Autonomous Nondeterministic Tour Guides: Improving Quality of Experience with TTD-MDPs

Andrew S. Cantino, David L. Roberts, and Charles L. Isbell
College of Computing
Georgia Institute of Technology
{cantino,robertsd,isbell}@cc.gatech.edu

Georgia Institute of Technology Technical Report
GIT-IIC-07-02

Abstract

In this paper, we address the problem of building a system of autonomous tour guides for a complex environment, such as a museum with many visitors. Visitors may have varying preferences for types of art or may wish to visit different areas across multiple visits. Often, these goals conflict. For example, many visitors may wish to see the museum's most popular work, but that could cause congestion, ruining the experience. Thus, our task is to build a set of agents that can satisfy their visitors' goals while simultaneously providing quality experiences for all.

We use *targeted trajectory distribution* MDPs (TTD-MDPs), a technology developed to guide players in an interactive entertainment setting. The solution to a TTD-MDP is a probabilistic policy that results in a specific distribution of trajectories through a state space. We motivate TTD-MDPs for the museum tour problem, then describe the development of a number of models of museum visitors. Additionally, we propose a museum model and simulate tours using personalized TTD-MDP tour guides for each kind of visitor. We explain how the use of probabilistic policies reduces the congestion experienced by visitors while preserving their ability to pursue and realize goals.

1 Introduction

In this paper, we discuss the creation of a system of interactive tour guides that gently guide visitors through a set of engaging experiences in complex social environments. Specifically, we consider tour guides for visitors to a museum. Museums are an interesting test bed because of their size, complexity of layout, the number of simultaneous visitors, and the variety of goals these visitors may pursue. One of the world's most famous museums, the Louvre in Paris, contains tens of thousands of art works in hundreds of rooms and is visited by over seven million people annually. Visitors may have preferences for different types of art or art from different time periods. Additionally, there are very famous pieces of art such as Leonardo da Vinci's "Mona Lisa" and Alexandros of Antioch's "Venus de Milo" that many visitors will want to see. There is not enough time to see everything in the museum's collection during any given visit, so many guests may be repeat visitors trying to see things they have not seen before.

Thus, in building a system of agents to act as tour guides in such a setting, we are forced to balance many competing desires. We want to take groups (or single individuals) on tours that enable them to see as much art as possible without overwhelming them. We want the groups to see the specific pieces of art that they are interested in as well as focus on the type of art that they prefer, but limit congestion. Finally, we want to allow visitors to ignore the tour guide's directions, while still ensuring that they reap the benefit of the tour guide's insight.

We consider a scenario where each group of museum visitors is given a small handheld device, such as a PDA, that will interactively guide them through the museum by suggesting actions that they might take. With this small device, we have limited processing power, limited memory, and limited communication; therefore, we must consider methods to reduce the computational demands imposed on the tour guide.

We have opted to use *targeted trajectory distribution Markov decision processes* (TTD-MDPs) [19]. TTD-MDPs are a class of Markov decision processes originally developed for coordinating agents engaged in interactive entertainment [10]. A solution to a TTD-MDP is a probabilistic policy that induces a specific distribution over trajectories. By using a probabilistic policy, we can precompute an expected goal set for different types of visitors and guide them on tours accordingly. This allows us to avoid fully modeling every visitor's history of visits and preferences. We use TTD-MDPs because they provide a method to target a distribution over desirable museum tours that enables the trade-off between autonomy and exploitation mentioned above.

In the next section, we formally introduce TTD-MDPs. We then show how TTD-MDPs can be adapted from their original formulation for use in the tour guide application. Next, we describe our application domain in some detail, motivating the assumptions we make about state representation and visitor modeling. We then describe our experimental setup and results. We demonstrate that TTD-MDPs provide a viable solution for building a system of multiple autonomous tour guides. Finally, we conclude with a discussion of related work and future directions.

2 TTD-MDPs

An MDP is a tuple $< \mathcal{S}, \mathcal{A}, P, R >$, where $\mathcal{S}$ is a set of states, $\mathcal{A}$ is a set of actions, $P : \{\mathcal{S} \times \mathcal{A} \times \mathcal{S}\} \rightarrow [0, 1]$ is a transition function, and $R : \mathcal{S} \rightarrow \mathbf{R}$ is a reward function. The solution to an MDP is a policy $\pi : \mathcal{S} \rightarrow \mathcal{A}$. An optimal policy ensures that the agent receives the maximum long-term expected reward.

A TTD-MDP is also a tuple $< \mathcal{T}, \mathcal{A}, P, P(\mathcal{T}) >$, with states $\mathcal{T}$ that are finite-length trajectories of MDP states, possibly including the history of actions as well; a set of actions $\mathcal{A}$; a transition model P; and a target distribution over complete trajectories $P(\mathcal{T})$. The solution to a TTD-MDP is a policy $\pi : \mathcal{T} \rightarrow P(\mathcal{A})$ providing a distribution over actions in every state. The optimal policy results in long-term behavior as close to the target distribution as possible.

Any finite-length discrete-time MDP can be converted to a TTD-MDP. Consider an MDP with a set of states $\mathcal{S}$ and sets of actions available in each state $\mathcal{A}_s$. The probability $P_{i+1}(s)$ that the process is in state s at time $i + 1$ is defined recursively by:

$$P_{i+1}(s) = \sum_{\forall s' \in \mathcal{S}, a \in \mathcal{A}_{s'}} (P(s|a, s') \cdot P(a|s') \cdot P_i(s')) \quad (1)$$

where $P(s|a, s')$ is the transition model encoding the dynamics of the world and $P(a|s')$ is the policy under the agent's control. During an actual episode, $P_i(s') = 1$.

Because we are interested in trajectories in TTD-MDPs, we can simply roll the history of the MDP states into the TTD-MDP trajectories, resulting in a TTD-MDP where each trajectory represents a sequence of states in the underlying MDP, optionally including a history of the actions taken.

