

Generating Qualitatively Different Plans through Metatheoretic Biases

Karen L. Myers Thomas J. Lee

SRI International
333 Ravenswood Ave.,
Menlo Park, CA 94025
myers@ai.sri.com tomlee@erg.sri.com

Abstract

Current methods for generating qualitatively different plans are either based on simple randomization of planning decisions and so cannot guarantee meaningful differences among generated plans, or require extensive user involvement to drive the system into different sections of the overall plan space. This paper presents a cost-effective method for automatically generating qualitatively different plans that is rooted in the creation of *biases* that focus the planner toward solutions with certain attributes. Biases are derived from analysis of a *domain metatheory* and enforced through compilation into preferences over planning decisions. Users can optionally direct the planner into desired regions of the plan space by designating aspects of the metatheory that should be used for bias generation. Experimental results are provided that validate the effectiveness of the biasing method for reliably generating a range of plans with meaningful semantic differences.

Introduction

Automated planning tools hold much promise as decision aids for humans charged with producing plans for large-scale, demanding applications. The value of the tools lies with their ability to help humans understand the complexity of the underlying problem, providing guidance in the determination of a solution that is well-suited to their needs and concerns.

For many real-world applications, the search space is dense with solutions. Air campaign planning (Thaler & Shlapak 1995; Lee & Wilkins 1996) and travel planning (Linden, Hanks, & Lesh 1997) provide two examples. For these applications, it is not difficult to find a solution; rather, the challenge is to produce a solution that is tailored to the preferences and needs of the user. One means by which to help users with this task is to provide a set of *qualitatively distinct* options that vary in meaningful ways. A set of solutions of this type would help the user understand the range of possibilities available to him or her.

Current automated planning tools can readily generate different plans, for example through repeated runs with randomized choices at decision points. The differences among

such plans, however, are difficult to extract and not necessarily semantically meaningful. Furthermore, different users will have different notions of what constitutes ‘meaningful’ differences. For example, with travel plans a budget traveler might like to see options with a range of costs while the business traveler might like to see options that maximize accumulation of frequent flier points. Ideally, a system for generating qualitatively different plans would allow the user to specify dimensions along which he or she would like to see variation.

To address the problem of meaningfulness, hill-climbing methods could be used to generate successive plans until they differed sufficiently along some defined evaluation function. Two problems arise with this approach. First, the complexity of plan generation makes it expensive to iterate through many solutions. Second, defining evaluation criteria for complex planning domains is problematic: evaluation metrics are generally difficult to elicit, multidimensional in nature, qualitative rather than quantitative, and often subjective (Gil 1998). As such, defining the ranking function required to drive a hill-climbing process will be difficult in many domains.

Recent work on mixed-initiative, interactive, and advisable planning enables users to drive the process of generating qualitatively different plans (Ferguson & Allen 1998; Tate, Dalton, & Levine 1998; Myers 1996). With these frameworks, however, the user must be involved extensively in an ongoing role to articulate desired differences and to manage the space of options.

This paper describes a framework for generating qualitatively different plans in a fully automated fashion, but that can accept user guidance to influence the types of solutions that are generated. Rather than searching through the space of plans directly, the algorithm leverages a *metatheory* of the planning domain, introduced previously to support user advisability of a planner (Myers 1996). This metatheory provides an abstracted characterization of the planning domain that highlights key semantic differences among operators, planning variables, and instances. This abstraction provides the ability to filter out irrelevant differences when generating plans, focusing instead on distinctions that are guaranteed to be semantically meaningful. Based on analysis of the metatheory, *biases* are generated that are designed to focus the planner toward solutions with certain characteristics.

These biases are enforced by imposing preferences on planning decisions. The result is plans that are guaranteed qualitatively different *by design*, rather than possibly (though not necessarily) distinct by randomization.

While the biasing approach does not require user input, it can readily accept guidance from the user to target certain subportions of the overall plan space. For example, users could indicate that they want to see plans within a range of cost and time values, while insisting on traveling by airplane (rather than train, boat, or car).

