A Benchmark Model for Decision-Theoretic Planning with Constraints

Kendra Renee Gehlbach*, **Brandon Laracuente***, **Cynthia Isenhour**†, **Judy Goldsmith***, **Beth Goldstein**‡
Mirosław Truszczyński*, {krgehl2,bglara2,Cynthia.Isenhour}@uky.edu, {goldsmit,mirek}@cs.uky.edu, bethg@uky.edu
University of Kentucky, Lexington, KY

Abstract

Work on planning under uncertainty has progressed from slow solvers on small models to solvers that can handle huge models through compact, factored representations and clever use of data structures. The next steps are to consider models with factored actions and complex constraints on action-value combinations and action/state combinations.

This paper introduces a simplified model of social welfare in Kentucky. The domain includes factored states and actions, and a variety of forms of constraints. We offer it as a benchmark for planning algorithms for factored Markov decision processes with complex actions and constraints.

1 Introduction

We present here a simplified model of social welfare in Kentucky. We refer to this domain as the welfare-to-work or WtW domain. Despite restrictions imposed on this model, it is of manifold interest. First, the model (and even more, the process of constructing it) helps in understanding the functionality needed in a modeling language to express features of a complex domain such as WtW. Second, the model offers a challenge and a test-bed for decision-theoretic planing algorithms, with the size of the model going beyond that of typical benchmarks used in decision planning literature, yet small enough to allow hope that it might be tractable by decision-theoretic algorithms currently under development both by our group and by others. Last but not least, the model concerns a socially important domain and, despite the limitations we imposed, captures its essential aspects.

* Department of Computer Science.
† Department of Anthropology.
‡ Department of Educational Policy.

The work we present here is part of a research project to develop decision-support software for case managers in the Kentucky social welfare system. Welfare case managers help their clients plan their activities such as volunteer work, participation in job readiness programs, substance abuse counseling, study in high school or college, etc. The case manager and client negotiate a *contract*, which specifies what official activities the client will engage in and what benefits she or he will receive. (Locally, most welfare recipients are single mothers, and a majority of the case managers are women. We will tend to use the female pronoun generically.)

Case managers' advice is guided by rules and regulations which specify what the client must do to stay eligible for the benefits, as well as the case manager's assessment of the client's goals, interests, skills, willingness to learn and other characteristics. It is also guided by the case manager's perception of how these characteristics, as well as the client's past performance, may affect the probability of success of the client in the proposed activities.

Decision-theoretic planning, the art of optimizing expected utility in the face of stochastic actions, describes most decision-making in the real world. Markov decision processes [7] provide a standard formalism for decision-theoretic planning, effectively capturing stochasticity of the domain of interest. However, there are many factors beyond stochasticity that complicate decision-making: constraints on states, actions and their results, uncertainty of the current state, and limited rationality, to name a few.

To model WtW domain, that is, the case managers' knowledge of possible activities for their clients and the predictors and effects of those actions, we chose the formalism of factored Markov decision processes [2]. We extend it by a formalism to represent constraints, addressing one of the limitations of Markov decision process model listed above.

The WtW system is subject to frequent changes as rules and regulations are revised, services appear or are eliminated, and job availability varies. The full model will need to handle such changes gracefully. We describe here a simplified version of the WtW domain model, developed while we are in the process of building a full and accurate model of the WtW system. This benchmark is suitable for testing algorithm prototypes.

There is a related paper, [6], also submitted to this workshop. The emphasis in that paper is on the interdisciplinary process. Both papers discuss the WtW system and our elicitation process. However, the emphasis in this paper is the simplified model that we will be making available to the community as a benchmark for MDP solvers that handle constraints on states and actions.

The paper is organized as follows. Section 2 outlines the underlying model, factored Markov decision processes, that we chose to model the case managers' knowledge of possible activities for their clients, and the predictors and effects of those actions. Section 3 outlines basic characteristics of WtW domain. The subsequent section gives a brief overview of the elicitation process we developed for this project, and the Bayesian network fragment that is elicited by our software. Section 5 discusses the simplifications we made to the elicited data in order to get the small, stable model that is the focus of this paper. The appendix gives a list of the actions and attributes of the model, as well as samples of the constraints and a sample probability table.

