

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

Andrew S. Cantino, David L. Roberts, Charles L. Isbell
College of Computing
Georgia Institute of Technology
Atlanta, GA

{cantino, robertsd, isbell}@cc.gatech.edu

ABSTRACT

In this paper, we address the problem of building a system of autonomous agents for a complex environment, in our case, a museum with many visitors. Visitors may have varying preferences for types of art or may wish to visit different exhibits on 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 high quality experiences for all.

Categories and Subject Descriptors

I.2.11 [Artificial Intelligence]: Distributed Artificial Intelligence—*Intelligent agents*; G.3 [Probability and Statistics]: Markov processes

General Terms

Algorithms, Measurement, Design

Keywords

Markov decision processes, interactive entertainment

1. INTRODUCTION

In this paper, we discuss the creation of a system of interactive agents that gently guide visitors through 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. Generally, there is insufficient time to see the whole museum during any given visit, so many guests may be repeat visitors who are trying to see previously unseen collections.

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. We describe a system that satisfies visitor goals while avoiding congestion and thus preserving the quality of experience for all.

We use *Targeted Trajectory Distribution Markov decision processes* (TTD-MDPs) [8]. TTD-MDPs are a class of Markov decision processes originally developed for coordinating agents engaged in interactive entertainment [4].

2. TTD-MDPS

A TTD-MDP is a tuple $\langle \mathcal{T}, \mathcal{A}, P, P(\mathcal{T}) \rangle$, with states $\mathcal{T}$ that are finite-length trajectories of MDP states (possibly including a 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 MDP can be converted into a TTD-MDP. 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. The trajectory space of the TTD-MDP forms a tree. We state the TTD-MDP equation as:

$$P(t) = \sum_{\forall a \in \mathcal{A}_{t'}} (P(t|a, t') \cdot P(a|t')) \cdot P(t') \quad (1)$$

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. For a complete introduction to TTD-MDPs, see [8].

3. DESIGNING TOUR GUIDES

3.1 Modeling a Museum

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) that represent potential goals of museum visitors. A trajectory through this grid world models a tour through a museum. Therefore, we consider trajectories to be sequences of rooms. There is an entrance where all trajectories through the museum begin and a gift shop where all trajectories end. We also model the visitor capacity of each room in the museum. When above capacity, a room becomes congested—decreasing visitors' enjoyment. Thus, a tour is represented by a sequence of (x, y, c) coordinates that indicate the rooms visited and

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

AAMAS'07 May 14–18 2007, Honolulu, Hawai'i, USA.
Copyright 2007 IFAAMAS.

whether they were congested during the visit. For example, one tour might be $\{(0, 0, false), (0, 1, true), (1, 1, false), \dots\}$.

We assume that through visitor input, RFID localization, or some other means, the tour guide agent can detect its current room. Further, we assume the agent can communicate locally with other agents to determine the congested state of the *closest* neighboring rooms.

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) pick a set of prototypical “good” tours. Combining the distance metric with these prototype induces a target probability distribution over all tours. In our case, we define a Gaussian mixture model over the set of prototypical tours with a distance metric based on *Levenshtein distance* or *edit distance* [5, 6]. Because this model is well defined for all possible trajectories, the tour guide can always make an intelligent decision even if the visitor has wandered far from the set of prototype tours.

4. MODELING VISITORS

We assume that different types of museum visitors have different goals, or known artworks that they wish to see. Therefore, we explicitly model locations in the museum as goals and model the visitors’ transitions as preferring to visit those goals. There are many types of visitors to museums, and they likely have varying sets of goals and varying willingness to follow the suggestions of tour guides. Therefore, we model visitors in three dimensions: 1) first time or returning visitor, 2) how much they know about art, and 3) willingness to follow tour guide suggestions.

These three dimensions yield numerous instantiations of visitor models. Specifically, we have a *naive visitor* type intended to represent a tourist that has little knowledge of art and knows mainly what they might have read in a popular guide book, and an *informed visitor* type that knows more about what the museum has to offer and has a larger number of goals. Additionally, we have a *new visitor* variant that has no completed goals when entering the museum, and a *returning visitor* variant that has already realized some percentage of the possible goals available to them. Lastly, we have four “levels” of willingness to follow guide suggestions. First, we have *ignore*, where the visitor does not follow tour guide suggestions. Second, we have *possibly* and *probably* visitors that follow the tour guide suggestions a given percentage of the time. These are likely more accurate models of visitor behavior. Lastly, we consider a visitor variant that will *definitely* follow tour guide suggestions, regardless of its own goals.

We classify visitors into new vs. returning and naive vs. informed before they start a tour, and we use TTD-MDPs specially trained on the appropriate combinations. The probabilistic policy of a TTD-MDP provides a good tour for each visitor that we hope will minimize both the number of goals that are not satisfied and the parts of the museum that are repeated or congested. Training on general classes of visitors allows us to avoid fully modeling every visitor’s history of visits and preferences, which provides a notable computational advantage.