The ability to generate qualitatively different plans is essential for the effective use of generative planning technology in producing solutions that satisfy user requirements. In particular, the technology should be viewed as a tool to assist human planners rather than as a replacement for them, with the user driving the kinds of solutions that the technology produces. In particular, after requesting a set of qualitatively different plans, the user would make recommendations on how to improve them. The idea is that the generated plans act as initial seeds, which users can subsequently refine to meet their needs (using, for example, *planning advice* (Myers 1996)).

Our biasing method for generating qualitatively different plans was implemented and evaluated within SIPE-2 (Wilkins 1988), a Hierarchical Task Network (HTN) planner (Erol, Hendler, & Nau 1994). The biasing method is not specific to HTN planning but rather applies to any form of operator-based planning. Experimental evaluation was performed in a travel domain that involves selecting itineraries, schedules, accommodations, modes of travel, and carriers for business and pleasure trips. The results show that the biasing method is effective for reliably producing a range of plans with meaningful semantic differences.

Measuring Qualitative Difference

The fundamental problem that we address is to produce n solutions to a given planning problem, such that those solutions do a good job of ‘covering’ the set of possible solutions. We are interested in small values of n (i.e., $2 \leq n \leq 5$), since a user of an automated planning system would generally consider only a small number of options at any point in time. We consider two notions of coverage, *dispersion* and *proximity*, both of which are grounded in the notion of a measurable distance between plans.

Plan Distance

We have opted to ground plan distance in *evaluation criteria*, which have the advantage of measuring aspects of the plan that are of significance to users. This contrasts with syntactic measures of plan distance (for example, the difference in plan length), which do not necessarily correlate with semantic differences among plans.

We assume a set of k evaluation criteria that define a k -dimensional space E^K in which to situate plans. For simplicity, we assume that evaluation values are normalized to lie in the range $[0, 1]$. With these evaluation criteria, we formally define our notion of *plan distance* in terms of the Euclidean distance between the corresponding points for those plans in evaluation space.

Definition 1 (Plan Distance) The distance between two plans P_1 and P_2 is defined to be

$$Dist(P_1, P_2) = \sqrt{\sum_{i=1}^k (Eval_i(P_1) - Eval_i(P_2))^2}$$

Dispersion

Dispersion is defined to be the average distance between plans in a plan set. As such, dispersion measures the degree to which solutions are spread apart from each other.

Definition 2 (Dispersion) The dispersion for a plan set P is defined to be

$$Disp(P) = \frac{\sum_{1 \leq i < j \leq n} Dist(P_i, P_j)}{\frac{n \times (n-1)}{2}}.$$

Proximity

Proximity is defined to be the average distance for a point in the evaluation space to its closest point in the evaluation of the plan set. As such, proximity measures the degree to which the solution set is ‘near’ all other points in the evaluation space.

Definition 3 (Proximity) The proximity of a point $e = \{e_1 \dots e_k\} \in E^K$ to a plan set $P = \{P_1 \dots P_n\}$ is defined to be

$$Prox(e, P) = \min_{1 \leq i \leq n} \left(\sqrt{\sum_{1 \leq j \leq k} (e_j - Eval_j(P_i))^2} \right).$$

The proximity for a plan set P is defined to be the average proximity to P from any point in E^K :

$$Prox(P) = Avg_{e \in E^K} (Prox(e, P)).$$

Closed-form solutions for computing proximity are not generally available in continuous evaluation spaces. We employ sampling methods to approximate proximity measures in this paper.

Discussion

Higher dispersion values generally indicate that a plan set does a better job of covering the extremities of the plan space. As such, highly dispersed plan sets are useful when users want to investigate the limits of the solution space. In contrast, lower proximity values correlate with plan sets that are more representative of the set of possible solutions in that they are ‘closer’ to all points in the evaluation space. As such, low proximity plan sets are valuable for presenting users with reasonable first-cut solutions that are likely to be close to what would be the user’s ideal solution.

For these reasons, we seek to generate solution sets that are highly dispersed with low proximity. These two objectives can be conflicting, depending on the distribution of plans through the evaluation space. For example, extremal points in the evaluation space will be maximally dispersed but are unlikely to yield low proximity values.

