

Touring Machines: Guide Agents for Sharing Stories about Digital Places

Katherine Isbister

Open Laboratory
NTT Communication Science Laboratories
2-4, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237
JAPAN
katherine@digitalcity.gr.jp

Patrick Doyle

Stanford University
Gates Computer Science Building 2A
Stanford, CA 94305-9020
pdoyle@cs.stanford.edu

Abstract

Human understanding of the physical world frequently derives from the narratives we construct about it. Analogously, we can incorporate narratives in virtual spaces to give visitors a sense of place and purpose. In this paper we describe two projects using interactive characters as a medium for story building and storytelling within and about the digital places they inhabit.

Introduction

According to (Schank and Abelson 1995), the bulk of human knowledge and memory is communicated and encoded in story form. However, there is currently little explicit use of story in the repertoire of interface agents. Many of the interface agents in widespread use are employed largely for simple data-driven tasks such as file searches or context-sensitive help, in which the agents are simply embodied, rather than menu- or prompt-driven, interfaces, but without possessing intelligent believability (Ball et al. 1997; Rist et al. 1997). There has also been active research on pedagogical or utilitarian agents that do exhibit intelligent behavior, though they create narratives only in the sense that their pedagogical requirements impose a structured sequence of behaviors (e.g. (Rickel and Johnson 1997; Lester and Stone 1997)). Finally, there has been work on the creation of stories through the interaction between the user and agents that are characters in the story (Bates 1992; Hayes-Roth et al. 1997). In these systems, story is both the mechanism of the interaction and its purpose.

In contrast, our interest is in the *explicit* use of narrative to help users organize familiar ideas, learn new material, and more effectively engage them with this content. In particular, we are interested in situated stories—stories that are tied closely to the environment at hand, which can illustrate it and bring it to life for the visitor. In this paper, we will discuss two projects that attempt to incorporate situated storytelling into the behavior of an agent.

Both of us are working on projects that involve creating interface agents that can act in the human role of “guide” to a digital location. In examining qualities a human guide must possess, we learned that “interpretation” of the loca-

tion is one of the most important tasks. Interpreting is bringing a place to life using one’s own perspective and artifacts at hand, and usually involves telling stories to visitors, to help make the account more vivid and emotionally engaging to them (Pond, 1993).

We will briefly summarize the structures of our projects and how we are attempting to build guides capable of this sort of interpretation. We discuss the direction this research is taking, and close with a list of challenges for agent tour guide design that we hope will also serve as guidelines for their construction.

Kyoto Digital City Tour Guide

Isbister is currently part of a team that is building a digital version of Kyoto, Japan (<http://www.digitalcity.gr.jp>). This online city is meant to be an evolving resource both for outsiders and for Kyoto residents, with a rich store of useful local information. It currently consists of a database of geographically-arranged links to local websites, with 2D and 3D interfaces for browsing these links. The 3D interface to the city is being constructed in a new development environment called 3DML. 3DML uses texture mapping and simple geometric forms to enable users to rapidly create photo-based 3D worlds. The worlds can be explored using Doom-like navigation. Visitors can also click on buildings to open related Web links.

One of Isbister’s tasks is to ensure that the Kyoto digital city will be an inviting place for real people to use—a truly useful and viable tool. As a visitor herself, she is focusing on making the site approachable and engaging for outsiders who might want to someday visit Kyoto. To help accomplish this goal, she is creating an agent-led group chat tour of the city. The tour will be a point of entry to the online resource and to Kyoto itself, ideally increasing visitor interest in and use of the digital city. Isbister hopes the group tour will also encourage dialogue and relationships among those who participate, as well as increase exposure to Kyoto’s history among those who are friends and family of people who participate in the tour.

In creating the agent, Isbister is focusing on developing storytelling strategies that will produce an engaging experience for tour takers. She derived a list of target abilities for the agent by researching the behavior of actual tour

guides in Kyoto. Isbister hopes that the process of developing the agent's storytelling abilities will lead to a contribution to the narrative intelligence/agent research domain, specifically involving timing and duration strategies for situated storytelling.