5. 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 closeness to targeted tours. For

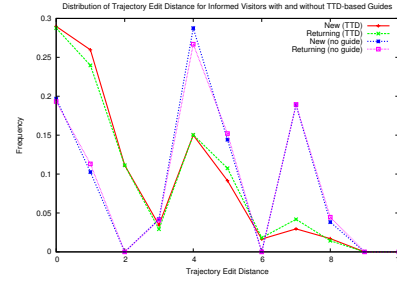


Figure 1: Distribution of Trajectory Edit Distance for Informed Visitors with and without TTD-based Guides.

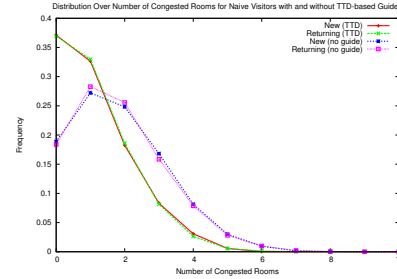


Figure 2: Frequency of Congestion for Naive Visitors with and without TTD-based Guides.

the experiments we present below, we assume that the visitors have limited time. Specifically, they take tours of no more than 10 steps. Experiments were run using all combinations of visitor types, and experiments were run to test various characteristics of the resulting tours; however, due to space limitations we cannot provide all of the details in this paper. Instead, we have opted to discuss some of the more intuitive findings for some of the visitor types. We leave more complicated results and discussion for future publications.

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 1, we plot a tour “edit distance histogram” for the informed visitor model—visitors with many goals—both new and returning 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 congestion occurs), and visitors with a fairly low probability of accepting tour guide suggestions (the *possibly* visitor category). In the low density case, visitors choose from half as many goals as 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 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, location of goals within the museum, and the set of prototype tours.

In Figures 2 & 3, we examine the frequency of congested rooms. In Figure 2, we compare the congestion rates experienced by the naive visitor model—visitors with few goals—in trials both with and without the benefit of the TTD-based guide. Note the relative

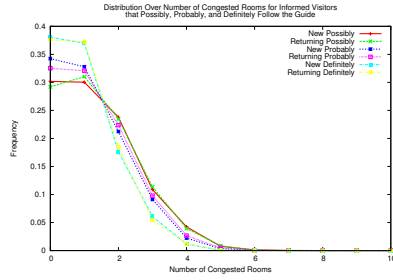


Figure 3: Frequency of Congestion for Informed Visitors with Varying Willingness to Follow the Guide's Suggestions.

position of the curves for the trails 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 those most unwilling to follow the suggestion of a guide. Consider Figure 3, 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.

Our presentation of results in this paper is brief due to space constraints; however, in other experiments, we found that there is a trade off to be made between authorial control (*i.e.* reduced congestion) and player satisfaction (*i.e.* goal satisfaction). The majority of these results are discussed in a longer version of this paper [2].

6. RELATED WORK

Much of the work related to TTD-MDPs can be grouped into two categories: drama management and probabilistic policies for MDPs. Work on tour guides is based mainly in the robotics and ubiquitous computing communities. The technical issues that arise in those communities are generally orthogonal to ours.

6.1 TTD-MDP Related Work

Using a drama manager to guide interactive entertainment was first proposed in 1986 by Laurel [4], formalizing the idea of an agent directing 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 [9] using an expecti-max game tree like search over plot point sequences, and then reformulated as a reinforcement learning problem by Nelson *et al.* [7]. That work led directly to the development of TTD-MDPs to enable variety of experience while preserving authorial intent.

6.2 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.* [3], 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.

7. CONCLUSION

Our autonomous tour guides dynamically construct personalized tours online in response to visitor's reactions to their suggestions. In this paper, we have presented results that show that the tours experienced by visitors who use these autonomous tour guides tend to both contain fewer congested rooms and remain closer to the prototype tours than those experienced by visitors who do not use the guides. Additionally, we find that our prototype-based mixture of Gaussians model allows for visitor autonomy while still resulting in a distribution of tours similar to the class of desirable “good” tours.

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.

8. REFERENCES

- [1] J. Bates. Virtual reality, art, and entertainment. *Presence: The Journal of Teleoperators and Virtual Environments*, 2(1):133–138, 1992.
- [2] A. S. Cantino, D. L. Roberts, and C. L. Isbell. Autonomous nondeterministic tour guides: Improving quality of experience with ttd-mdps. Technical Report GIT-IIC-07-02, Georgia Institute of Technology College of Computing, 2007.
- [3] 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.
- [4] B. Laurel. *Toward the Design of a Computer-Based Interactive Fantasy System*. PhD thesis, Drama department, Ohio State University, 1986.
- [5] V. I. Levenshtein. Binary codes capable of correcting deletions, insertions and reversals. *Doklady Akademii Nauk SSSR*, 163(4):845–848, 1965.
- [6] V. I. Levenshtein. Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady*, 10(8):707–710, 1966.
- [7] 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.
- [8] D. L. Roberts, M. J. Nelson, C. L. Isbell, M. Mateas, and M. L. Littman. Targetting specific distributions of trajectories in MDPs. In *Proceedings of the Twenty-First National Conference on Artificial Intelligence (AAAI-06)*, Boston, MA, 2006.
- [9] P. Weyhrauch. *Guiding Interactive Drama*. PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 1997. Technical Report CMU-CS-97-109.