High-Stake Planning

Sven Koenig
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332-0280, USA
skoenig@cc.gatech.edu

Abstract

Preference models determine which one of several plans to prefer. It is important that planners use the same preference models as rational human decision makers because planners should make the same decisions as their human users (as long as they are rational), otherwise the planners are not of much use. While artificial intelligence planning has made lots of progress in the areas of representations of planning tasks and planning methods, it has not yet paid sufficient attention to the preference models of human decision makers. Utility theory is a formal framework for modeling the preferences of human decision makers and making rational decisions in high-stake decision situations. This paper reports on first steps in the direction of building planners that fit the risk attitudes of human decision makers in high-stake planning domains better than current planners, by combining descriptive methods from utility theory with constructive methods from artificial intelligence planning and operations research, thus combining the strengths of the two decision-making disciplines and extending the applicability of planners from artificial intelligence.

Introduction

Artificial intelligence planning has developed knowledge-based planners. These planners can have advantages over methods from other decision-making disciplines because they exploit more of the structure of large-scale planning tasks. Planning methods from artificial intelligence, for example, represent search spaces implicitly (for example, with STRIPS rules) and exploit the resulting decomposability. Artificial intelligence planning has made lots of progress over the past couple of year in the areas of representations of planning tasks as well as planning methods. However, artificial intelligence planning has not yet paid sufficient attention to the planning objectives, which are still primitive. In deterministic domains, planners from artificial intelligence have traditionally been used with the objective to find any plan that achieves the goal. To make their preference model richer, planners then began to associate execution costs with plans and preferred plans that achieve the goal with minimal plan-execution cost, that is, minimal consumption of one limited resource such as time, energy, or money. In probabilistic domains, planners from artificial intelligence usually either minimize the average plan-execution cost or, if the goal cannot be achieved for sure, maximize the probability of goal achievement. However, these preference models are often too simplistic to model the preferences of human decision makers adequately. How to plan with more realistic preference models, however, is a topic that has been neglected in the literature on artificial intelligence planning. It is an important topic because the recommendations of planners should reflect the opinions of their users correctly (as long as they are rational – we are not interested in irrational decision making). After all, the planners make suggestions for how to act and should make the same suggestions that the users would have made themselves. Otherwise the planners would not be very helpful.

Our research program therefore investigates how to build efficient planners that fit the preference models of rational human decision makers better than current planners, by combining descriptive methods from utility theory with constructive methods from artificial intelligence planning and operations research, thus combining the strengths of the two decision-making disciplines and extending the applicability of planners from artificial intelligence. We are interested, for example, in planning with deadlines and other resource limits as well as planning with multiple attributes, such as energy, cost, time, probability of goal achievement, prestige, and so on. In this paper, we report on a very first step in the direction of building planners with more realistic preference models. We study how to plan in high-stake decision situations with one resource, taking the risk-attitude of decision makers into account. High-stake decision situations occur in domains in which huge wins or losses of money, equipment, or even human life are possible. Many NASA domains are high-stake domains, including planning for autonomous space craft (Pell *et al.* 1997), and many human decision makers prefer to avoid the huge losses that are possible in these domains. We are also interested in how the risk attitude changes the optimal plan, for example, its influence on how long to plan before starting to act and how frequently to sense. Utility theory is a formal framework for modeling the preferences of human decision makers and making rational decisions in high-stake decision situations. It suggests that decision makers maximize their average utility in these domains, where the utility is a strictly monotonically decreasing but often nonlinear function of the plan-execution cost. However, it specifies only what optimal plans are but not how they can be obtained efficiently. The difficulty we need to overcome is how to combine ideas from utility theory, artificial intelligence, and operations research, which is nontrivial but allows us to exploit the structure of planning tasks to make planning with nonlinear utility functions efficient. In the past decades, artificial intelligence researchers have acquired a large body of knowledge about how to plan efficiently with the current preference models, by utilizing the structure of planning tasks. We will demonstrate that, in some cases, these insights can be used to plan efficiently with nonlinear utility functions. This promises to make planning with some nonlinear utility functions as fast as planning with standard planners from artificial intelligence or operations research and enables one to participate in performance improvements achieved by other researchers in the currently very active field of probabilistic planning, while extending the applicability of existing planners from artificial intelligence or operations research. It also promises to allow for an easy integration of risk attitudes into existing decision-support systems, automated planning systems, and agent architectures.

