery indication is not directly beneficial to anyone, but is necessary for the robot to operate. If the robot delays recharging too long and runs out of charge, someone will be required to intervene. For this reason, recharging is characterized by a function containing a negative exponential component, making the utility prohibitively negative after an initial delay. This delay represents the time before the robot would start to lose power (Eq 6).

Hero utility functions (delay in seconds)

$$Utility(printer, delay) = \frac{100}{1+e^{\frac{delay-200}{10}}} \quad (3)$$

$$Utility(delivery, delay) = \frac{200}{1+e^{\frac{delay-200}{10}}} \quad (4)$$

$$Utility(collectCup, delay) = 10 \quad (5)$$

$$Utility(recharge, delay) = -e^{delay-200} \quad (6)$$

The following three examples illustrate the usefulness of the methods described.

4.1 Cup Collection Example

Suppose the robot is attending to a low utility goal when a new high utility request is received. The decision to be made is whether to continue with the current plan or suspend it until the new high utility goal is accomplished. In this example, the robot has an initial plan to collect a cup and receives a new request to deliver printer output. The room layout for the example is given in Figure 3. Objects are placed in the room in such a way that the cup and the bin

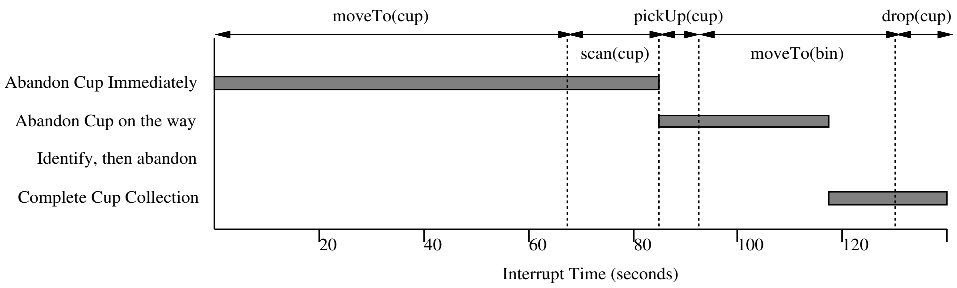

Figure 2: **Best Goal Ordering versus Interrupt Time, Cup Collection**

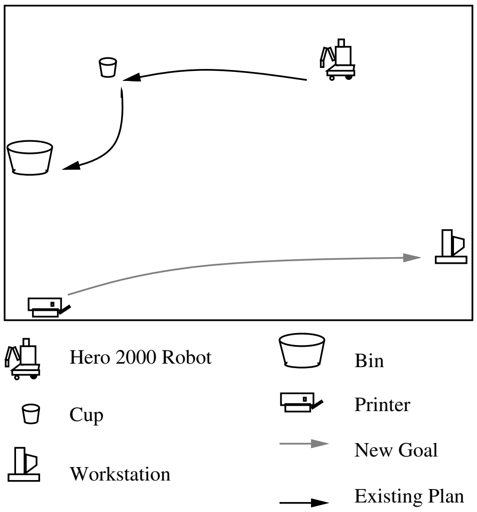

Figure 3: **Cup Collection Example.**

are only short detours on the way to the printer from the initial robot location.

Figure 2 shows how the preferred plan varies as a function of the new goal's activation time. This graph was produced by running the simulation with different interrupt times for the new printer delivery goal. The graph shows that before the robot has picked up the cup, it will suspend cup collection in favour of the printer request. Once the cup is picked up, it will be dropped off on the way to the printer, unless the robot is sufficiently close to the bin to make putting the cup in the bin worthwhile. The distance at which it becomes worthwhile to complete the cup collection is affected by the relative utilities of cup collection and printer output delivery, by the discount rate, by the cost of re-acquiring the cup, and by the relative positions of the robot, the bin and the printer. Abandoning the cup collection task incurs an extra cost to retrieve the cup since the robot must put it down in order to deliver the output. The relative positions are significant since moving towards the bin may move the robot towards or away from the printer.

The time line in Figure 5 describes the "Abandon Cup Collection" plan. The time delay for each goal is the time interval between when the request is received and when the goal is accomplished. The utility of accomplishing each goal is determined by substituting the corresponding goal delay time into the appropriate utility equation (Eqs 3 – 6). The discount interval for a goal is the amount of time over which the utility of each goal must be discounted in order to find its present value. This will be different from the delay time if some time has elapsed since the request time, as in the case with the cup collection goal. The utility values and the discount intervals are used in formula (2) to determine the net present value of each plan.