Dealing with trajectories means that the "state" space of the TTD-MDP forms a tree. Note that we can restate Equation 1:

$$P(t) = \sum_{\forall a \in \mathcal{A}_{t'}} (P(t|a, t') \cdot P(a|t')) \cdot P(t') \quad (2)$$

In other words, for every partial or full trajectory t, the transition probability $P(t|a, t')$ is nonzero for exactly one $t' \sqsubset t$ that is its prefix. Thus, the summation must only account for possible actions that can be taken in the prefix trajectory rather than actions in multiple MDP states. Further, each trajectory has a fixed length and can therefore appear at only one specific time.

When it is possible to build a policy exactly matching the target distribution, Algorithm 1—an online variant of the algorithm from [19]—will compute the optimal policy for every partial or complete trajectory. Unfortunately this is not always possible. There may be no vector $\vec{P}(a|t)$ that exactly satisfies the linear system in Step 7. Also, even when there is an exact solution, the elements of $\vec{P}(a|t)$ may not be realizable as probabilities. In that case, because the constraint forces the elements of the vector to sum to 1.0, at least one action a will have $P(a|t) < 0.0$ and at least one other action a' will have $P(a'|t) > 1.0$. Intuitively, achieving the desired distribution would require that action a be "undone" some percentage of the time. This is impossible, so in practice, we follow [19], zeroing out any negative values and re-normalizing.

In the following sections we describe the use of TTD-MDPs to guide simulated visitors through a modeled museum environment. The TTD-MDPs are solved online with Algorithm 1. At each step of the simulation, actions (modeled as suggestions to the visitors) are drawn probabilistically from $\vec{P}(a|t_i)$ and presented to the visitors for consideration. Visitors follow these suggestions according to their own specific preferences. The TTD-based guides are designed to provide visitors with a quality experience while minimizing congestion and allowing visitors the freedom to ignore suggestions.

Algorithm 1 Online Algorithm for solving TTD-MDPs

Require: A TTD-MDP tuple $< \mathcal{T}, \mathcal{A}, P, P(\mathcal{T}) >$

Ensure: The trajectory t resulting from this episode is drawn from the distribution $P(\mathcal{T})$.

$i \leftarrow 0$

Let t_i be the partial trajectory consisting of only the start state.

while t_i is not a complete trajectory **do**

  for Every child trajectory $t_{i:c}$ of trajectory t_i **do**

    Condition Equation 2 on t_i:

$$P(t_{i:c}|t_i) = \sum_{\forall a \in \mathcal{A}_{t_i}} \left(P(t_{i:c}|a, t_i) \cdot P(a|t_i)\right)$$

  end for

  This forms a system of $|\mathcal{T}_{t_{i:c}}|$ linear equations in $|\mathcal{A}_{t_i}|$ unknowns:

$$\vec{P}(t_{i:c}|t_i) = \vec{P}(t_{i:c}|a, t_i) \cdot \vec{P}(a|t_i)$$

  which can be solved for $\vec{P}(a|t_i)$ using standard linear algebra. This can optionally be memoized for future episodes.

  Draw an action a from $\vec{P}(a|t_i)$ and apply it. Let c_a be the outcome of applying the action.

  $t_{i+1} \leftarrow t_{i:c_a}$

  $i \leftarrow i + 1$

end while

---

3 Designing Tour Guides

In building a system of autonomous museum tour guides, we must consider several factors. In this section we discuss how to model museums, visitors, and tours. We aim to include enough information about the museum to enable a tour guide to compute tours tailored to the individual preferences of each visitor while maintaining enough simplicity so as to make the computation feasible. In particular, our approach must allow for fast action selection using only limited computation and communication.

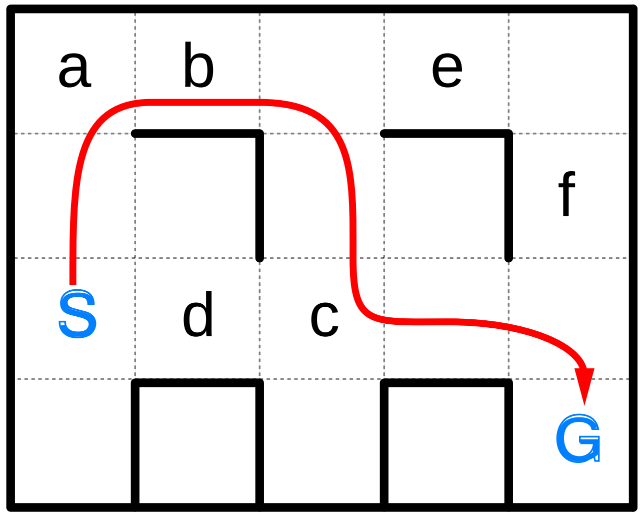

Figure 1: We model a museum as a grid world with walls that prevent some transitions. **S** is the start room of all trajectories (tours) through the museum, while **G** (gift shop) is the end room. The arrow from **S** to **G** shows one of the centroids from our Gaussian mixture model and represents a prototype tour, discussed in Section 3.2. The letters **a**, **b**, **c**, **d**, **e**, and **f** are goals that a museum visitor may have – they represent particularly famous or interesting works of art.

3.1 Modeling a Museum

As is shown in Figure 1, we model a museum as a 4x5 grid with walls preventing certain transitions and where some of the rooms contain objects of particular interest (like famous works of art). A trajectory through this grid world models a tour through a museum. Therefore, we consider trajectories to be sequences of rooms—in this case (x, y) coordinates in the grid. In Figure 1, **S** is the start room of all trajectories through the museum, while **G** (the gift shop) is the end room. We also model the visitor capacity of rooms in the museum. When above capacity, a room becomes congested. Thus, a tour is represented by a sequence of (x, y, c) coordinates that indicate the rooms visited and whether they were congested during the visit. For example, one tour might be $\{(0, 0, false), (0, 1, true), (1, 1, false), \ldots\}$.