2 Decision-Theoretic Planning with Constraints

A *Markov decision process (MDP)*, in the AI world, consists of a finite set of *states*, S, a finite set of actions, A, transition probabilities, P, for each state/action pair being mapped to a next state, and a utility function $r : S \times A \rightarrow \mathbf{R}$. When state spaces and transitions are represented explicitly, we call the representation *flat*. If the state space is represented as a product of state variable values, and the transition function is represented by dynamic Bayesian networks or some equivalent model, we say that the representation is *factored*.

A *plan* or policy for an MDP is a mapping $\pi : S \rightarrow A$ from states to the recommended action for that state. The most common approaches for finding an optimal policy are either to transform the problem into a linear programming problem, or to use a dynamic programming-based algorithm such as value iteration.

There are several approaches that have been tried already for decision-theoretic planning with constraints. One is to convert the problem to a SAT or Constraint Satisfaction Problem (see, for example, Dechter and Mateescu's work [3]) and use the corresponding heuristics. Alternately, if the preferred solution method is linear programming, and the constraints map to linear constraints (or even quadratic constraints, see [1]), it is straightforward. It is our intuition that more algorithms are likely to be forthcoming. Therefore, it is important to have benchmarks on which to compare such algorithms.

3 The Benchmark Model

The WtW planning domain was described in a paper on the knowledge acquisition process [5]. In this domain, states are described by means of 50 client characteristics, or attributes, whose domains consist of two or more values. A state can change as a result of actions with probabilistic effects. Our knowledge acquisition process identified 16 different elementary actions, or activities that clients might engage in. At any time, actions available to the client are combinations of one or more activities. Our domain description also contains descriptions of constraints, which specify sets of actions and client characteristics that are not allowed.

In its full description, the WtW model is a formal representation of a rich and important application domain in which decision making is of fundamental importance. It will be the input for automated planning algorithms that will determine optimal courses of action (policies) for individual clients based on their characteristics and goals. Our goal here is to derive from the full model of the WtW domain a restricted version that preserves most of its richness and basic characteristics, yet is small enough to be useful as a testing ground for prototype planners. The restrictions we adopted focus on all components of the full WtW domain: attributes, their domains, actions, and constraints. These restrictions are described in detail in Section 5.

4 The Elicitation Process

The data on the full WtW domain was elicited from welfare case managers. A pilot study was completed using the action "Get GED", where case managers listed client characteristics which would affect the result of this action and which were affected by the results of this action. The list of attributes was then combined with the results of previous expert interviews, then simplified into a list of fifty variables.

This elicitation methodology was then repeated for the other fifteen actions with the help of special elicitation software.

4.1 High Level Elicitor (HELL) Tool

The special purpose software used during the elicitation was the High Level Elicitor (HELL). The HELL tool allows the user to assign attributes which contribute to the success or failure of an action as well as attributes which are affected by the outcomes of that actions.[1]

In early discussions with the case managers, it turned out that they specified attributes representing basic needs, such as food and shelter, as affecting the likelihood of success of every action. In fact, these needs are the first that the case managers handle for new clients. So we made the simplifying assumption that clients had basic needs met before attempting any of the other activities. Thus, those characteristics were omitted from the elicitation process.

The HELL elicitation process allowed the case managers to define the relative importance of each attribute to completing each action and to specify which attributes are affected by said actions. Subsets of this information were used to make simplifications in the benchmark model.

4.2 Bowties

We call the underlying representation used in HELL a *bowtie model*. The best way to understand this representation is by looking at an example. Consider the action *Literacy Training* from the simplified model. The contributing attributes were determined to be the client's commitment, if they have a learning disability, and their literacy level. The affected attributes are the client's aptitude, goals, and literacy level. Given these relations the bowtie network is shown in Figure 4.2.