High-Stake Decision Situations

High-stake decision situations occur in domains in which huge wins or losses are possible. In high-stake decision situations, rational human decision makers usually do not minimize the average plan-execution cost or maximize the average reward because they take risk aspects into account. This is why many human decision makers buy insurance even though the insurance premium is usually much larger than the average loss from the insurance cause. Another example is the following simple decision situation with two alternatives, one of which has the larger average pay-off and the other one of which has the smaller variance. Consider the decision situation shown in Figure 1, where you can participate in one and only one of the following two lotteries at no charge. When human decision makers have to decide whether they would like to get 4,500,000 dollars for sure or get 10,000,000 dollars with fifty percent probability (and nothing otherwise), many human decision makers prefer the safe alternative although its average pay-off is clearly lower – they are risk-averse. (Similarly, there are decision makers that are risk-seeking.) It is important to realize that this is perfectly rational behavior. Risk-averse human decision makers are willing to accept a smaller mean of the pay-off for a decrease in variance because they fear for the worst case. They are trying to avoid catastrophes, and a small variance avoids pay-offs that are much smaller than average. If a planner chose the lottery with the larger average pay-off, then many human decision makers would be extremely unhappy half of the time. It is therefore important that planners reflect the opinions of rational human decision makers correctly. After all, planners make suggestions for how to act and should make the same suggestions that the human decision makers would have made themselves. (This is the reason why investment advisors ask for your risk attitude before making investment recommendations.) However, artificial intelligence planning has not studied how to determine plans that correctly reflect the risk attitudes of rational human decision makers in high-stake decision situations.

Preference Models

Bernoulli and Von Neumann/Morgerstern's utility theory (von Neumann & Morgenstern 1947; Bernoulli 1738) has investigated how rational human decision makers make decisions in high-stake decision situations. Utility theory can explain why they often do not minimize the average plan-execution cost. It suggests that it is rational to choose plans with maximal average utility, where the utility $u(c)$ is a strictly monotonically decreasing function of the plan-execution cost c. Human decision makers sometimes deviate from utility theory because, different from other theories, such as Kahneman and Tversky's prospect theory (Kahneman & Tversky 1979), utility theory does not model human inadequacies in decision making and thus is not able to explain all empirical findings about human decision making. This is not a problem for planners since planners are supposed to follow a theory of rational (normative) rather than empirical decision making. Maximizing average utility and minimizing average plan-execution cost result in the same decisions if either the domains are deterministic or the utility functions are linear. These assumptions, however, are often not satisfied. For example, nonlinear utility functions are necessary to account for the risk-averse attitudes of many rational human decision makers for the lottery example above. The lottery example can be explained as follows: Assume that a human decision maker is risk-averse and has the concave exponential utility function shown in Figure 2 and no other assets. This human decision maker associates utility (here: pleasure) 0.00 with a wealth of 0 dollars, utility 0.74 with a wealth of 4,500,000 dollars, and utility 0.95 with a wealth of 10,000,000 dollars. Then, the (average) utility of getting 4,500,000 dollars for sure is 0.74, whereas the average utility of getting 10,000,000 dollars with fifty percent probability is only 0.475. In this case, the safe alternative maximizes the average utility for this human decision maker, which explains why this human decision maker chooses the safe alternative over the one with the larger average pay-off. Other human decision makers can have other utility functions and thus arrive at different conclusions.