Domain Metatheory

A standard planning domain is modeled in terms of three basic types of elements: *individuals* corresponding to real or abstract objects in the domain, *relations* that describe characteristics of the world and individual world states, and *operators* that describe ways to achieve objectives.

The domain metatheory captures high-level attributes of planning operators, variables, and individuals, thus providing users with the means to describe desired solution characteristics in terms that are natural to them. The domain metatheory was developed originally to provide a language in which users could construct *advice* for a planning system (Myers 1996) but is not advice-specific: it describes general properties of elements of a planning domain and can be employed to support a variety of uses, including the generation of qualitatively different plans. The metatheory is built around three main constructs: *roles*, *features*, and *measures*.

A *feature* designates an attribute of interest for an operator that distinguishes it from other operators that could be applied to the same task. For example, among operators that can be used to refine tasks of moving from location X to location Y, there can be some that involve travel by air, land, or water; each of these media could be modeled as a feature. Because there can be multiple operators that apply to a particular task, features provide a way of abstracting from the details of an operator up to distinguishing attributes that might be of interest to users. Note that features differ from operator preconditions in that they do not directly restrict use of operators by the planner.

Related features are grouped into *feature categories*. For example, the features {Air Land Water} mentioned above define a **Transport-Media** category. Feature categories themselves can have interesting properties. Just as planning operators reflect a hierarchical structure, features and feature categories can be organized in hierarchical fashion. Certain categories may be *mutually exclusive* in that at most one feature from the category can be assigned to any given operator; this is the case for the feature category **Transit-Ownership** containing the elements {Public Private}. Other categories may support overlapping features; for example, there may be an operator that involves both Air and Land travel.

A *role* corresponds to a capacity in which an individual is to be used within an operator. For instance, a transportation activity within the travel domain could have roles such as Origin and Destination, and Carrier. Roles correspond to individual variables within a planning operator.

Feature categories can have associated *measures*. A measure corresponds to an ordering (possibly partial) of features within the category with respect to some designated criteria. For example, consider the feature category **Transit-Ownership** with features {Public Private}. For the measure COMFORT, the feature Private would rank higher than Public; for the measure AFFORDABILITY, the order would be reversed.

A single measure can be used across different feature categories. For example, AFFORDABILITY would apply to a broad range of feature categories. Thus, by expressing preferences on measure, it is possible to influence a broad range of plan generation decisions.

Vacation-Scope = {Overseas National Regional}
 AFFORDABILITY: (Overseas National Regional)
 TIME-EFFICIENCY: (Overseas National Regional)

Accommodation = {Hotel Motel Camp}
 COMFORT: (Camp Motel Hotel)
 AFFORDABILITY: (Hotel Motel Camp)

Transport-Media = {Air Land Water}
 AFFORDABILITY: (Water Air Land)
 TIME-EFFICIENCY: (Water Land Air)

Land-Transport-Mode = {Auto Bus Shuttle Taxi Train Limo}
 AFFORDABILITY: (Limo Train Auto Taxi Shuttle Bus)
 TIME-EFFICIENCY: (Bus Shuttle Auto Limo Taxi Train)
 COMFORT: (Bus Shuttle Taxi Train Auto Limo)

Transit-Ownership = {Public Private}
 COMFORT: (Public Private)
 AFFORDABILITY: (Private Public)
 TIME-EFFICIENCY: (Public Private)

Transit-Capacity = {Solo Shared}
 COMFORT: (Shared Solo)
 AFFORDABILITY: (Solo Shared)

Figure 1: Sample Feature Categories and Associated Measures from the Travel Domain

Figure 1 presents an excerpt from the metatheory for the travel domain that shows sample feature categories and associated measures. Each block defines a feature category, with the first line listing the name of the feature category followed by its constituent features. The remaining lines declare a measure associated with that feature category, and provide the ranking of the features for that measure and category. Here, we show only measures that completely order the features (although partial orders are possible).