Researching the Tour Guide Role

In preparation for creating the tour guide agent, Isbister went on several guided tours of Kyoto, making notes about how the tour guides did their work. She found that tour guides made use of illustrative stories frequently, supplementing the rich visual environment of Kyoto with explanations of how Japanese people, both past and present, made use of the settings that were visited. Stories included things such as: descriptions of how a given site was constructed and its history of destruction and reconstruction; descriptions of peak historic events that happened at a given site; and descriptions of seasonal events and customary activities that occur at the site. Tour guides also incorporated material from previous tours, discussing what other visitors said and did when they visited the sites.

The tour guides would introduce the stories upon arrival at the site, and at specific points in the site that were directly relevant to the particular story. While visitors looked at the buildings, and took things in visually, the guide would create a narrative context for the site with these stories, providing visitors with stories that they could share with their fellow tour members, as well as with people back home.

The tour guide would time his or her storytelling to the visitors' absorption with the site at hand. That is, the guide would provide story context while the visitors were engaged with the site, and move on when it became clear that the visitors were through looking at the site and were ready for a new venue. In addition, the guide would provide follow-up stories to those that were met with particular interest by visitors.

Isbister once worked as an exhibit sign writer at a zoo, and noticed another quality that should be considered in developing tour guide agents. At the zoo, parents often read the exhibit text and then retold the contents of the signs to their children on the spot. For example, children were most likely to be interested in stories about how animals behave or their unusual characteristics. So a parent would find a fact like this on the sign, and turn and tell the child this particular piece of information, omitting the rest. This kind of content repurposing for particular audience members is an important goal to aim for in designing one-on-one agent guides, and will be discussed a bit later on in the section on the virtual art museum tour guide agent.

To summarize, Isbister observed the following qualities in human tour guide storytelling to groups:

1. Stories were told about particular locations, while visitors looked at those locations.

2. Some stories included material that tour guides incorporated from previous tours.
3. Stories that the guide selected seemed to be those that would lead to easy retelling by visitors to others.
4. Guides adjusted the timing of storytelling as well as choice of follow-up stories based on visitor interest level.

Agent Construction Strategy

Isbister is currently compiling a database of stories that are related to the sites that will be part of the Kyoto Digital City. She is categorizing these stories along several dimensions: what specific location the story is linked to, type of story (historic event, construction-based, tour anecdote, etc.), and length of story version. Isbister plans to have short, medium, and long versions of each story, to allow for length adaptation by the agent to tour member interest level. She also hopes to tag these stories with general interest ratings once she has feedback from actual prototype tour takers about which stories are the most interesting to them. Finally, stories will be grouped based on similarity, to allow the agent to tell a series of related stories.

Isbister will use this database of stories to create an agent-led tour of the 3D interface to the digital city. The tour will be implemented within the World Wide Web environment, using I-Chat and Microsoft's Peedy Agent (see Figure 1). Peedy's tour narrative will be triggered by the group's approach to each of the sites on the tour. Isbister is currently planning to vary Peedy's narrative to adapt to each visitor group's interest level on the following dimensions:

1. Story duration.
2. Number of stories.
3. Length of stay at a given site.

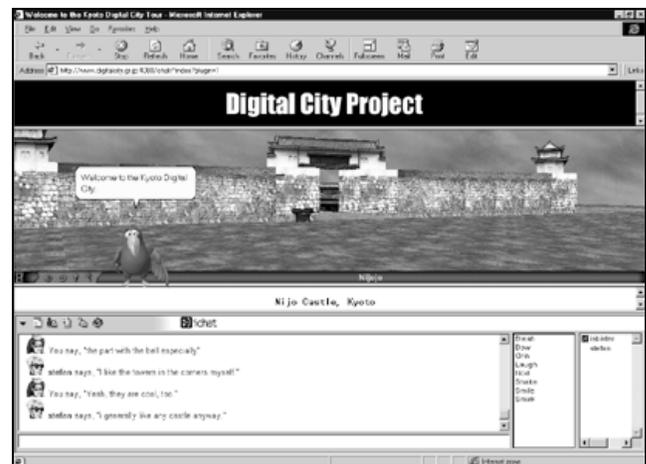


Figure 1. Prototype of the Kyoto Digital City Tour

Currently, plans for tracking visitor interest include noting the amount of conversation among visitors within the

chat environment, searching for positive key words in conversation while at the site, as well as directly asking visitors in prototype tours to rate stories and to indicate that they would like more information and/or stories about a site. Isbister plans to develop and incorporate direct feedback mechanisms into the final interface to the guide agent, as well.