We assume that through visitor input, RFID localization, or some other means, the agent can detect the current room. Further, we assume the agent can communicate with other guides or the museum itself to determine whether surrounding rooms are congested.

We represent the congested state of a room's neighbors as a *configuration*, $C = \{n_c, e_c, s_c, w_c\}$. We could consider configurations to be a part of the state space; however, there are two problems with this approach. First, it is unclear why one would want to construct a tour that depended directly upon how crowded neighboring rooms are. Second, such a scheme would require solving a linear system of 2^7 equations in Step 7 of the algorithm. Alternatively, we can treat configurations as observations and condition on them in the solution to the TTD-MDP. This allows us to deal with a system of at most four equations.[1] In this case, Equation 2 becomes:

$$P(t) = \sum_{\forall a \in \mathcal{A}_{t'}} \left(P(t|a, t', C_{t'}) \cdot P(a|t', C_{t'}) \right) \cdot P(t') \tag{3}$$

Therefore the system of linear equations to be solved in Step 7 of the algorithm becomes:

$$\vec{P}(t_{i:c}|t_i, C_{t_i}) = \vec{P}(t_{i:c}|a, t_i, C_{t_i}) \cdot \vec{P}(a|t_i, C_{t_i}) \tag{4}$$

3.2 Tour Probabilities

When using traditional MDPs, the designer achieves a desired behavior by selecting an appropriate reward signal. With TTD-MDPs, the designer achieves a desired behavior by properly selecting a target probability distribution over trajectories.

In the museum domain, we 1) define a distance metric between tours and 2) collect a set of prototypical "good" tours. Combining the distance metric with our prototype tours induces a target probability distribution over all possible tours.

We have chosen to base our distance metric on *Levenshtein distance* or *edit distance*. Edit distance measures the minimum number of insertions, deletions, or substitutions needed to transform one trajectory into another [11, 12]. The edit distance is a generalization of the Hamming distance [6] that is defined over strings of unequal length. It can be computed using an efficient $O(nm)$ dynamic programming method where n is the length of one trajectory and m is the length of the other.

We are concerned with two kinds of differences between trajectories: "room distance" and "congestion distance." We define $d_R(t, t')$ to be the edit distance between trajectories t and t' defined over the sequence of rooms (represented as their (x, y) coordinates). Similarly, we define $d_C(t, t')$ to be the edit distance defined over just the congestion indicators of the trajectories. For example, if two trajectories t and t' visit the same rooms in the same order, but t visits only uncongested rooms while t' visits three congested rooms, then $d_C(t, t') = 3$.

If $l(t)$ is the length of trajectory t and $\rho(t, x)$ is the prefix of length x of trajectory t where $\rho(t, x) = t$ when $l(t) < x$ we define a vector:

$$\vec{D}_\mu(t) = \begin{bmatrix} (1 + l(\mu) - l(t)) \cdot d_R(t, \rho(\mu, l(t))) \\ (1 + l(\mu) - l(t)) \cdot d_C(t, \rho(\mu, l(t))) \end{bmatrix} \tag{5}$$

that provides a measure of the difference between two trajectories. $\vec{D}_\mu(t)$ has two desirable properties. First, when the trajectories are of equal length, the value in each dimension is exactly the edit distance. Second, as the difference in trajectory length increases (or decreases), the value in each dimension increases (or decreases) even if the edit distance remains the same.

[1] There are at most four directions available from each room. Because we are conditioning on the *current* configuration of surrounding rooms, we assume that whether the neighboring rooms are congested will not change immediately, leaving only at most four possible next states. As a practical matter, this assumption holds because of the asynchronous nature of movement between rooms.

It is worth noting that this model provides a probability distribution over distances among trajectories rather than a probability distribution over the trajectories themselves. To account for this, we simply work with conditional probabilities as described in Equation 4 by normalizing the probability of every trajectory subsequent to the one under consideration.

To construct a distribution, we define a Gaussian mixture model over the set of prototypical tours, μ_i. We consider μ_i to be the centroid of a multivariate Gaussian distribution with covariance matrix Σ_i. Then, the probability of any trajectory t is given by

$$P(t) = \sum_{i=1}^{N} \tilde{P}(\mu_i) \cdot \mathcal{N}(t; \mu_i, \Sigma_i) \tag{6}$$

where

$$\mathcal{N}(t; \mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma|}} \exp^{-\frac{1}{2}(\vec{D}_\mu(t)^T \Sigma^{-1} \vec{D}_\mu(t))} \tag{7}$$

$\tilde{P}(\mu_i)$ is the prior weight given to each centroid, expressed as a probability, and $|\Sigma|$ is the determinant of Σ. Σ is constructed to reflect whatever tradeoff we would like to make between visiting rooms in a particular order and avoiding congestion.

4 Modeling Visitors

We assume that different types of visitors have different goals for their visits to the museum. There are analogies in other applications of TTD-MDPs. For example, in the drama management domain where TTD-MDPs were originally developed, one might expect different types of game players: there is the player who is trying hard to win the game, the player who is trying to explore the game world, and the player who is trying hard to "break" the game. We want the experience for each of these types of players to be a good one just as we want the experience for each of the types of museum visitors to be a good one.

In Section 4.1 we describe and motivate four museum visitor models. Then, in Section 4.2, we describe our visitor transition models and show how they reflect visitors' goals.

4.1 Visitor Types

We want to account for both naive and informed visitors. The naive visitor is modeled after a tourist who does not have a particular preference for any of the museum's exhibits other than what they may have read in a guide book. The informed visitor represents a dedicated art spectator.