Choices	Probability	Pay-Off	Average Pay-Off	Utility	Average Utility
Choice 1	50 percent	10,000,000 dollars	5,000,000 dollars	0.95	0.475
	50 percent	0 dollars		0.00	
Choice 2	100 percent	4,500,000 dollars	4,500,000 dollars	0.74	0.740

Figure 1: Decision Situation

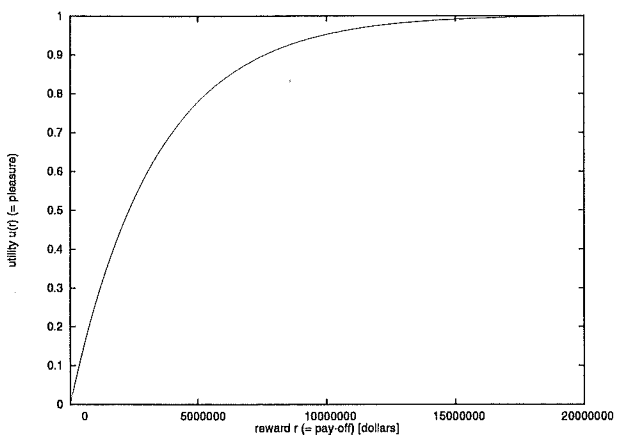

Figure 2: Risk-Averse Utility Function

Exponential Utility Functions

Utility theory specifies only what optimal plans are but not how they can be obtained efficiently, that is, other than by enumerating every trajectory of every possible plan. Operations research and control theory use dynamic programming methods to find plans with maximal average utility (Marcus *et al.* 1997; Whittle 1990) but these methods often do not exploit the structure of planning tasks completely. Artificial intelligence planning has developed knowledge-based planners that plan efficiently in large domains but usually either minimize the average plan-execution cost or, if the goal cannot be achieved for sure, maximize the probability of goal achievement. It might seem that one could simply replace the resource consumptions with their utilities and then continue to use planners from artificial intelligence. This is certainly true for one-stage decision situations, such as the lottery example. However, different from the lottery example, real-world planning tasks are usually much more complex and consequently involve much more complex decisions. For these multi-stage decision situations, it is impossible to simply replace the resource consumptions with their utilities and then continue to use planners from artificial intelligence. Assume, for example, that one incurs cost c_1 at the first time step and cost c_2 at the second time step, which is also the last one. Then, one obtains $u(c_1) + u(c_2)$ if one replaces the costs with their utilities whereas the correct utility is $u(c_1 + c_2)$. Instead, our multiplicative planning-task transformation (Koenig 1998) makes use of existing planners from artificial intelligence or operations research by transforming planning tasks with exponential utility functions to planning tasks that planners from artificial intelligence or operations research can solve. The transformed planning tasks can be solved by finding plans with maximal probability of goal achievement or minimal average plan-execution cost. The multiplicative planning-task transformation is such that optimal plans for the transformed planning task are also optimal for the original one, and good ("satisficing") plans for the transformed planning task are also satisficing for the original one. The multiplicative planning-task transformation works for convex and concave exponential utility functions. Convex exponential utility functions are of the form $u(c) = \gamma^{-c}$ for parameter $\gamma > 1$, and concave exponential utility functions are of the form $u(c) = -\gamma^{-c}$ for parameter γ with $0 < \gamma < 1$. These utility functions are expressive because the parameter γ can be used to trade-off between minimizing the best-case, average, and worst-case plan-execution cost (Watson & Buede 1987). As γ approaches infinity, the human decision makers become more risk-seeking and thus more interested in plans with small best-case plan-execution cost (under appropriate assumptions) (Koenig & Simmons 1994; Koenig 1998). As γ approaches one, the human decision makers become more interested in plans with small average plan-execution cost. Finally, as γ approaches zero, the human decision makers become more risk-averse and thus more interested in plans with small worst-case plan-execution cost. Thus, exponential utility functions can express a continuum of risk attitudes, that includes the utility functions of the lottery example above.

Crisis Management

Our current application area is managing environmental crisis situations such as oil spills (Desimone & Agosta 1994). The goal of planning in the oil-spill domain is to determine how to manage resources such as human teams, vessels, and equipment to contain and clean up oil spills, taking into account all costs incurred until they are cleaned up completely (expressed in dollars). Crisis management domains have several advantages: Efficient planning methods for them directly benefit the public. They are important high-stake planning domain for which many human decision makers are very risk averse since they prefer to avoid the huge losses that are possible in these domains. Thus, planners that minimize the average cost do not use the same preference model as many human decision makers and arrive at different courses of action.