Just as measures can be employed to rank features (and hence operators with those features), they can also be employed to rank instances. For measures on instances, an ordered set of *measure values* is defined. For each measure, a given individual can (optionally) be assigned one of these values, thus inducing a partial order over instances. In the travel domain, for example, the measure AFFORDABILITY has the values (*Extravagant Expensive Moderate Inexpensive Cheap*) in increasing order from left to right. The individual Ritz of class Hotel has the AFFORDABILITY value *Extravagant*, while the individual Motel6 of class Motel has the value *Cheap*, thus Motel6 ranks higher than Ritz with respect to AFFORDABILITY.

We define the *domain* of a measure to be the set of (partially) ordered values employed by the measure. For measures defined over feature categories, the domain is the set of features that comprise the feature category. For measures defined over instances, the domain is the set of measure values that can be assigned to instances.

Biasing Algorithm

Our approach to generating n qualitatively different plans uses the domain metatheory to establish n sets of *biases* that can direct the planner toward different sections of the overall plan space. In addition to n , our algorithm takes as input a subset of the measures provided by the domain metatheory. These measures can be selected by the user in order to influence the types of qualitative differences among the generated plans.

Bias and Region Creation

The biasing method involves partitioning the domains of the selected measures into *intervals*, and then grouping the intervals (one from each measure) into *cases* designed to force the planner into different sections of the overall plan space. We use the term *region* to refer to the collection of biases for a given case.

A bias is defined by a measure and an interval of values within the domain of the measure. To simplify manipulation and bias enforcement (described below), a proportional, order-preserving mapping from each measure domain onto the interval $[0, 1]$ is defined. For simplicity, we restrict attention to connected intervals.

Different strategies for partitioning and grouping are possible. For the results described in this paper, each measure domain is partitioned into n subintervals of equal length (relative to $[0, 1]$). A set of n regions is created by, for each measure, randomly assigning each of the n intervals for the measure to a different region.¹ This strategy provides systematic coverage of measures in that each interval appears in exactly one region. The use of random selection for region assignment is important to avoid potential problems that can arise due to correlations among measures.

Bias Enforcement

Biases are enforced in a heuristic manner: rather than imposing hard constraints, choices available to the planner are ordered to reflect the preferences inherent in the biases. Because the enforcement of biases prioritizes choices rather than filtering them, it does not restrict the set of plans that could be produced. As such, the biases can be viewed as *relaxable* constraints on plan generation.

Two types of planning decisions are influenced by biases: *operator selection* and *instance selection*.

Operator Selection Multiple operators could be used to refine a goal within a plan.

Instance Selection Instantiation of variables is performed in two situations. First, instances are selected for variables left uninstantiated after the original task has been reduced

to a primitive task network. Second, certain operators dictate that variables be instantiated, although a unique value may not be determined by accumulated constraints on that variable. (Such *early commitment* is often used to reduce the complexity of constraint reasoning (Myers & Wilkins 1998).)

For each of these decision types, a scoring function defines an ordering that reflects the degree to which the choices satisfy the stated biases for the current case. Choices are made according to this order, with random selection amongst choices with equivalent scores. In particular, the most highly ranked choice will be selected first; successive choices may be made in the event that backtracking through that decision occurs.

The scoring functions take into account whether a given bias is *relevant* to a given instance or operator. We say that a bias is relevant to an instance I if the bias measure is defined for I . Similarly, a bias is relevant to an operator Opr if the operator has some feature in a feature category F for which the measure is defined.

The calculation of *bias distance* lies at the heart of the scoring mechanism. Bias distance measures the extent to which a choice satisfies a stated bias. For an instance I , the function $BDist(I, B)$ is the distance d between the interval of measure M defined by B and the measure value for I in M . If the measure is not relevant to the instance, then the distance is defined to be \perp ; otherwise, the distance is the absolute difference between the measure-defined mapping of I and the interval onto $[0, 1]$. The function $BDist(Opr, F, B)$ for bias distance for an operator Opr and feature category F is defined similarly.