In our domain, we explicitly model the destination goals that each class of visitors has. These goals represent works of art of particular interest to a visitor. The union of the sets of goals of all informed visitors is a strict super set of the union of the goal sets of all naive visitors. In addition, we consider two variants of these visitor types (for a total of four visitor models). These variants are the first-time visitor and the returning visitor. We model first-time visitors as having no history of satisfied goals, while returning visitors have some percentage of the possible goals satisfied already (we use 35% in our experiments). In the 4x5 museum world described in Section 3.1, we select 10 out of the 20 rooms to contain potential goals for the informed visitor and 6 to contain potential goals for the new visitor. Figure 1 shows the 6 possible goals for new visitors. For each of the visitor types, we assign 3 goals to be "hidden" goals, or goals that the visitor will enjoy but does not know to pursue. The gift shop (**G** in Figure 1) is also added as a possible goal for all visitors.

4.2 Visitors' Transition Models

The TTD-MDP tour guides lead visitors by suggesting actions for them to take according to $\vec{P}(a|t_i)$. The available actions are selected from the set $\{n, s, e, w, noop\}$. The $noop$ action means the guide makes no suggestion.

We construct a transition model where visitors usually move toward a goal location when they are close to it, regardless of the tour guide's suggestions. When not near a goal location, visitors are more likely to follow the guide's suggestions. Further, visitors prefer not to revisit rooms whenever possible.

We divide visitors' willingness to follow suggestions into three categories: those who *possibly*, *probably*, or *definitely* will follow tour guide suggestions. These categories of visitors are modeled by variations in the visitors' transition probabilities.

During tours, we obtain actual transition probabilities through sampling. When in a room, we simulate a random population of visitors consistent with the current visitor's model. We obtain a transition matrix by querying each random visitor for her response to each action that the tour guide may take. This has the desired effect of scaling the local transition probabilities by the relative probability of goals (*e.g.* if there are potential goals in neighboring rooms, the transition probabilities will be skewed in those directions proportionally to the probability that a randomly sampled population of visitors will want to pursue those goals.)

5 Experimental Design

We perform a number of experiments to characterize the efficacy of our approach for building autonomous tour guides. We explore two general measures of performance. First, we want to know how closely we can match the desired distribution over trajectories through the museum. Second, we want a measure of how "satisfied" visitors are. We compare results for our TTD-MDP tour guide to three other approaches. The first two use no tour guide: 1) *wander*, a simple, randomly wandering visitor, and 2) *ignore*, an otherwise wandering visitor who pursues a goal when one step away. The third approach, *random*, augments *ignore* with a tour guide that chooses actions uniformly.

5.1 Setup

We select a set of prototype trajectories for both the naive and the informed visitors. These represent "good" tours, perhaps created by a museum curator. Naive visitors have two prototype trajectories while informed visitors have three. We assign the same set of prototype trajectories for both the new and returning visitors of each type. The trajectories are chosen subject to two conditions: 1) a prototype trajectory must begin in the entrance room and end in the gift shop; and, 2) every possible goal must lie on at least one prototype trajectory. Condition (2) is motivated by the desire not to have the tour guide be unfairly penalized for not guiding a visitor to a goal if that goal is not available on some prototype trajectory. We expect museum curators will be able to articulate sets of prototype trajectories that fully cover the space of potentially interesting artworks (or visitor goals).

We select an initial uniform distribution over visitor types. Visitors enter the museum model at a constant rate of n per simulation step. We select a room capacity to reflect the average number of visitors that we expect to be in the museum at any given time step. If this is set too high, then none of the rooms will ever be congested. On the other hand, if it is too low, then all of the rooms will be congested and none of the realized tours will reflect the prototype tours closely. Below we present results that empirically verify this fact.

During each simulation step, we select a random ordering over all visitors currently in the museum and allow them to move in this order. Before a visitor is allowed to move, we update the congestion state of all rooms to reflect any changes in configuration. We do this because, in reality, visitors do not synchronously move from room to room at the same time. Once every visitor in the museum has had a chance to move, we advance the simulation step and repeat the process.

5.2 Success Metrics

We wish to measure both how closely we match the target distribution over tours and how many visitors satisfied their goals.

5.2.1 TTD Performance

To characterize the first type of performance, we look at aggregate statistics on the distribution of trajectories. We also compute a measure of policy error. Specifically, we consider the L_2-norm of the desired policy at every step with the obtained dynamics. Recall the system of linear equations from Step 7 of Algorithm 1: $\vec{P} = T \cdot \vec{\pi}$. We are using $\|\vec{P} - T \cdot \vec{\pi}\|_2$ as our error metric. We report these measures as an average over all local computations made during an evaluation trial.

Additionally, we are interested in looking at the behavior in terms of individual visitors. That is, we want to characterize the distribution over realized trajectories. To accomplish this, we look at the distribution of distances of each tour obtained during evaluation from the centroid it is closest to. To be more precise, we create histogram bins for each distance $0, \ldots, l$ (where l is the maximum permissible trajectory length). Then, for each tour encountered, we select the centroid that it is closest to and increment the histogram bin associated with that distance.

5.2.2 Goal Realization

To measure how effective our tour guides are at realizing the goals of visitors, we consider three summary statistics: first, the percent of a visitor's known goals that are achieved; second, the percent of a visitor's hidden goals that are achieved; third, the frequency of congested rooms experienced by each visitor. We compute these measures in aggregate, but also account for visitor type and visitor responsiveness to tour guide suggestions. This enables us to illustrate how our system reacts in different situations.

6 Results

Here, we summarize the results obtained for a number of experiments. We present data to illustrate the effects of using a TTD-based tour guide on congestion and visitors' realization of goals. Additionally, we highlight the complex tradeoff between autonomy of visitors and the resulting quality of experience.