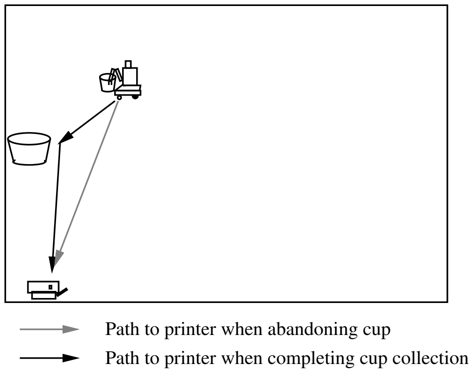

Figure 4: **Printer Path Difference.**

The tradeoff that is being made can be seen more clearly by considering the differences in the paths the robot would take to get to the printer. Figure 4 shows the two paths: one that goes directly to the printer and one that goes by way of the bin. The direct path will always be shorter, but as the robot approaches the bin, the difference becomes arbitrarily small. The corresponding added delay in deferring the printer goal approaches zero, as does the corresponding opportunity cost. The point at which this opportunity cost equals the opportunity cost associated with deferring the cup collection goal corresponds to the point at which the preferred plan switches from abandoning the cup to completing the cup collection. In Figure 2, the point at which the "complete cup collection" plan becomes the preferred plan occurs when the robot is about 4.5 feet from the bin.

The example illustrates the advantage of this method over fixed priority schemes. If a fixed priority scheme were

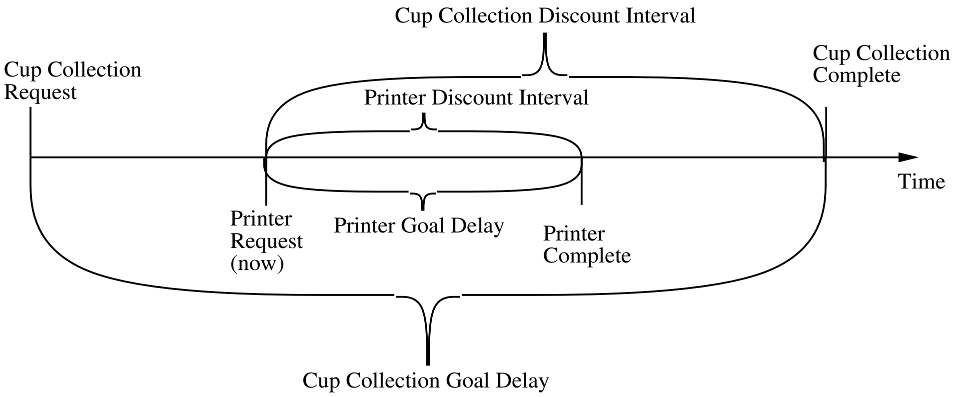

Figure 5: **Time line for Abandoning Cup Collection.**

used, as was done in the original Hero system, it could result in situations where the robot drops the cup beside the bin rather than expend the extra few seconds needed to drop it in the bin. This would get the printer output delivered a few seconds earlier, but it requires the robot to expend a significantly greater amount of time to return, re-acquire the cup and finish the task.

This example also serves to show some of the limitations of heuristic-based approaches. As stated above, the distance at which the cup collection should be completed depends on a number of factors. For a heuristic method to take these factors into account would require either a large number of simple heuristics [Feldman and Sproull, 1977] or a few heuristics that essentially embody the same calculations used in our method.

4.2 Delivery Example

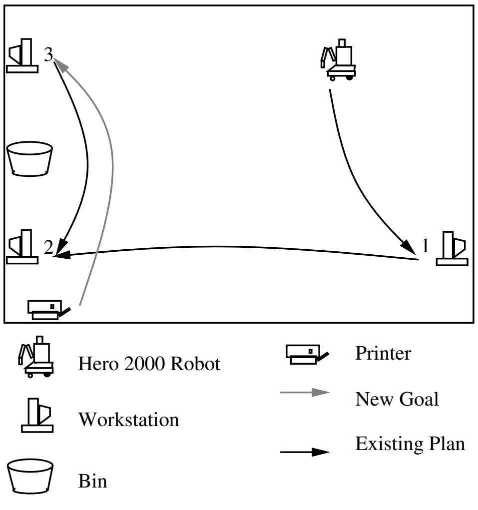

Figure 6: **Delivery Example.**

By properly ordering the achievement of goals, the robot can take advantage of the opportunities derived from the spatial relationships between goals. This result follows naturally from the utility-based approach. Consider the situation shown in Figure 6. The robot is making a sequence of deliveries, one from workstation 1 to workstation 2 and a second from workstation 3 to workstation 2. If a printer request arrives for workstation 3 then the robot can reduce its amount of travel by picking up the output on the way to workstation 3.