Sensor Planning

An interesting observation in crisis management situations is that human decision makers gather large amounts of information even if it is costly and thus might not be part of a plan that minimizes the average plan-execution cost. In the oil-spill domain, sensing operations include, for example, sending out helicopters to gather information about how oil spills drift. The knowledge that an oil spill drifts towards a nature preserve can be used to concentrate resources in the surrounding sea sectors to prevent the oil from reaching sensitive shores. Thus, sensing provides information that can reduce the plan-execution cost but comes at a cost itself. Sensor planning for risk-averse or risk-seeking human decision makers involves the same trade-off as the lottery example above (Koenig & Liu 1999). For example, adding more sensing operations than is necessary to minimize the average plan-execution cost increases the mean of the plan-execution cost (because sensing is expensive) but also reduces its variance (because the information obtained can be used to avoid catastrophes). Consequently, we speculate that risk-averse human decision makers add more sensing operations than is necessary to minimize the average plan-execution cost.

Case Study

In the following, we provide a case study of how the risk attitude influences the sensing frequency to test our hypothesis that risk-averse human decision makers add more sensing operations than is necessary to minimize the average plan-execution cost. At the same time we demonstrate that our multiplicative planning-task transformation can be combined with existing planners that minimize the average plan-execution cost to yield planners that maximize average utility for exponential utility functions. We apply the multiplicative planning-task transformation to the sensor planner by Hansen (Hansen 1997) that combines methods from operations research (namely, policy iteration (Howard 1964)) and artificial intelligence (namely, the A* search method (Nilsson 1971)) to find plans with minimal average plan-execution cost. In (Koenig & Liu 1999), we present the resulting sensor planner and a proof of its correctness. Here, we present the results of a case study where we apply the resulting sensor planner to simple artificial robot-navigation tasks with actuator uncertainty. Notice that these robot-navigation tasks are not high-stake planning domains and *we do not suggest to apply our planning methods to planetary rover navigation in the way we do it here.* We use robot-navigation tasks merely as a test bed because they are much simpler to solve than the oil-spill domain (and we can thus run a much larger number of experiments in a reasonable amount of time), they allow us to visualize the planning results much more easily, and they have been studied before in the context of sensor planning (Hansen 1997) and can therefore be considered good test problems for new sensor planners. We consider robot-navigation tasks with the following properties. The robot has to navigate from a given start location to a given goal location in a known environment. Since motion is noisy, the robot can deviate from the nominal path but it can always opt to sense its current location. Sensing provides certainty about its current location but is costly (for example, consumes energy). We assume that there is a finite set of locations L. The robot knows that it starts at location $l_{start} \in L$ and its task is to navigate to location $l_{goal} \in L$ and be sure that it stops at exactly that location. There is a finite set M of movement actions, all of which can be executed at all locations. Motion uncertainty is modeled with conditional probability distributions. Executing movement action $m \in M$ results with cost $c(l, m) > 0$ and probability $p(l'|l, m)$ in location l'. The robot receives no feedback as to what its new location is (which makes the simplifying assumption that the cost of the executed actions cannot be observed directly) but there is one sensing action o that can be executed at all locations. Executing it incurs cost $c(l, o) > 0$ and reports the current location of the robot with certainty. We assume that it is possible to reach every location from every other location. Figure 3 (left) shows the gridworld that we use in our experiments, where the locations are squares. The start location is C1 and the goal location is J1. The robot can always sense its current location (O) or move north (N), east (E), south (S), or west (W) to an adjacent square. If the robot attempts to move in a certain direction (say, move east in square C1), then it either moves as intended (C2, with probability 0.6) or strays off by one square to the left (B2, with probability 0.2) or right (D2, with probability 0.2) due to actuator noise and not facing precisely in the right direction. The robot does not move when it bumps into the border of the gridworld. The movement cost ranges from 1.0 to 10.0. It is low for roads (white) and high for muddy terrain (darker colors). The sensing cost is always 0.2. Notice that even very risk-averse robots have to trade off between minimizing the worst-case and average plan-execution cost in our example domain even if they want to approximate plans with minimal worst-case plan-execution cost, a planning objective popular in robotics (Lozano-Perez, Mason, & Taylor 1984). It is not possible to minimize the worst-case plan-execution cost directly in our example domain because all plans cycle with some probability and thus have a worst-case plan-execution cost that is infinite. So, even very risk-averse robots need less risk-averse planning objectives than minimizing the worst-case plan-execution cost but more risk-averse planning objectives than minimizing the average plan-execution cost. Maximizing the average utility for an exponential utility function with γ sufficiently close to zero provides such a planning objective.