A Virtual Art Museum

Doyle has arrived at the same problem from a different direction. His recent work has focused on increasing the effectiveness of intelligent characters in virtual spaces by adding *annotations* to those environments. Annotations are declarations and procedures embedded in the environment and made accessible to wandering agents in order to give those agents useful information about available content and activities, as well as to assist them in the choice and timing of their actions.

Previous explorations have taken place in text-based environments called MUDs, where agents have used annotations to solve puzzles, play games, and enhance their emotional intelligence with respect to the environment (Doyle and Hayes-Roth 1998a; Doyle and Hayes-Roth 1998b). While the World Wide Web does not offer as sophisticated or flexible a platform for agent behavior, its ubiquity and obvious amenability to simple forms of annotation make it a natural choice for exploring the annotation approach.

One of our current testbeds is a Web-based art museum. The museum consists of a set of galleries, each of which contains rooms filled with artworks. There are currently two galleries: the *Nativity Exhibit* houses medieval and Renaissance religious art revolving around the birth of Christ, while the *Pre-Raphaelite Exhibit* features 19th century artworks in the style of that movement.

Within these galleries, annotations provide a visiting agent with details about their artworks, including their historical contexts, the lives of their artists, their relationships to other artworks in the gallery, and details about specific characteristics of the art. Any agent familiar with the artwork ontology can explore the site together with a user, lecturing, providing commentary, playing games, and guiding the user according to its own lights, the site designer's, or the user's preferences.

The Web agent used in the virtual art museum is a character called Merlyn, aptly named after the forgetful old wizard of T.H. White's *The Once and Future King*. The animation and speech synthesis systems are provided by the Microsoft Agent package. Figure 2 shows Merlyn in the Briar Rose Room of the Pre-Raphaelite gallery, in the process of telling a story to the user. Merlyn is designed to accompany young children, and the control interface and his patter reflect this. When he is provided with prewritten text by the Web site (as is the case with the site tours), he performs simple syntactic manipulations on the text, such as pauses, stutters, or interjections, to personalize its flavor.

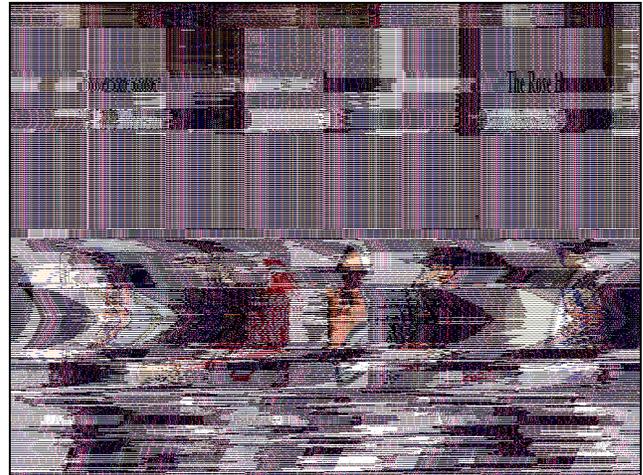


Figure 2. Merlyn in the Art Museum

The annotation framework provides a simple structure for implementing tours across different parts of the site. For these Web experiments, annotations are written in the Extensible Markup Language (XML). Each Web page in the site has a corresponding XML file, which the agent requests from the server when the page is loaded. XML has the advantage of being simple and highly flexible, and its growing use on the Web gives us hope that we may be able to test our agents on sites designed by others.

Figure 3 shows a fragment of a tour annotation associated with one of the rooms in the museum. Tours presently consist of a linear sequence of stops, at each of which there is a set of texts available for presentation. Texts are categorized according to user age or interest (e.g., "child," "adult," "historian," "artist"). Each artwork may have multiple tours associated with it, though currently no artwork may appear within a single tour more than once.

```
<TOURS>
<TOUR NAME="Sleeping Beauty" TYPE="story">
<TOPIC>the legend of Sleeping Beauty</TOPIC>
<TEXT TARGET="ALL">
Once upon a time, there lived a
beautiful princess in a great castle...
.
.
.
</TEXT>
<STOP>
<NAME>Briar Rose Room South Wall</NAME>
<URL>gallery/briar/briar4.html</URL>
</STOP>
<NEXT-STOP>
<NAME>Briar Rose Room North Wall</NAME>
<URL>gallery/briar/briar2.html</URL>
</NEXT-STOP>
</TOUR>
</TOURS>
```

Figure 3. Fragment of an XML Tour Annotation

Notice that the annotation mechanisms assume no special character-like qualities of any visiting agent, so they can be useful to Web spiders, search engines, or purely functional agents of other kinds. However, our interest is in embodied, personable tour guides, to which we restrict this discussion.