Figure 7 shows how the preferred strategy changes with the new goal activation time. Inserting the printer output request between the two deliveries reduces the total amount of travel needed by taking advantage of the fact that the printer request ends at the location where the second delivery begins. Note that the robot will initially go back to the printer even after starting towards workstation 3 to do the second delivery. It is advantageous to do so as long as the robot has not moved too far away from the printer.

A priority based scheme, if used in this example, would not be able to take advantage of the spatial relationships between the goals. The printer request would always be serviced last since it has the lowest utility. A heuristic-based approach could be used to suggest ordering goals such that the destination of one was the start of the next. However, this would not take into account situations as in the example where workstation 2 is only near the printer and not at the same location.

4.3 Contingency Example

In the course of working with the Hero Robot, an informal experiment was run to see how people handle the same tasks as the robot. It was observed that sometimes people would elect to "recharge" rather than go collect a cup, even though they had sufficient battery charge. Invariably, the reason given was that they wanted to have enough charge in reserve to be able to handle a possible printer or delivery request.

The techniques described in this paper can be applied to model this type of contingency planning. The situation in the experiment was modeled as a choice between two

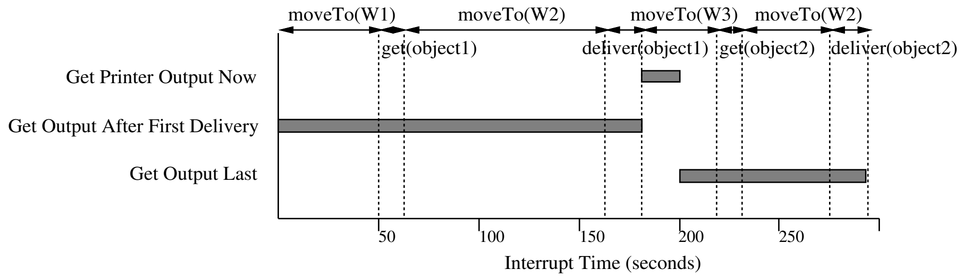

Figure 7: **Best Goal Ordering versus Interrupt Time, Deliveries**

$$NPV_1 = \int_0^T \frac{P(printer) * Plan_1'(t)}{(1+d)^t} dt + (1 - (P(printer) * T)) * NPV(Plan_1) \quad (7)$$

plans: plan_1, to collect the cup first, and plan_2, to recharge first. These plans were evaluated taking the possibility of a printer request into account. It was assumed that there was a constant probability *P(printer)* of a printer request arriving in any minute, (the possibility that two or more requests would arrive was ignored). Let $Plan_1'(t)$ be the net present value at time t of the plan that would be selected if $\text{plan}_1$1 were used and a printer request arrived at time *t*. Multiplying $Plan_1'(t)$ by the probability that a request will arrive at time *t* and discounting it back to time zero gives the current net present value weighted by its probability. Integrating gives the total net present value. (See equation 7) A similar calculation gives the result for the plan_2.

Using the utility values selected for the domain and a 20% probability of a printer request arriving from workstation 3 in any minute results in a preference for the plan that recharges. If the possibility of a printer request is not taken into account, the other plan is preferred. Obtaining this result using full numerical integration in Maple is computationally expensive requiring a few minutes of elapsed time on a sparc II workstation. The example does serve to suggest, however, that the approach may be applicable using a more efficient implementation and further approximations.

5 Conclusions

This paper has presented some initial results on rational planning for mobile robots. The examples presented show that a mobile robot can take advantage of opportunities as they arise if it can interrupt and reformulate its plan of action. A decision theoretic approach to plan reformulation is more general than heuristic based methods and produces more rational results than do fixed priority schemes. The use of a net present value decision criterion for the mobile robot domain has some advantages over a benefit-cost ratio criterion when dealing with limited resources and non-independent alternatives.

Decision criteria using present values of utilities are highly sensitive to the discount rate. Choosing a discount rate is still a matter of experimentation. Further work is needed to determine how the characteristics of the domain should be taken into account when selecting a discount rate. One possibility is to have the robot adjust its discount rate as it refines its time and probability estimates.

The planning framework chosen has a number of limitations. Using the linearity assumption and considering only a limited number of plans serve to reduce the amount of computation needed. These are fundamental limitations, however, and they could be removed with a corresponding increase in computation time. The question to be studied is when it is advantageous to do so.