For the experiments we present below, we assume that the visitors has limited time. Specifically, they took tours of no more than 10 steps. If the visitor had not reached the gift shop within 10 steps he immediately moved there.[2]

6.1 Characterizing Tours

We remind the reader that in our model, prototype tours represent a hypothetical museum curator's view of what makes a good tour. Thus, it makes sense to examine how closely visitors have followed those prototypes. In Figure 2, we plot an "edit distance histogram" for the informed visitor (new and returning) both with and without the benefit of a TTD-based tour guide. The data for this plot was obtained from experiments run with a low goal density, a room capacity of four visitors (beyond which the room becomes congested), five visitors added to the museum per simulation time step, and visitors with a fairly low probability of accepting tour guide suggestions (the *possibly* visitor category). In the low density case, visitors choose from only half of the possible goals available in the high density case. Notice the relative shape of the distribution of distances for the trajectories obtained using the TTD-based tour guides (*i.e.* a Gaussian that has been cut in half). This illustrates nicely that despite the relative lack of cooperativeness of this visitor type, we still see a distribution over distance that roughly matches the shape we desire and expect from our mixture of Gaussians model. The data for the informed visitor without the tour guide does not exhibit this behavior. The dips at distance three and six in this plot are attributable to the structure of the museum and the set of prototypes. Specifically, once the visitor enters a particular part of the

[2] We have also run experiments without imposing limits on the length of visitor trajectories. In those experiments, our prototype tours were still of length 10. We found that this does not significantly change our results when visitors reach the goal in more than 10 steps; however, visitors tend to meet more goals in this situation because those that do not stay on our prototype trajectories may wander randomly around the museum, realizing unmet goals along the way. On the other hand, once a visitor has significantly deviated from all prototype tours by visiting many more than 10 rooms, the tour guide's suggestions cease to be meaningful.

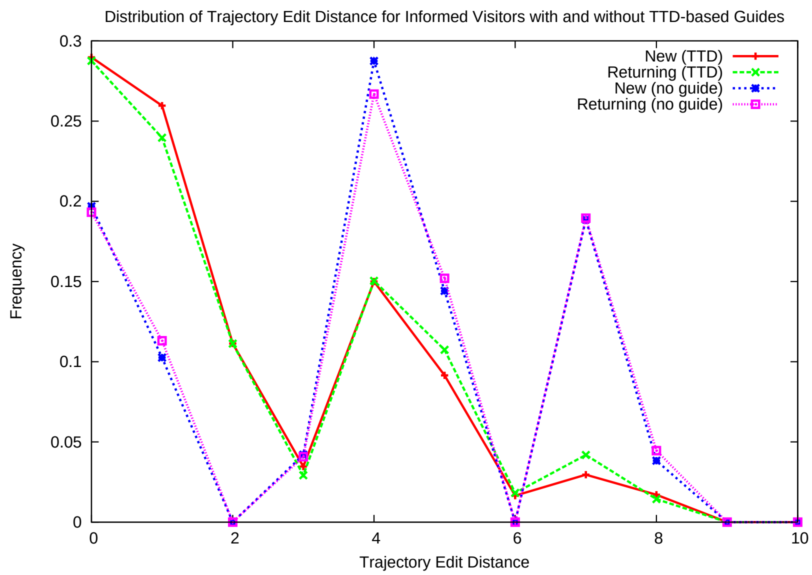

Figure 2: Distribution of Trajectory Edit Distance for Informed Visitors with and without TTD-based Guides.

space (*e.g.* the top left or top right corner in Figure 1) there are a limited number of locations from which they can diverge to another path, thus making deviation less likely in these regions.

In Figures 3 & 4 we examine the frequency of congested rooms. In Figure 3, we compare the congestion rates experienced by the naive visitor (new and returning) in trials both with and without the benefit of the TTD-based guide. There are two interesting points here. First, the rate of congestion is almost identical for the new and returning visitors in each case. Second, note the relative position of the curves for the trials with and without the guides. Visitors with guides experienced less congestion, with a histogram peak at 0 congested rooms, instead of 2 for visitors without a guide.

Thus far, we have highlighted the relationship between visitors with no tour guide and the most unwilling visitors with a guide. Consider Figure 4, where this unwilling visitor is compared to more willing variants. Here, we see that all visitors exhibit the "half-Gaussian" shape noted previously, but the curves for the visitors who listen to their guides have lower variance than the curves of those who do not. Thus, those who listen to their guides trend toward experiencing less congestion. Furthermore, in general we see the desired changes to the shape of the half-Gaussian in response to varying parameters.

6.2 Goals

In Table 1, we consider the results of experiments with and without TTD-based guides as well as with *wander* and *random* (see Section 5). In this table, as in the subsequent

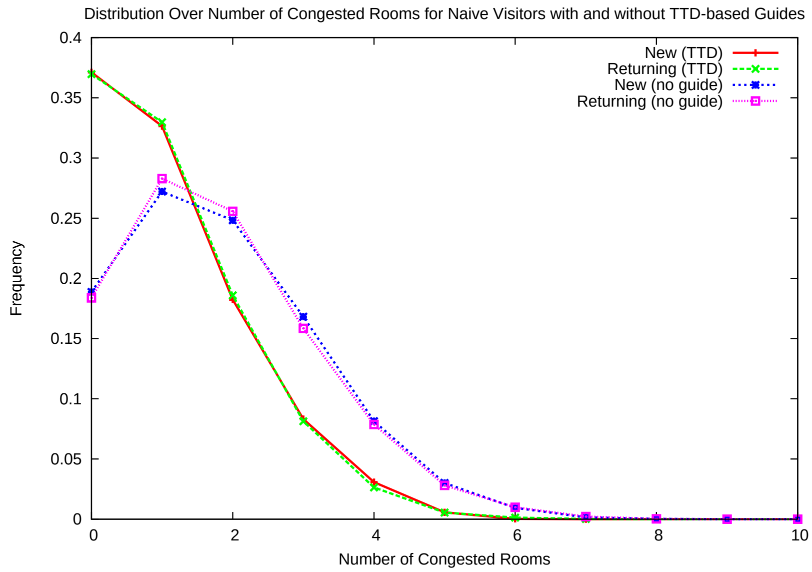

Figure 3: Frequency of Congestion for Naive Visitors with and without TTD-based Guides.

