Time-Dependent Planning for Buses, Trains, and Telescopes

Justin A. Boyan

Artificial Intelligence Laboratory
Massachusetts Institue Of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

Collaborators: Michael Littman (AT&T Labs), Leslie Kaelbling

The Problem: Imagine trying to plan a route from home to work that minimizes expected time. One approach is to use a tool such as "Mapquest", which annotates maps with information about estimated driving time, then applies a standard graph-search algorithm to produce a shortest route. Even if driving times are stochastic, the annotations can be expected times, so this presents no additional challenge. However, consider what happens if we would like to include public transportation in our route planning. Buses, trains, and subways vary in their expected travel time *according to the time of day*: buses and subways come more frequently during rush hour; trains leave on or close to scheduled departure times. In fact, even highway driving times vary with time of day, with heavier traffic and longer travel times during rush hour.

If we consider actions with time-dependent stochastic durations, we can no longer use a straightforward "shortest-path" approach to route planning. Instead, we must model the problem in an MDP whose states include both a discrete location component and a real-valued time component. Our work gives an approach to representing and solving such a time-dependent MDP, which we term a TMDP [1]. With appropriate restrictions, we guarantee that the optimal value function at each location is a piecewise linear function of time, which can be represented exactly and computed efficiently by value iteration.

Note that the TMDP model is more general than semi-Markov decision processes [3], which have no notion of absolute time. With absolute time in the state space, we can model a rich set of domain objectives beyond minimizing expected time, such as maximizing the probability of making a deadline, or maximizing the dollar reward of a path subject to a time deadline. In fact, using the time dimension to represent other one-dimensional quantities, our approach supports planning with non-linear utilities (e.g., risk-aversion), or with a continuous resource such as battery life or money.

Examples: Figure 1 illustrates an example TMDP for a commute from San Francisco to NASA Ames. The 14 discrete states model both location and observed traffic conditions: shaded and unshaded circles represent heavy and light traffic, respectively. Observed transition times and traffic conditions are stochastic, and depend on both the time and traffic conditions at the originating location. At locations 5, 6, 11, and 12, the "catch-the-train" action induces an arrival distribution, reflecting the train schedules.

The domain objective is to arrive at Ames by 9:00. We impose a linear penalty for arriving between 9 and noon, and an infinite penalty for arriving after noon. There are also linear penalties on the number of minutes spent driving in light traffic, driving in heavy traffic, and riding on the train; the coefficients of these penalties may be adjusted to reflect the commuter's tastes.

Figures 2 and 3 present the optimal value function and policy for location #10, "US101&Bayshore in heavy traffic". There are two actions from this state, corresponding to driving directly to Ames and driving to the train station to wait for the next train. Driving to the train station is preferred (has higher value) at times that are close—but not too close!—to the departure times of the train.

We have also applied the TMDP model to the problem of telescope observation scheduling [2]. The optimal observation plan must accommodate significant stochasticity in observation durations, while adhering to the absolute-time constraints on the hours of night during which the various targets are visible. The TMDP-based solver significantly outperformed a conventional solution method based on discretizing the time dimension.

The commuting example is solved in well under a second by our current implementation, but we note that it is manipulating around 700 linear segments in the final iteration. Our current research concerns efficient approaches to

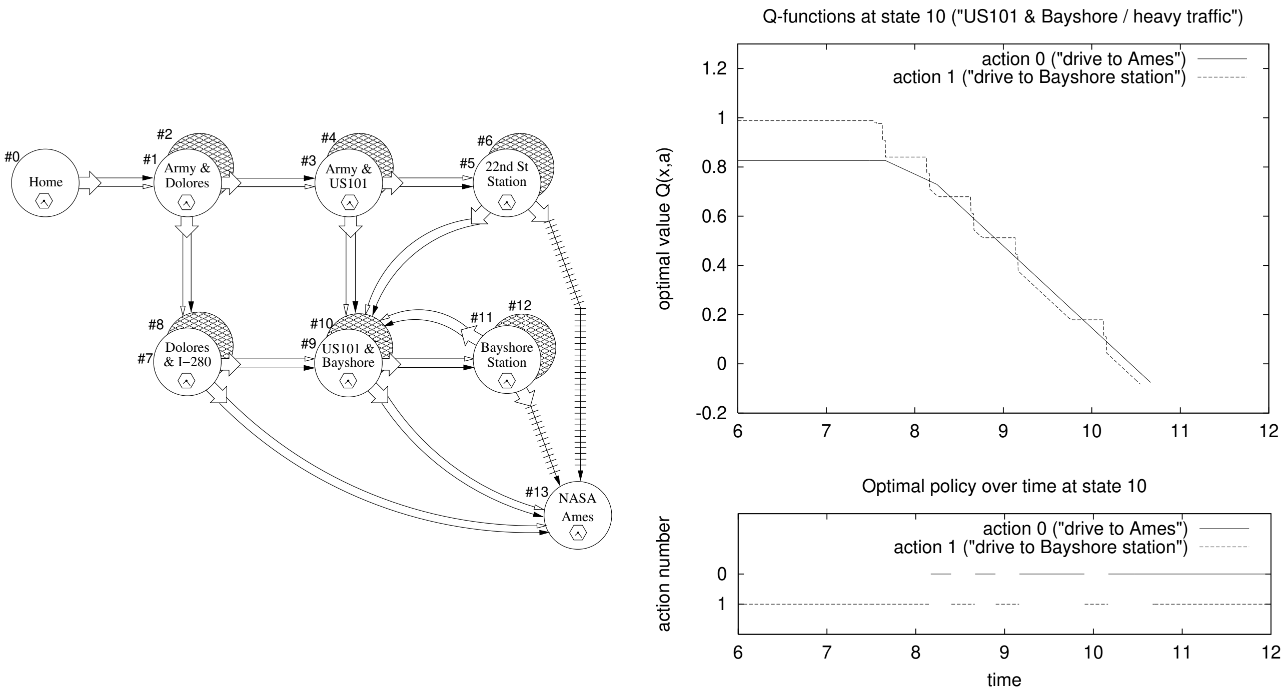

Figure 1: The San Francisco to Ames commuting example. At right, the optimal value function and policy at location 10.

propagating upper and lower bounds to the value function. This would allow the system to compute an optimal or provably near-optimal policy without necessarily identifying all the twists and turns in the optimal value function. We hope to scale our approach to much larger problems—on the scale of commuting throughout the entire area of Boston, or telescope scheduling over the whole night sky.

Research Support: NASA Ames Research Center.

References:

[1] J. A. Boyan and M. L. Littman. Exact solutions to time-dependent MDPs. In *Advances In Neural Information Processing Systems*. MIT Press, 2001. to appear.

[2] John Bresina, Mark Drummond, and Keith Swanson. Managing action duration uncertainty with just-in-case scheduling. In *Decision-Theoretic Planning: Papers from the 1994 Spring AAAI Symposium*, pages 19–26, Stanford, CA, 1994. AAAI Press, Menlo Park, California.

[3] Martin L. Puterman. *Markov Decision Processes—Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., New York, NY, 1994.