A decision theoretic approach to plan evaluation is useful when dynamically reordering multiple active goals. Coupled with the use of net present value and consideration of opportunity costs, it provides the basis for effective operation of a mobile robot.

Acknowledgments

We wish to express our thanks to Long-Ji Lin and Lonnie Chrisman for their work on the initial design and implementation of the Hero robot system. This research is supported by NASA under contract NAGW-1175 and in part by a Natural Sciences and Engineering Research Council of Canada 1967 Science and Engineering Scholarship.

References

[Boddy and Dean, 1989] Mark Boddy and Thomas Dean. Solving time-dependent planning problems. In *IJCAI-89 Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*. AAAI, August 1989.

[Bratman *et al.*, 1988] M. E. Bratman, D. J. Isreal, and M. E. Pollack. Plans and resource-bounded practical reasoning. *Computational Intelligence*, 4(4), 1988.

[Chapman, 1987] D. Chapman. Planning for conjunctive goals. *Artificial Intelligence*, 32, 1987.

[Char, 1987] Bruce W. Char. A tutorial introduction to maple. Technical Report CS-85-56, University of Waterloo, 1987.

[Chrisman and Simmons, 1991] Lonnie Chrisman and ,Reid Simmons. Senseful planning: Focusing perceptual attention. In *Proceedings, Ninth National Conference on Artificial Intelligence*. AAAI, July 1991.

[Currie and Tate, 1985] K. Currie and A. Tate. O-plan-control in the open planning architecture. *Expert Systems*, 1985.

[Dawes, 1988] Robyn M. Dawes. *Rational Choice in an Uncertain World*. Harcourt Brace Jovanovich, Toronto, 1988.

[Doyle, 1988] Jon Doyle. Artificial intelligence and rational self-government. Technical Report CMU-CS-88-124, Carnegie Mellon University, March 1988.

[Etzioni, 1989] Oren Etzioni. Tractable decision-analytic control. In *Proceedings, First International Conference on Principles of Knowledge Representation and Reasoning*, May 1989.

[Feldman and Sproull, 1977] Jerome A. Feldman and Robert F. Sproull. Decision theory and artificial intelligence ii: The hungry monkey. *Cognitive Science*, 1, 1977.

[Garey and Johnson, 1979] Michael R. Garey and David S. Johnson. *Computers and Intractability*. W. H. Freeman and Company, 1979.

[Hansson *et al.*, 1990] Othar Hansson, Andrew Mayer, and Stuart Russell. Decision theoretic planning in bps. In *Proceedings of the AAAI Symposium on Planning*. AAAI, 1990.

[Horvitz *et al.*, 1988] Eric J. Horvitz, John S. Breese, and Max Henrion. Decision theory in expert systems and artificial intelligence. *International Journal of Approximate Reasoning*, 2, 1988.

[Raiffa, 1968] Howard Raiffa. *Decision Analysis Introductory Lectures on Choices under Uncertainty*. Addison-Wesley, Reading Mass., 1968.

[Sassone and Schaffer, 1979] Peter G. Sassone and William A. Schaffer. *Cost-Benefit Analysis A Handbook*. Academic Press Inc, 1979.

[Schoemaker, 1980] Paul J. H. Schoemaker. *Experiments on Decisions Under Risk: The Expected Utility Hypothesis*. Kluwer Nijhoff, 1980.

[Secerdoti, 1977] Earl Secerdoti. *A Structure for Plans and Behavior*. American Elsevier, New York, NY, 1977.

[Simmons and Krotkov, 1991] Reid Simmons and Eric Krotkov. An integrated walking system for the Ambler planetary rover. In *Proc. IEEE International Conference on Robotics and Automation*, pages 2086–2091, Sacramento, CA, April 1991.

[Simmons *et al.*, 1990] Reid Simmons, Long Ji Lin, and Chris Fedor. Autonomous task control for mobile robots. In *Proc. IEEE Symposium on Intelligent Control*, Philadelphia, PA, September 1990.

[Simon, 1982] Herbert A. Simon. *Theories of Decision-Making in Economics and Behavioral Science*, pages 287–317. The MIT Press, 1982.

[Sinn, 1983] Hans-Werner Sinn. *Economic Decisions Under Uncertainty*. North-Holland, 1983.

[Wellman, 1988] Michael P. Wellman. Formulation of tradeoffs in planning under uncertainty. Technical Report MIT/LCS/TR-427, MIT, August 1988.