Figure 4 shows that the sensing frequency (that is,

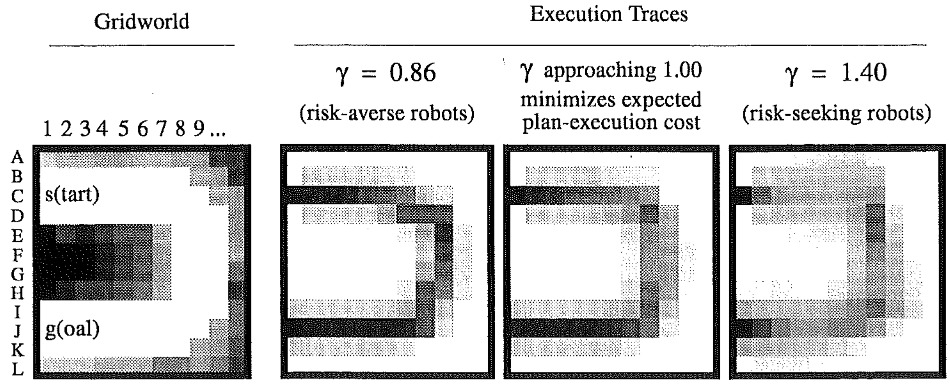

Figure 3: Gridworld and Execution Traces

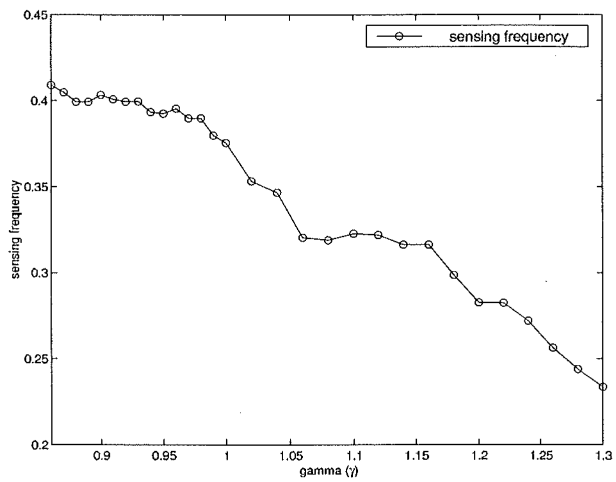

Figure 4: Sensing Frequency

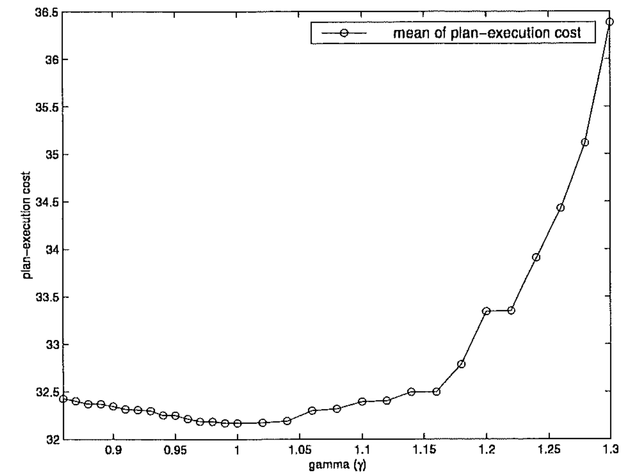

Figure 5: Mean of Plan-Execution Cost

the percentage of sensing actions among all executed actions) increases as the robots become more risk-averse and γ decreases. This is also illustrated in Figure 7, that shows optimal sensor plans for two different values of γ.[1] The sensor plans are depicted as gridworlds, each location of which is annotated with an action sequence. These action sequences are used as follows: After the robots have executed a sensing action, they look up the action sequence that corresponds to the sensed location, execute it, and repeat the process, until they sense that they are at the goal location. For example, for $\gamma = 0.86$, the action sequence of location B2 is SEO. Consequently, after the robot has sensed that it is at location B2, it first moves south (S), then moves east (E), and finally senses again (O). Locations whose action sequences are not used for getting from the start location (C1) to the goal location (J1) are left blank. Figure 3 (right) shows how often the robots visit each grid square during two million runs for three different values of γ. Darker colors indicate a larger number of visits. Thus, more risk-averse robots are more likely to stay on the road and close to the nominal path, which is possible due to the increased sensing frequency. That more risk-averse robots are more likely to stay on the road can be explained as follows: By staying on the road, the robots are likely able to avoid the large cost necessary for getting out of the mud, which risk-averse robots consider to be important. On the other hand, smaller sensing frequencies and attempts to cut the cor-

[1] There are some exceptions to this trend, for example, in the vicinity of the goal. This can be explained as follows: Risk-seeking robots assume that short action sequences that have a chance of reaching the goal location will indeed reach it. Thus, they execute these action sequences followed by a sensing action to confirm that they have reached the goal location. For example, for $\gamma = 1.40$, the action sequence of location I2 is SO. The robot hopes that it will drift to the goal location J1 as it moves south, although this is less likely than moving to location J2. More risk-averse robots are more cautious and execute longer action sequences. For example, for $\gamma = 1.86$, the action sequence of location I2 is SWO, which reaches the goal location with higher probability than the action sequence SO. This phenomenon and similar phenomena contribute to the small local minima in the graph of Figure 4.

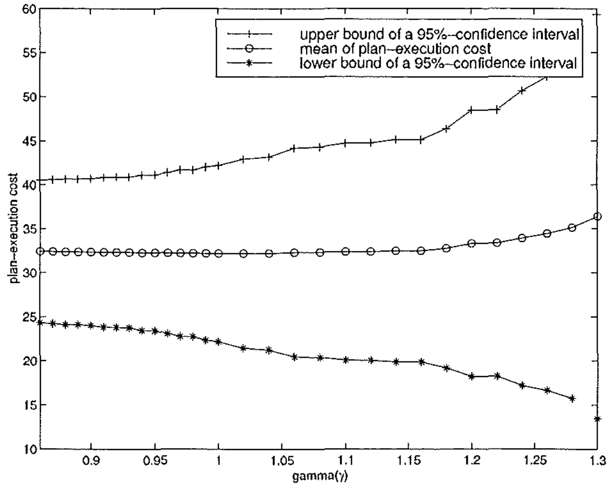

Figure 6: Mean with Confidence Interval