Merlyn provides narrative in two ways. First, by his descriptions of artworks, artists and entities, which are either offered autonomously or upon user request. These are produced with simple template-driven natural-language generation that uses the XML data as source. Merlyn has a high degree of control over the form of this storytelling.

Secondly, Merlyn produces narrative through the site tours that he offers. The XML for the tours now contains complete stories written in English. Merlyn presents these with superficial changes, and expands upon them with more details using the other XML content according to the user's interests. Here we are sacrificing some personality for the sake of a well-scripted tale; ultimately we would like the tour annotations represented in a deep, rather than surface, manner, but the complexity of that problem is beyond the scope of our work.

Currently Merlyn relies solely upon tours embedded in the environment, though we are in the process of allowing users to create their own stories. Using the *Scrapbook*, which records snapshot images and basic facts about every artwork the user sees, it is possible for the user to build new stories by specifying sequences of artworks and appropriate corresponding text. Merlyn will remember these tales and offer them to their author or any new children with whom he interacts. In this way the users enable the agent to become a more entertaining guide just as the environment does.

Challenges

While interactive tour guides offer the possibility of increased user attentiveness, retention, and enjoyment (Lester et al. 1997), these benefits require we make these agents believable to their users, both as living entities and as expert guides. Our research suggests that the following four traits are critical for creating believable and compelling guides. The first three points have emerged from Doyle's work on annotation-based tour guide storytelling; the last has emerged from Isbister's investigation of human tour guides leading group tours.

- *Intelligent reincorporation.* Reincorporation in this context means the reintroduction of ideas or entities that have been seen earlier in the tour, with some reference to how they relate to the topic under discussion. In improvisational theater (Johnstone 1992) it is well understood that reincorporation is a key to building a story satisfying to the audience; a sequence of unrelated events does not make a story. Similarly, we suspect that a key to creating an effective tour narrative is

reincorporation of earlier material. This requires that the agent not only track what material has been seen but also when connections are either pedagogically or dramatically effective.

- *Empathy with the content.* (Elliott et al. 1997; Elliott 1998) has argued that understanding of and reasoning about narrative is strongly associated with reasoning about emotion. If we think of stories as descriptions of sequences of events that have emotional associations for the listener, then we can enhance the significance of our tour content to the user by infusing emotion into the presentation. Beyond incorporating emotions in the presented material, we can build our agents to react emotionally to what they present. This not only heightens the significance of these narratives, it also enhances the believability of the tour guide.
- *Presentation through personality.* The most obvious flaw of an intelligent character that relies upon pre-written text is that the text must either be designed for that particular personality, rendering it dangerous for other personalities to use, or so devoid of character that the recitation seems stilted and unbelievable. Ultimately, one would desire an annotation language and an agent powerful enough that the concepts could be explained, reinterpreted, and formed into dialog by the agent. Unfortunately, this is still a hard unsolved problem. Syntactic sugar is a simple approximation that frequently provides good results, as users come to associate idiosyncratic verbal behaviors with the agent's personality. Nonetheless, a more intelligent mechanism for integrating content and presentation will ultimately be required for widespread use.
- *Artful timing/delivery.* As noted in the section on human tour guides, successful storytelling involves knowing when to begin a story, how long it should last, and whether one should elaborate with related stories. Making good decisions about timing and delivery requires the ability to detect user interest, and react appropriately to it. Human beings use many subtle cues to indicate low or high engagement, and detection of user interest is currently very primitive. This will continue to be an important area of research, both for development of characters and for development of satisfactory and subtle interactions with interfaces, in general.