While Figure 4.2 is the most human-intelligible representation of the Bayesian network fragment, all of the output nodes have earlier versions of themselves as implicit parents. Figure 4.2 shows all parents explicitly.

5 The Benchmark

The simplified model was created by expanding the simplifying assumptions of the HELL model, decreasing the number of actions, reducing the number of attributes, and decreasing the attributes' domains.

[1]We used the name "HELL" as an inside joke, saying, "We have to change this before the case managers use the software." However, it turned out to amuse the case managers, so we kept the name.

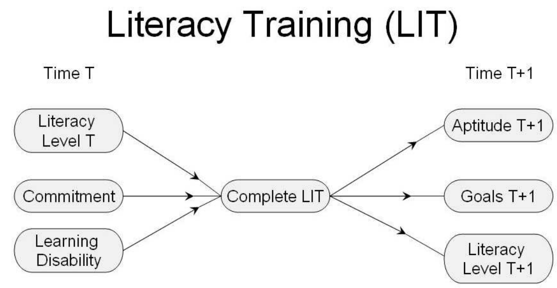

Figure 1: Bowtie model of *Literacy Training*

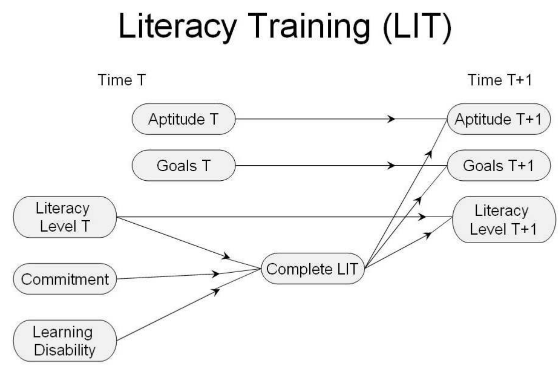

Figure 2: The implicit DBN

Aside from the assumption made in HELL that all critical barriers have been addressed, the simplified model adds the assumption that the client is at least 21 years old in a single-parent household with the youngest child at least 6 years old. This assumption permitted the number of hours in the work requirement be set at thirty, rather than varying based upon the age of the youngest child and the number of parents within the household, and reduced the number of possible methods of attending secondary school.

5.1 Simplifications of the Attributes and Actions

Several methods were used to reduce the number of actions from sixteen to seven. The first simplification method used was to combine similar actions that often shared pools of permissible time. For example, community service and volunteer service together are permitted for twelve months, rather than having limits for each type of service. In this way community service and volunteer service were combined into one action (Community and Volunteer Service), GED and high school into another action (Secondary Education), and community college, university, vocational school, and short term training into a third action (Post-Secondary

Education). The second method of reducing actions was eliminating actions which were limited to small subsets of the populations, such as substance abuse counseling and English as a second language.

Simplification of attributes was more involved. Some of the attributes were already obsolete due to the simplification of the actions; all attributes that appeared only in actions which had been removed were themselves removed. Going back to the elicitations from the case managers, the three or four client attributes that were listed as the most important for the input and the output of each action were selected. Where actions had been combined, attributes were chosen if they appeared in the lists of attributes for all of the combined actions. Of the remaining attributes, those that appeared only once were removed from the list. Attributes that were listed as the output for actions but were seldom or never used as inputs of actions were removed, since it would be pointless to keep track of changes that never affected an outcome. Attributes which were similar were combined into one attribute. As an attribute was removed, the next highest attribute for the input or output of that action bowtie was used to replace it, until we had a stable list of actions, each with three or four attributes as inputs and three or four attributes as outputs.

The next restrictions were on attribute variable domains. Several of the domains already contained only two or three values. They were not changed. Most of the rest of the domains were simplified by combining values until they contained only two or three values. The domains of the variables keeping track of the number of hours for each action were restricted to multiples of ten. In terms of schooling, the simplifying assumption was made that all students would attend school for the maximum number of hours, that being ten hours a week for secondary school students and full time, completely fulfilling the work requirement, for post secondary students. This reduced the previous domains of 0 to 10 hours and 0 to 30 hours, respectively, to Boolean domains.