ners decrease the probability that the robots stay on the road but also decrease the plan-execution cost in the best case, which risk-seeking robots consider to be important. (In fact, the action sequences of the start location get longer and longer as the robots become more risk-seeking until the action sequences are able to move the robots to the goal location in the best case.) This explanation suggests that there is a mean-variance trade-off in our example domain. As pointed out earlier, mean-variance trade-offs are often used as crude but easy-to-understand explanations for trade-offs between minimizing the worst-case, average, and best-case plan-execution cost. For example, the graph in Figure 5 shows the mean of the plan-execution cost, and the difference of the upper and lower graphs in Figure 6 corresponds to four times the standard deviation of the plan-execution cost. The mean of the plan-execution cost increases but the variance decreases as the robots become more risk-averse and γ decreases from one to zero. The variance decreases because more risk-averse robots stay on the road and close to the nominal path. The mean-variance trade-off can be explained as follows: More risk-averse robots are willing to accept a larger mean of the plan-execution cost for a decrease in variance because they fear for the worst case. A small variance avoids a plan-execution cost that is much larger than average. Figure 6 illustrates this using the upper bound of a 95-percent-confidence interval (that is, mean plus twice the standard deviation) as an approximation of the worst-case plan-execution cost. The upper bound indeed decreases as the robots become more risk-averse since the decrease of the variance outweighs the increase of the mean. On the other hand, more risk-seeking robots are willing to accept a larger mean for an increase in variance since a larger variance promises a chance to realize a plan-execution cost that is much smaller than the average plan-execution cost. Figure 5 illustrates this using the lower bound of a 95-percent-confidence interval (that is, mean minus twice the standard deviation) as an approximation of the best-case plan-execution cost. The lower bound indeed decreases as the robots become more risk-seeking since the increase of the variance outweighs the increase of the mean. Our sensor planner is not only as easy to implement as the sensor planner that it extends but also almost as efficient. For $\gamma = 0.86$, our sensor planner expands 2,071 nodes and needs an average of 4.6 milliseconds per node expansion on a Sun Ultra 1 running Solaris 7. The original sensor planner by Hansen, that our sensor planner extends, corresponds to the case where γ approaches one. It needs 2,815 node expansions and 2.0 milliseconds per node expansion. For $\gamma = 1.40$, our sensor planner needs 5,808 node expansions and 5.2 milliseconds per node expansion. The number of node expansions depends on the sensing frequency. It increases as the sensing frequency of the optimal plans decreases. Our sensor planner has a slight run-time disadvantage per node expansion compared to the original sensor planner by Hansen because it has to calculate exponentials and logarithms, and its heuristic search method cannot calculate the heuristic values quite as efficiently as the original sensor planner.