Future Plans

Both Isbister and Doyle are interested in incorporating visitor reactions and stories into the repertoire of tour guide agents. Isbister is interested in allowing web-site hosts (such as shop owners or temple leaders) to provide stories that can be retold by the tour guide agent to visitors during

the tour. Isbister is currently prototyping a web-based input form, that will allow site hosts to add their own stories for the agent to retell when their site is visited. She is also experimenting with ways to solicit direct visitor input about interest and engagement with stories, during the tour, without interrupting the flow of the experience.

Doyle is expanding the virtual art museum with additional areas in order to more fully explore the problems associated with annotated environments as those environments grow in size. Merlyn is being rewritten as part of an effort to build a unified basic annotation-sensitive believable agent architecture, one that can support a variety of interfaces and types of environmental annotations. He hopes to employ adaptive modeling of the user both to enhance believability and improve the agent's effectiveness as a guide.

Conclusions

Storytelling is an important part of the human role of tour guide. By telling stories, they convey content in structures that give it meaning. Using reincorporation, emotion and personality, and artful timing, human guides can present the entities their users see effectively. We anticipate that the lessons we are learning designing interface agents according to these principles will allow us to produce sophisticated storytelling agents, and will be useful to any designers building characters to interact in social settings and tasks.

Acknowledgments

Patrick Doyle's work was supported by NSF Grant IRI-9424261, gifts from Microsoft and Intel, and an Intel Foundation Fellowship.

References

Ball, G.; Kurlander, D.; Miller, J.; Pugh, D.; Skelly, A.; Stankosky, A.; Thiel, D.; van Dantzich, M.; and Wax, T. 1997. Lifelike computer characters: the Persona Project at Microsoft Research. In J. Bradshaw, ed., *Software Agents*. Menlo Park, CA: AAAI Press.

Bates, J. 1992. The nature of characters in interactive worlds and the Oz project. Technical Report CMU-CS-92-200, School of Computer Science, Carnegie Mellon University.

Doyle, P., and Hayes-Roth, B. 1998a. Guided exploration of virtual worlds. In F. Sudweeks, M. McLaughlin and S. Rafaeli, eds., *Network and Netplay: Virtual Groups on the Internet*. New York: MIT Press.

Doyle, P., and Hayes-Roth, B. 1998b. Agents in annotated worlds. In *Proceedings of the Second International Conference on Autonomous Agents*. New York: ACM Press.

Elliott, C.; Lester, J.; and Rickel, J. 1997. Integrating affective computing into animated tutoring agents. In *Notes of the IJCAI '97 Workshop on Animated Interface Agents: Making Them Intelligent*, 113-121.

Elliott, C.; Brzezinski, J.; Sheth, S.; and Salvatoriello, R. 1998. Story-morphing in the Affective Reasoning paradigm: Generating stories automatically for use with "emotionally intelligent" multimedia agents. In *Proceedings of the Second International Conference on Autonomous Agents*. New York: ACM Press.

Hayes-Roth, B.; van Gent, R.; and Huber, D. 1997. Acting in character. In R. Trappl and P. Petta, eds., *Creating Personalities for Synthetic Actors*. Berlin: Springer-Verlag.

Johnstone, K. 1992. *IMPRO: Improvisation and the Theater*. New York: Routledge.

Lester, J.; Converse, S.; Kahler, S.; Barlow, T.; Stone, B.; and Bhogal, R. 1997. The persona effect: affective impact of animated pedagogical agents. In *Proceedings of CHI '97*. New York: ACM Press.

Lester, J., and Stone, B. 1997. Increasing believability in animated pedagogical agents. In *Proceedings of the First International Conference on Autonomous Agents*, 16-21. New York: ACM Press.

Pond, K. 1993. *The Professional Guide: Dynamics of Tour Guiding*. New York: Van Nostrand Reinhold.

Rickel, J., and Johnson, W. L. 1997. Integrating pedagogical capabilities in a virtual environment agent. In *Proceedings of the First International Conference on Autonomous Agents*, 30-38. New York: ACM Press.

Rist, T.; André, E.; and Müller, J. 1997. Adding animated presentation agents to the interface. In *Proceedings of the 1997 International Conference on Intelligent User Interfaces*, 21-28. New York: ACM Press.

Schank, R. and Abelson, R. 1995. Knowledge and memory: the real story. In R. Wyer, ed., *Advances in Social Cognition*, vol. VII. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 1-86.