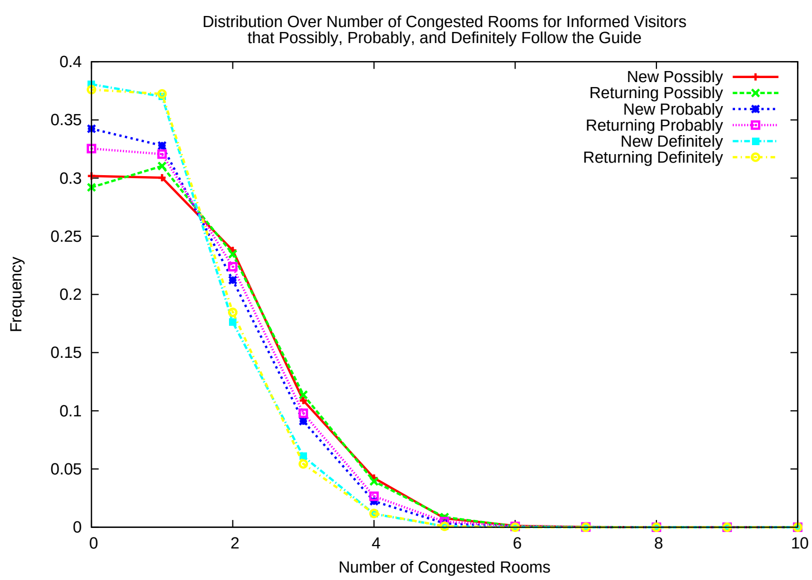

Figure 4: Frequency of Congestion for Informed Visitors with Varying Willingness to Follow the Guide's Suggestions.

Measure	Congestion		New Goals		Hidden Goals	
Model	L	H	L	H	L	H
TTD:	**0.135**	**0.153**	**0.476**	**0.598**	0.289	0.351
ignore:	**0.209**	**0.202**	**0.497**	**0.608**	0.290	0.374
wander:	0.517	0.517	0.113	0.271	0.118	0.273
random:	0.287	0.247	0.398	0.554	0.226	0.342

Table 1: Aggregate statistics for visitor models with low and high goal density.

Measure	Congestion		New Goals		Hidden Goals	
Capacity	L	H	L	H	L	H
inf	0	0	0.501	0.569	0.313	0.346
6	0.021	0.029	0.487	0.564	0.307	0.344
5	0.049	0.062	0.472	0.560	0.297	0.342
4	0.116	0.133	0.441	0.543	0.272	0.338
3	0.244	0.259	0.416	0.534	0.253	0.333

Table 2: Aggregate Statistics for Visitors that *probably* follow guide instructions with varying room capacity limits.

two tables, we use "L" and "H" to represent low and high goal density trials. The results in these tables are shown averaged across all visitor models (naive and informed in both the new and returning variants). The *wander* and *random* baselines do not perform well in any of the categories. In the case of the *wander* baseline, this is attributable to a lack of goal directed behavior. For the *random* tour guide, however, this is more attributable to the willingness of the visitor to follow the guides random suggestions. In comparison to those baselines, the *ignore* and TTD cases yield very promising results. Specifically, we see a noticeable reduction in congestion as well as a significant increase in goal realization that is even more pronounced when the TTD guides are used.

6.3 Capacity and Visitor Autonomy

In Table 2, we summarize the effects of room capacity. In particular, we see that the effects of room capacity on goal realization are more pronounced in the low goal density case than in the high goal density case. Particularly, as capacity decreases, the percentage of realized goals in the high density case remains essentially the same. The

Measure	Tour Length		Congestion		New Goals		Hidden Goals		Policy Error	
Model	L	H	L	H	L	H	L	H	L	H
ignore	9.883	9.861	**0.209**	**0.202**	0.497	0.608	0.290	0.374	n/a	n/a
possibly	9.968	9.960	**0.135**	**0.153**	0.476	0.598	0.289	0.351	**0.206**	**0.271**
probably	9.981	9.976	**0.116**	**0.133**	0.441	0.544	**0.274**	**0.338**	**0.141**	**0.177**
defi nitely	9.993	9.990	**0.091**	**0.090**	0.364	0.450	**0.315**	**0.385**	**0.071**	**0.066**

Table 3: Aggregate Statistics for Visitors with varying willingness to follow suggestions.

effect in the low goal density case is exaggerated because the same number of visitors are sharing a desire to achieve fewer goals, thus increasing congestion. As a result of the guide's tendency to suggest alternates to congested rooms and the visitor's tendency to follow those suggestions, we also see a reduction in goal satisfaction.

Consider the effect of willingness to respond to guide suggestions in concert with the data in Table 3. The data in the table was obtained by varying the visitor's willingness to follow advice in both high and low goal density scenarios. The detailed information presented in Figure 4 is restated in aggregate in the "Congestion" column of the table. Note how the rate of congestion is slightly lower for the low goal density case. This is attributable to the visitors not gravitating toward as many rooms. This information taken together with the percentage of satisfied goals is very interesting. We see that the more willing a visitor is to follow the tour guide, the less congestion they will encounter, but the fewer goals they will realize; however, this tradeoff may be worthwhile—a 26.0% reduction in goal satisfaction accompanies a 55.4% reduction in congestion (in the high goal density case).

Additionally, we see that although the frequency of realization for hidden goals generally decreases as visitors more willingly follow their guides, if they always follow their guides they begin to realize more goals again. This occurs because the guides have some sense of where hidden goals may be, due to the museum curator's well-constructed centroid tours.