Bias distances are scored so that values in the interval are heavily rewarded while values outside the interval are penalized in proportion to the distance from it:

$$BDistScore(d) = \begin{cases} 0 & \text{if } d = \perp \\ 1 & \text{if } d = 0 \\ -d & \text{otherwise} \end{cases}$$

The score for an instance I relative to a region R is defined to be the sum of the score for I relative to each bias B in R :

$$Score(I, R) = \sum_{B \in R} InstScore(I, B)$$

$$InstScore(I, B) = BDistScore(BDist(I, B))$$

The score for an operator Opr is defined similarly. However, an operator may have features from multiple feature categories that are relevant to the stated biases. For this reason, the average score across those relevant feature categories (denoted by F_B) is used:

$$Score(Opr, R) = \sum_{B \in R} OprScore(Opr, B)$$

$$OprScore(Opr, B) =$$

$$Avg_{F \in F_B} BDistScore(BDist(Opr, F, B))$$

¹Measures with fewer than n elements are problematic in that there will be multiple intervals that contain the same domain elements. As a result, it is possible that multiple regions will contain combinations of intervals that are effectively identical. To avoid this problem, measures with fewer than n elements are combined (for region creation only) into an artificial composite measure defined by the cross-product of their domains. Partitioning and grouping are performed relative to this composite measure.

Evaluation

To evaluate our algorithm for biasing plan generation, we compared dispersion and proximity values for plan sets of varying size that were produced using biases to those generated by simply randomizing planning decisions within SIPE-2 (i.e., choice of operators and variable instances).² Generated plans ranged in length from 2 to 14 actions.

The evaluation space was defined by three evaluation criteria:

- the overall *cost* (measured in dollars),
- the overall *time* (measured in hours),
- the *distance* covered (measured in kms).

Evaluations for generated plans were normalized relative to assigned minimum and maximum values for each criterion. Minimal values were derived through analysis of the planning domain; maximum values were obtained by adding a small buffer (roughly 5%) to the maximum values seen throughout the course of experimentation (encompassing several thousand plans). Determining exact maximum bounds was not feasible because of the overall size of the plan space.

Figures 2 and 3 display the results for an experiment that evaluated four different generation strategies for plan sets:

- biases derived from the 2 measures {AFFORDABILITY, TIME-EFFICIENCY}.
- biases derived from the 2 measures {AFFORDABILITY, COMFORT}
- biases derived from the 3 measures {AFFORDABILITY, COMFORT, TIME-EFFICIENCY}
- randomization of planning decisions (*no biasing*)

For each strategy, 100 plan sets of sizes 2 through 9 were generated (i.e., 800 plan sets in total for each strategy).

Figure 2 plots the mean dispersion value for each trial of each method as a point, embedded in a vertical line showing the standard deviation. Figure 3 provides a similar plot for estimated proximity values, obtained by averaging proximity values for points on a uniform k -dimensional grid with 10 grid points within the evaluation space (i.e., for a total of 10^k sample points). To provide perspective for these charts, note that the maximal plan distance in the evaluation space is 1.73 while the mean plan distance is .58.³

The experimental results show that the two biasing methods produce solution sets with significantly better dispersion (i.e., *higher*) and proximity (i.e., *lower*) measures, from $n = 2$ through $n = 6$. Even more significant, however, are the differences in standard deviation which show that the