5.2 The Constraints

Both the full and simple WtW domains have constraints, which provide a mapping of the space of actions and attribute variable values $\{0, 1\}$. If a set of elementary actions and client characteristics is mapped to 0, that means it is not allowed. These constraints are written as clauses of the logic PS+ [4], a language which can be processed by its own tools or translated (with no significant increase in size) into smodels [8], and processed with tools developed there. As with the actions and attributes, the constraints for the simplified model were condensed. Many constraints were perforce simplified when actions or attributes they discussed were removed or combined with other actions. Many of the simplifications discussed in Section 5.1 were done specifically to reduce the number of constraints. For example, making the assumption that the client was at least 21 removed constraints specifying that high school cannot be attended by over 19 years old, that GED classes count as full-time if the client is no older than 20, and that GED classes can count for no more than ten hours if the client is older than 20. The constraints of the benchmark WtW model can be categorized into three classes of constraints.

The first and most far-reaching constraint is that the work requirement must be fulfilled. This can be accomplished in several ways. Often this is accomplished when the client accumulates thirty countable or allowable hours from secondary education, community/volunteer service, literacy training, job and secondary education, and literacy training. Certain activities also automatically fulfill the work requirement, including secondary school attendance and group job search.

The second group of constraints pertain to duration. Most of the activities are only permitted for certain periods of time, not indefinitely. Group job search, for example, is only permitted one month out of every twelve. Community/volunteer service is permitted for no more than six months at a time and twelve months total. Other activities are permitted indefinitely, but only count toward fulfilling the job requirement for certain periods of time. Secondary school, for example, is countable for ten hours, but only for a period of twelve months. Any GED classes taken after that might be to the client's benefit, but do not count towards fulfilling the work requirement. Post-secondary school is countable for twelve months, then allowable for another twelve. Beyond those two years, a client must fulfill their work requirement with other activities while attending post-secondary school. Many times clients will transition from post-secondary school to job and post-secondary education, wherein they work an easy job on campus while completing their degree.

The final type of constraint is an eligibility constraint. Certain programs require that a client have a specific attribute value or combination of attribute values before she is permitted to take part in that program. For example, a client cannot take secondary education if she already has a high school degree or a GED. Similarly, a client cannot take post-secondary education, or job and post-secondary education, unless she already has a high school degree or GED. A client also cannot be permitted to attend literacy training if there is no problem with her literacy level. Finally, a client cannot attend group job search if she already has a job, a constraint which occurs with both on-the-job training and job and post-secondary education.

6 Conclusions

We have presented a brief overview of the United States social welfare system as implemented in KY, elicited from Fayette County welfare professionals, and simplified into a small and mathematically formal model. While this simplified model does not capture the complexity through which clients must navigate, it is sufficiently complex to be suitable for testing algorithms for decision-theoretic planning with constraints.

The real world is awash in uncertainty and constraints. We hope that AI planning will catch up with reality enough to begin planning under those conditions. If we, or others, are able to contribute planners that make social welfare case managers' jobs easier, all the better. (For a discussion of social issues, both within our research group and between the researchers and the case managers, see [6].)

The lists of attributes and actions, and a sample of the constraints, are included in the appendix. The full model will be available shortly from our web page, http://www.csr.uky.edu/wtw/.

Acknowledgments

This work was partly supported by NSF grant ITR-0325063 and a generous grant from the Dean of the College of Engineering at the University of Kentucky. The opinions expressed here are those of the authors and do not represent the Foundation, the University, or any social welfare offices.

References

[1] Christopher Amato, Daniel S. Bernstein, and Shlomo Zilberstein. Solving pomdps using quadratically constrained linear programs. In *Proc. 9th FLAIRS*, 2005.

[2] C. Boutilier, T. Dean, and S. Hanks. Decision-theoretic planning: Structural assumptions and computational lev erage. *Journal of Artificial Intelligence Research (JAIR)*, 11:1–94, 1999.