In Table 3, we also present policy error. Here, we report the L_2 error (see Section 5.2) averaged over each of the local computations made during a simulation run. What this measures is the difference between the dynamics that the policy will yield and the ideal dynamics. Here, "dynamics" refers to the probability of seeing one room given that you are in another room. In this context, L_2 measures the distance between what we want visitors to do and what we can actually get them to do (subject to the accuracy of the visitor model). In these experiments, we see a reduction in the average local error as visitors more willingly follow requests. Additionally, we see lower local policy error in the low goal density case. As before, this is attributable to the fact that there are fewer goals and therefore fewer rooms that are neighbors of rooms with goals, which results in fewer instances of visitors ignoring the guide's suggestions to pursue those goals.

As visitors have more autonomy, they achieve more of their goals because of their willingness to ignore the tour guide and pursue a known goal; however, this gives rise to a tragedy of the commons: when visitors always act only in their own immediate interest, they end up in crowded parts of the museum lessening the quality of the experience for everyone. On the other hand, if visitors always listen to the tour guide, we find that they experience less congestion at the expense of realizing fewer of their known goals. Somewhere in the middle of these extremes is a "sweet spot" where visitors exercise enough autonomy to implicitly express desires but listen enough to take advantage of the tour guide.

7 Related Work

Here we describe work related to both TTD-MDPs and tour guides. Much of the work related to TTD-MDPs can be grouped into two categories: drama management and probabilistic polices for MDPs. Work on tour guides is based mainly in the robotics and ubiquitous computing communities. As we will shall see, the technical issues that arise in those communities are orthogonal to ours.

7.1 TTD-MDP Related Work

Using a drama manager to guide interactive entertainment was first proposed in 1986 by Laurel [10], formalizing the idea of an agent directing the action in response to visitor's actions. The inspiration for TTD-MDPs was based on a particular formalism for drama management proposed by Bates [1]. It was later formulated as a search problem by Weyhrauch [21] based on an expecti-max game tree like search over plot point sequences and then reformulated as a reinforcement learning problem by Nelson *et.al.* [15]. That work led directly to the development of TTD-MDPs to include the capability for controlling variety of experience.

TTD-MDPs share common ground with work on non-deterministic policies such as Isbell *et. al.*'s work on Cobot [7] and Littman's work on Markov games [13]. Additionally, the idea of using trajectories to solve for a policy is closely related to the work of Kearns, Mansour, & Ng [8] where sampled trees of trajectories are used to estimate state sequences and transition probabilities in a partially observable environment.

7.2 Robotic Tour Guides

As robotic technology has become increasingly accessible, researchers have begun to focus on robot-human interaction in social environments. In particular, one line of research involves the creation of sophisticated robotic tour guides that greet visitors, entertain them with antics or conversation, and lead them to their destination (see Kim *et. al.* [9], for example). In work on robotic tour guides, however, the specific tours given are fixed ahead of time and the autonomy of the tour taker is not considered.

In contrast to our work, the technical issues involved there are in navigation, localization, and speech recognition. Our TTD-MDP based tour guides are carried by museum visitors and therefore do not need the level of situational awareness that self-navigating robots do. The few citations we provide here are intended as an overview of technical areas under investigation by the robotics community rather than an in depth look at techniques used for robotic tour guides. For example, Prodana & Drygajlo describe a system using Bayesian Networks for interpreting multimodal signals [17]. In this system, they attempt to fuse and interpret signals from a laser scanner, cameras, and microphones in order to give the robot an understanding of its environment. Similarly, much work has gone into the development of voice enabled interfaces for robotic tour guides [5, 18]. Other work examines the problem of determining absolute position from sensors [20].

7.3 Ubiquitous Computing

Some work exists in the ubiquitous computing community on constructing mobile tour guides. Early efforts in this space relied on wireless networking and central repositories of data [14, 16]. Subsequent research has focused almost exclusively on content delivery. Specifically, an interest exists in context aware applications that can sense location either through use of localization technologies like GPS or by explicit interaction with the user [4].

Lastly, some additional research exists that explores the transition of these technologies to handheld devices [3, 2] and the testing of these systems in real world situations. As with the work from the robotics community discussed above, we are unaware of any existing research into the endowment of these tour guides with autonomy or decision making capabilities.

8 Future Work

The approach we have described here is not limited to museum tours. For example, we wish to explore the application of this technology to dynamic web site restructuring. If we consider a sequence of page views to be a trajectory through an MDP, we can select conversion goals and construct a distribution over desired navigation paths. By taking actions to make link positions more or less prominent, we can influence the navigation path that visitors are likely to take.

Additionally, we would like to apply the approach in this paper to models of congestion in highway transportation. The same sort of tradeoffs that exist in our museum model between visitors' desire to achieve goals and the museum curator's desire to avoid congestion exist in transportation systems. In these systems, the individual's desire to reach a destination can conflict with the group's desire to minimize traffic congestion. Imagine cars with navigation systems that direct users on "good" routes to their destinations while trading off a small amount of driving time for a large overall drop in system congestion.

We also intend to further explore the tradeoff between congestion and the preservation of a desired trajectory distribution. The covariance matrix Σ_i from our Gaussian mixture model allows us to vary how strongly our prototype trajectories are preferred over similar trajectories. Additionally, it allows us to tradeoff between commitment to the prototype trajectories and avoidance of congestion. Initial experiments are promising.

Lastly, when an agent is tasked with interaction in a social environment, it must be able to reason about its surroundings and interaction partners. When this interaction is specifically targeted to a human, this task can rapidly become more difficult—especially given that models of human behavior are not one-size-fits-all. In this work, we have utilized visitor models that are based on equivalence classes of behaviors. When using these models in a real world application, there are two important questions that need to be answered. First, which of these equivalence classes is most appropriate for the current interaction? Second, how is the selected model class tuned specifically for the current interaction? In answering these two questions, we find it necessary to consider online model adaption.