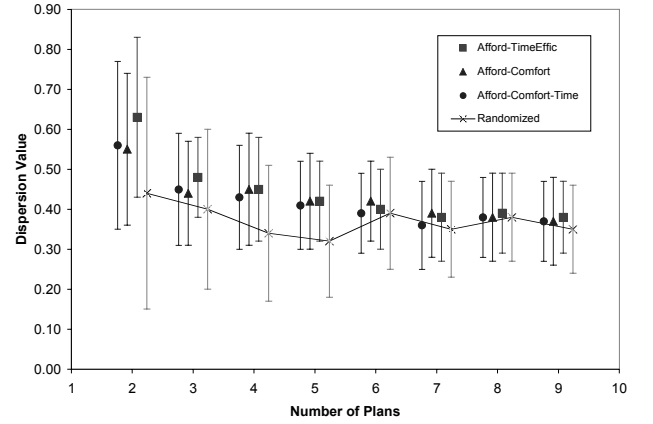


Figure 2: Plan Set Dispersion Measures for 100 Trials

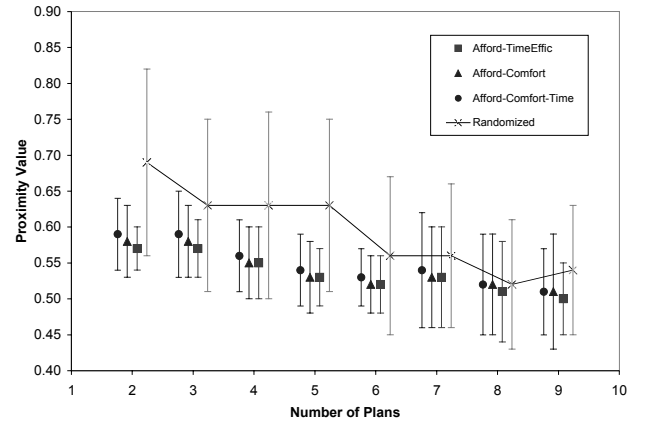


Figure 3: Plan Set Proximity Measures for 100 Trials

biasing approaches vary much less than the randomized approach. These effects are highly pronounced for $2 \leq n \leq 6$, tapering off as n increases beyond that point. As one would expect, dispersion and proximity values for all methods tend to decrease as n increases.

These experiments validate the idea that by using biases rather than randomization to direct the search process, plan sets can be generated for small values of n with significantly better coverage, and for larger values of n that are no worse. Furthermore, the *reliability* is far superior in that the user is unlikely to get bad results (i.e., low dispersion or high proximity) since the standard deviation is significantly smaller.

Future Work

As already noted, a variety of strategies is possible for constructing regions of biases. While the particular method selected worked well within our experiments, alternative approaches merit investigation. For example, the use of *dynamic feedback* among regions could lead to better results: after creating a solution for one region, its evaluation scores could lead to adaptation of other regions to promote better

²An alternative form of evaluation would be to compare the amount of time required to achieve certain levels of coverage. This definition seems less suitable: while relative comparison of dispersion and proximity values are possible, the interpretation of absolute values for these two measures is unclear.

³For a normalized k -dimensional space, the maximum distance between points is \sqrt{k} , and the expected mean distance between points is $\sqrt{\frac{k}{9}}$.

proximity or dispersion.

Additional aspects of the domain metatheory could be leveraged as well. For example, *role* information is not directly used in the current framework but could provide another input to the overall biasing algorithm. Thus, users could indicate that they want to see a range of options for air carriers, for example.

Another key area for future work is in highlighting differences among plans. Such explanatory capabilities are critical for helping users make informed choices about their options. The domain metatheory provides a means for abstracting from the details of planning decisions and could be used to concisely summarize key differences between plans. For example, summaries along the following lines could be readily extracted using the domain metatheory:

Accommodations were chosen that ranked high on *affordability*, *United* was chosen as the *carrier* for air travel, and *public transit* was used whenever possible.

Discussion

Focused Biasing The experiments presented above employed biasing over the full extent of measure domains in order to provide coverage of the entire solution space. In general, users will be interested in restricting attention to subregions of the overall space (e.g., overseas vacations in the low to medium affordability range). The biasing approach presented here can readily accommodate this narrowing of the solution space by allowing the user to designate portions of the measure domains over which biases should range.

Measure-Measure Correlation Certain correlations exist among the measures defined in the metatheory for the travel domain. The measures AFFORDABILITY and COMFORT have a strong inverse correlation for feature categories and instances where they are both defined. However, there are several domains where one is defined but not the other. COMFORT and TIME-EFFICIENCY have a mild positive correlation when they overlap, but such overlaps are limited. The use of AFFORDABILITY and TIME-EFFICIENCY overlap significantly, but show a range of positive, negative, and mixed correlation,

The specific biasing method described in this paper employs random selection among measure intervals to avoid effects due to correlations among measures. However, knowledge about correlations could be used as the basis for more sophisticated biasing methods that try to leverage known correlations among measures.