Conclusions

We described a method for creating planners that find plans with maximal average utility for a given exponential utility function and thus reflect the risk attitude of rational human decision makers in high-stake decision situations better than plans with minimal average plan-execution cost. Our method generalizes the planning objectives of traditional planners from artificial intelligence that often either minimize the average or the worst-case plan-execution cost or, if the goal cannot be achieved for sure, maximize the probability of goal achievement. Our multiplicative planning-task transformation is a fast simple context-insensitive representation change that can be performed locally on various representations of planning tasks, including rule-based (STRIPS) representations (Fikes & Nilsson 1971) and (totally and partially observable) Markov decision process models. We applied the multiplicative planning-task transformation to an existing sensor planner. The resulting sensor planner is not only as easy to implement as the sensor planner that it extends but also almost as efficient. Our case study showed that the frequency of sensing depends on the trade-off between minimizing the best-case, average, and worst-case plan-execution cost. More risk-averse human decision makers tend to sense more frequently, and planners should reflect this behavior accurately. This research is beneficial for NASA since many NASA domains are high-stake planning domains. Human decisions makers often want to make high-level decisions themselves, for example, because they have background knowledge that planners do not have (or because of political considerations). However, there are disadvantages to having humans in the loop for every decision on-board of un-

$\gamma = 0.86$ (risk-averse robots)

	1	2	3	4	5	6	7	8	9	10	11
A								SSSO			
B		SEO	SEO	SEO	SEO	SEO	SO	SSO			
C	EO	EO	EO	EO	EO	EO	SEO	SO	SO	WSO	
D		NEO	NEO	NEO	EEEO	EEO	EO	ESO	SO	WSO	WWSO
E							EESO	ESO	SO	WSO	
F								SSSO	SO	WSO	
G								SSO	SO	WSO	
H						SSWO	SSWO	SO	WSO	WO	
I	SO	SWO	SWO	SWO	SWO	SWO	SWO	SWO	WO	WWO	WWWO
J	goal	WO	WO	WO	WO	WO	WO	WO			
K	NO	NWO	NWO	NWO	NWO	NWO	NWO	NWO	NWWO		
L							NNWO	NNWO			

$\gamma = 1.40$ (risk-seeking robots)