9 Conclusion

In this paper, we consider building automated tour guides for visitors to a museum. We attempt to provide autonomy to visitors while simultaneously increasing the quality of their experience. We construct our tour guides using TTD-MDPs, where each tour defines a trajectory that is represented by a sequence of rooms through the museum. The guides act by sometimes suggesting movements to the visitors. We define the TTD-MDP target distribution via a Gaussian mixture model over distances from hand-crafted prototype trajectories. We derive a distance function based on Levenshtein edit distance. Using an online variant of the existing TTD-MDP solution technique, we obtain a probabilistic policy over actions and apply it to simulated visitors in a simulated museum. We tailor suggestions to a variety of visitor models that capture variability knowledge, experience and likelihood of listening to suggestions.

Our autonomous tour guides dynamically construct tours online in response to visitor's reactions to their suggestions. We are able to show that when visitors cooperate, even only occasionally, visitors achieve about as many goals as visitors who never cooperate while significantly reducing overall museum congestion.

We also find that our tour guides perform best when visitors occasionally choose to ignore their suggestions. This implicit communication from visitors gives the tour guides feedback that can be incorporated into the decision making process while respecting visitor autonomy.

Acknowledgments

This research was performed while on appointment as a U.S. Department of Homeland Security (DHS) Fellow under the DHS Scholarship and Fellowship Program, a program administered by the Oak Ridge Institute for Science and Education (ORISE) for DHS through an interagency agreement with the U.S Department of Energy (DOE). ORISE is managed by Oak Ridge Associated Universities under DOE contract number DE-AC05-06OR23100. All opinions expressed in this paper are the author's and do not necessarily reflect the policies and views of DHS, DOE, or ORISE. We also acknowledge the support of DARPA under contract No. HR0011-06-1-0021.

References

[1] J. Bates. Virtual reality, art, and entertainment. *Presence: The Journal of Teleoperators and Virtual Environments*, 2(1):133–138, 1992.

[2] J. Broadbent and P. Marti. Location aware mobile interactive guides: usability issues. In *Proceedings of the Fourth International Conference on Hypermedia and Interactivity in Museums (ICHIM97)*, Paris, 1997.

[3] K. Cheverst, N. Davies, K. Mitchell, A. Friday, and C. Efstratiou. Developing a context-aware electronic tourist guide: some issues and experiences. In *CHI*, pages 17–24, 2000.

[4] R. Cox, M. O'Donnell, and J. Oberlander. Dynamic versus static hypermedia in museum education: an evaluation of ilex, the intelligent labelling explorer. In *Proceedings of the Artificial Intelligence in Education conference*, Le Mans, 1999.

[5] A. Drygajlo, P. Prodanov, G. Ramel, M. Meisser, and R. Siegwart. On developing voice enabled interface for interactive tour-guide robots. *Journal of Advanced Robotics*, 2003.

[6] R. W. Hamming. Error-detecting and error-correcting codes. *Bell System Technical Journal*, 29(2):147–160, 1950.

[7] C. L. Isbell, Jr., C. R. Shelton, M. Kearns, S. Singh, and P. Stone. A social reinforcement learning agent. In *Proceedings of the Fifth International Conference on Autonomous Agents (Agents-01)*, pages 377–384, 2001.

[8] M. Kearns, Y. Mansour, and A. Y. Ng. Approximate planning in large POMDPs via reusable trajectories. *Advances in Neural Information Processing Systems*, 12, 2000.

[9] G. Kim, W. Chung, K.-R. Kim, M. Kim, S. Han, and R. H. Shinn. The autonomous tour-guide robot jinny. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 4, pages 3450–3455, 2004.

[10] B. Laurel. *Toward the Design of a Computer-Based Interactive Fantasy System*. PhD thesis, Drama department, Ohio State University, 1986.

[11] V. I. Levenshtein. Binary codes capable of correcting deletions, insertions and reversals. *Doklady Akademii Nauk SSSR*, 163(4):845–848, 1965.

[12] V. I. Levenshtein. Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady*, 10(8):707–710, 1966.

[13] M. L. Littman. Markov games as a framework for multi-agent reinforcement learning. In *Proceedings of the Eleventh International Conference on Machine Learning (ICML-94)*, pages 157–163, 1994.

[14] S. Long, R. Kooper, G. D. Abowd, and C. G. Atkeson. Rapid prototyping of mobile context-aware applications: The cyberguide case study. In *Proceedings of 2nd ACM International Conference on Mobile Computing*, Rye NY, 1996.

[15] M. J. Nelson, D. L. Roberts, C. L. Isbell, and M. Mateas. Reinforcement learning for declarative optimization-based drama management. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-06)*, 2006.

[16] M. D. Pinkerton. Ubiquitous computing: Extending access to mobile data, 1997.

[17] P. Prodanov and A. Drygajlo. Multimodal interaction management for tour-guide robots using bayesian networks. In *Proceedings of the International Conference on Intelligent Robots and Systems*, 2003.

[18] P. Prodanov and A. Drygajlo. Decision networks for repair strategies in speech-based interaction with mobile tour-guide robots. In *Proceedings of the IEEE International Conference on Robotics and Automation*, 2005.

[19] D. L. Roberts, M. J. Nelson, C. L. Isbell, M. Mateas, and M. L. Littman. Targetting specific distributions of trajecotries in MDPs. In *Proceedings of the Twenty-First National Conference on Artificial Intelligence (AAAI-06)*, Boston, MA, 2006.

[20] R. Thrapp, C. Westbrook, and D. Subramanian. Robust localization algorithms for an autonomous campus tour guide. In *Proceedings of the IEEE International Conference on Robotics and Automation*, volume 2, pages 2065–2071, 2001.

[21] P. Weyhrauch. *Guiding Interactive Drama*. PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 1997. Technical Report CMU-CS-97-109.