Measure-Evaluation Correlation Correlations exist between certain of the metatheory measures and the evaluation criteria used to adjudicate qualitative difference. While this correlation can be strong (as between the measure AFFORDABILITY and evaluation criteria *Cost*), it is quite weak for others. Such linkage is to be expected, since the metatheoretic biases used to distinguish planning options should relate to evaluation criteria of interest to users.

Biasing Cost The cost of creating biases and regions is negligible. Similarly, enforcement of biases adds only the

cost of identifying the most highly ranked choice among available options.

Related Work

Work to date on generating qualitatively different plans that is not rooted in randomization relies extensively on the user to drive the planner to different sections of the plan space. For example, TRIPS (Ferguson & Allen 1998) and O-Plan (Tate, Dalton, & Levine 1998) support mixed-initiative planning in which a user can explore different planning options by either explicitly telling the system what to do next, or imposing constraints on the plan or planning process. Similarly, the Advisable Planner (Myers 1996) provides the means to sketch desired characteristics of a single plan at high levels of abstraction.

The Automated Travel Assistant (ATA) (Linden, Hanks, & Lesh 1997) generates a sequence of plans that vary in response to user feedback on the content of earlier plans. The objective within this system is to find an acceptable solution quickly that ranks highly with respect to stated and inferred user preferences. The basic approach involves incrementally building an improved model of user preferences by analyzing user critiques of solutions that are suggested by the problem-solving system. Each generated solution reflects the evolving preference model, in a way that parallels our method of generating a single plan that reflects certain biases. However, the ATA frames planning as an *attribute selection* problem: finding values for a fixed set of variables. This problem is very different from the open-ended plan generation employed by the biasing algorithm described in this paper.

Conclusions

For many significant planning domains, users are reluctant to relinquish full control of the planning process. Rather, they would like automated planning tools that can help them understand their options. Tools that generate plans with guaranteed significant semantic differences can be of great value to these users, enabling them to make an informed selection from the space of possible solutions.

The biasing technique presented in this paper provides a simple, low-cost mechanism to reliably generate plans with meaningful semantic differences. An additional advantage of the method is that it enables the user to direct the overall process without having to be involved continuously in detailed decision-making.

Acknowledgments This research has been supported by DARPA Contract F30602-95-C-0259 as part of the ARPA/Rome Laboratory Planning Initiative.

References

- Erol, K.; Hendler, J.; and Nau, D. S. 1994. Semantics for hierarchical task-network planning. Technical Report CS-TR-3239, Computer Science Department, University of Maryland.

- Ferguson, G., and Allen, J. 1998. TRIPS: Towards a mixed-initiative planning assistant. In *Proceedings of the AIPS Workshop on Interactive and Collaborative Planning*.
- Gil, Y. 1998. On evaluating plans. Technical report, USC/ISI Expect Project.
- Lee, T. J., and Wilkins, D. E. 1996. Using SIPE-2 to integrate planning for military air campaigns. *IEEE Expert* 11(6):11–12.
- Linden, G.; Hanks, S.; and Lesh, N. 1997. Interactive assessment of user preference models: The Automated Travel Assistant. In *Proceedings of the Sixth International Conference on User Modeling*.
- Myers, K. L., and Wilkins, D. E. 1998. Reasoning about locations in theory and practice. *Computational Intelligence* 14(2).
- Myers, K. L. 1996. Strategic advice for hierarchical planners. In Aiello, L. C.; Doyle, J.; and Shapiro, S. C., eds., *Principles of Knowledge Representation and Reasoning: Proceedings of the Fifth International Conference (KR '96)*. Morgan Kaufmann Publishers.
- Tate, A.; Dalton, J.; and Levine, J. 1998. Generation of multiple qualitatively different plans. In *Proceedings of the Fourth International Conference on AI Planning Systems*.
- Thaler, D. E., and Shlapak, D. A. 1995. Perspectives on theater air campaign planning. Technical report, Rand Corporation.
- Wilkins, D. E. 1988. *Practical Planning: Extending the Classical AI Planning Paradigm*. Morgan Kaufmann.