	1	2	3	4	5	6	7	8	9	10	11
A	SSEEEO				SSEO	SSEO	SSSO	SSSO			
B	SEEEEO			SEEO	SEEO	SEO	SSO	SSO	SSSSSSSWO		
C	EEEEEEESO			EEEESO	EEESO	EESO	SO	SO	SSSSSSWO		
D	EEEEEEESO			EEEESO	EEESO	EESO	ESO	SSSSSWO	SSSSSWO	WSSO	WWSSO
E	NNEEEEO	NEEEEEO	NEEEEO	NEEEO	NEEO	EESSSO	ESSSO	SSSSWO	SSSSWO	WSSO	WWSSO
F	SSSSO	SSSSO	SSSSO	SSSSWO	SSSSWWO	EESSO	ESSSO	SSSWO	SSSWO	WSSO	WWSO
G	SSSO	SSSO	SSSO	SSSWO	SSSWWO	SSSWWWO	SSSWWWWO	SSWO	SSWO	WSO	WWSO
H	SSO	SSO	SSWO	SSWWO	SSWWO	SSWWWO	SSWWWWO	SSWWWWWO	SWO	WWO	WWWO
I	SO	SO	SWO	SWWO	SWWWO	SWWWWO	SWWWWWO	SWWWWWWO	WO	WWO	WWWO
J	goal	WO	WWO	WWWO	WWWWO	WWWWWO	WWWWWWO	WWWWWWWO	WWWWWWWWO	WWWWWWWWWO	WWWO
K	NO	NO	NWO	WWWO	WWWWO	WWWWWO	WWWWWWO	WWWWWWWO	WWWWWWWWO	WWWWWWWWWO	WWWWWWWWWWO
L	NNO	NNO	NNWO	NWWO	NWWWO	NWWWWO		NWWWWWWO			

Figure 7: Optimal Sensor Plans

manned spacecraft, including decisions in crisis situations, due to the resulting time delay. In these cases, it is important that artificial intelligence planners make similar decisions as rational human decision makers. The research that we reported here is only a first step in the direction of planning with more realistic preference models by combining descriptive methods from utility theory with constructive methods from artificial intelligence planning. Future work includes how to plan with multiple attributes (such as time, energy, and money) and resource limits.

Acknowledgments

This paper presents a high-level overview of work reported in more detail in (Koenig & Liu 1999).

References

Bernoulli, D. 1738. Specimen theoriae novae de mensura sortis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae* 5. Translated by L. Sommer, *Econometrica*, 22: 23–36, 1954.

Desimone, R., and Agosta, J. 1994. Oil spill response for plannig in stochastic domains: The application of artificial intelligence planning technology. In *Proceedings of the Simulation Multiconference*.

Fikes, R., and Nilsson, N. 1971. STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence* 2:189–208.

Hansen, E. 1997. Markov decision processes with observation costs. Technical Report CMPSCI 97-01, Department of Computer Science, University of Massachusetts, Amherst (Massachusetts).

Howard, R. 1964. *Dynamic Programming and Markov Processes*. MIT Press, third edition.

Kahneman, D., and Tversky, A. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47:263–291.

Koenig, S., and Liu, Y. 1999. Sensor planning with nonlinear utility functions. In *Proceedings of the European Conference on Planning*, (in press).

Koenig, S., and Simmons, R. 1994. How to make reactive planners risk-sensitive. In *Proceedings of the International Conference on Artificial Intelligence Planning Systems*, 293–298.

Koenig, S. 1998. Representation changes for planning with exponential utility functions. In *Proceedings of the Symposium on Abstraction, Reformulation, and Approximation*, 79–84.

Lozano-Perez, T.; Mason, M.; and Taylor, R. 1984. Automatic synthesis of fine-motion strategies for robots. *International Journal of Robotics Research* 3(1):3–24.

Marcus, S.; Fernàndez-Gaucherand, E.; Hernàndez-Hernàndez, D.; Colaruppi, S.; and Fard, P. 1997. Risk-sensitive Markov decision processes. In et. al., C. B., ed., *Systems and Control in the Twenty-First Century*. Birkhauser. 263–279.

Nilsson, N. 1971. *Problem-Solving Methods in Artificial Intelligence*. McGraw-Hill.

Pell, B.; Gat, E.; Kaesing, R.; Muscettola, N.; and Smith, B. 1997. Robust periodic planning and execution for autonomous spacecraft. In *Proceedings of the International Joint Conference on Artificial Intelligence*.

von Neumann, J., and Morgenstern, O. 1947. *Theory of games and economic behavior*. Princeton University Press, second edition.

Watson, S., and Buede, D. 1987. *Decision Synthesis*. Cambridge University Press.

Whittle, P. 1990. *Risk-Sensitive Optimal Control*. Wiley.