[3] Rina Dechter and Robert Mateescu. Mixtures of deterministic-probabilistic networks and their and/or search space. In *Proc. UAI '06*, 2006.

[4] D. East and M. Truszczyński. Predicate-calculus based logics for modeling and solving search problems. *ACM Transactions on Computational Logic*, 2006. To appear, available at `http://www.acm.org/tocl/accepted.html`.

[5] Krol Kevin Mathias, Cynthia Isenhour, Alex Dekhtyar, Judy Goldsmith, and Beth Goldstein. Eliciting and combining influence diagrams: Tying many bowties together. Technical Report TR 453-06, University of Kentucky, March 2006.

[6] Krol Kevin Mathias, Cynthia Isenhour, Alex Dekhtyar, Judy Goldsmith, and Beth Goldstein. When domains require modeling adaptations, 2006. Also submitted to this workshop.

[7] Martin L. Puterman. *Markov Decision Processes.* John Wiley & Sons, New York, 1994.

[8] P. Simons, I. Niemelä, and T. Soininen. Extending and implementing the stable model semantics. *Artificial Intelligence*, 138:181–234, 2002.

Appendix

We provide here a description of several components of the model: actions, attributes and constraints.

The Actions

- Secondary Education
- Post-Secondary Education
- Job and Post-Secondary Education
- Community/Volunteer Service
- Literacy Training
- Group Job Search
- On the Job Training

The Attributes

- Committment (none, low, high)
- Goals (false, true)
- Skills (false, true)
- Maturity (none, low, high)
- Literacy Level (remedial, average, high)
- Aptitude (false, true)
- Learning Disabilities (none, low, high)
- Previous Employment/Work Ready (bad, average, good)
- Education Level (pre-high school, high school diploma/GED, some post-secondary, post-secondary degree)
- Consecutive Months Taken Community/Volunteer Service (0–6)
- Months Since Last Group Job Search (0–11)

- Hours Taken (0, 10, 20, 30)[2]

- Complete Action (not yet attempted, in progress, failed, succeeded)[3]

Sample Constraints

These are some of the constraints on eligibility for actions, written first in English, then in ps+. Four non-numeric domains are mapped to numbers, following the order of values given within parentheses in the attributes above, with 0 being the one farthest left. We use variables t, m and n: the domain of t is all possible times; the domain of m is the possible values of the Education Level variable, and the domain of n is the possible values of the Literacy Level variable.

One cannot attend secondary education unless one has pre-high school education:
$complete\_action\_secondary\_education(t, 1),$
$education\_level(t, m), m > 0 \rightarrow .$

One cannot attend post-secondary education with a pre-high school education level:
$complete\_action\_post\text{-}secondary\_education(t, 1),$
$education\_level(t, 0) \rightarrow .$

One cannot do job and post-secondary education with a pre-high school education level:
$complete\_action\_job\_and\_post\text{-}secondary\_education(t, 1),$
$education\_level(t, 0) \rightarrow .$

One cannot take literacy training unless one's literacy level is remedial:
$complete\_action\_literacy\_training(t, 1),$
$literacy\_level(t, n), n > 0 \rightarrow .$

Sample Probability Table

This probability table corresponds to the fragments given in Figure 4.2. Full probability tables for all success nodes and outcome nodes will be available on our website by the time the UAI '06 workshop.

[2]There is an Hours Taken attribute for each of Job and Post-Secondary Education, Community/Volunteer Service, Literacy Training, and On the Job Training. The rest of the actions are assumed to be taken either full-time or for the maximum number of hours possible.

[3]There is one Complete Action attribute for each action. For example, Complete Action Secondary Education.

Aptitude time t	Complete Literature Training	Aptitude time $t+1$ $= T$	Aptitude time $t+1$ $= F$
T	Succeeded	0.9	0.1
T	Failed	0.9	0.1
T	In Progress	0.9	0.1
F	Succeeded	0.7	0.3
F	Failed	0.2	0.8
F	In Progress	0.7	0.3