The VISTA Project: An Agent Architecture for Virtual Interactive Storytelling

Elizabeth Figa

School of Library and Information Sciences
University of North Texas
P.O. Box 311068
Denton, Texas 76203
E-mail: efiga@lis.admin.unt.edu
WWW: http://courses.unt.edu/efiga

Paul Tarau

Department of Computer Science
University of North Texas
P.O. Box 311366
Denton, Texas 76203
E-mail: tarau@cs.unt.edu
WWW: http://www.cs.unt.edu/∼tarau

Abstract. VISTAs, Virtual Interactive Story Telling Agents, interact with users through natural language query/answer patterns derived from the analysis of narrative content developed from multimedia. This paper describes the rationale for agent development, their software components, and the overall architecture for VISTAs, which are used as a form of high-level information retrieval for educational and entertainment purposes. The VISTA agent architecture merges multiple technologies to build an information system with a query/answering front-end user interface and a back-end natural language and knowledge processor provided by the Jinni 2002 Prolog complier.

1 Introduction

Life is, truly, the grand narrative. Some theorists believe humans are born with stories living inside us and that we develop language for the purpose of being able to tell those stories. Imagine a group of people sitting around an evening campfire – the atmosphere is rich with mood and tone – and while around that fire, storytelling will naturally emerge. Stories enable children and adults to understand and make meaning in their lives [1]. Storytellers tell tales of life past, present, and future: *"story is the richest heritage of human civilizations"* [2].

Storytelling has emerged from the grande oral tradition into "modern day" platforms represented in books, dance, music, theatre, movies, etc. With the advent of the Internet/World Wide Web and new technologies to support storytelling, narrativity in the "virtual campfire" has been transformed into a more

dynamic and powerful system of communication with sound, music, visuals, and interactivity. Story preserves, perpetuates, and transforms culture and is finding new applications in education, corporations, industry, and entertainment – in settings in which people interact or seek to "escape". Some call it *virtual storytelling* others call it *digital storytelling*; whatever it is called, it is an emerging frontier with compelling possibilities. Narratives can be captured, displayed, indexed, retrieved, and transformed for new uses and applications for enjoyment, work, and research.

1.1   University of North Texas Digital Storytelling Project

The VISTA Project is part of the University of North Texas Digital Storytelling Project, both developed, in part, to support the online teaching of storytelling [3] to over 150 students per year in a Web-based graduate-level course. The nature of teaching storytelling as an art form and technique for communication has been proven to work very well in the Web-based environment. WebCT, a proprietary software product, serves as the home base for the computer supported cooperative work necessary to effectively teach this content online. This platform offers asynchronous threaded discussion forums, synchronous online chat sessions, email, and high-level interactivity.

Storytelling is a performance-based learning experience and the use of digital storytelling technologies (digital video and digital audio recording software) is of critical pedagogical importance. Students submit three digital storytelling performances, which are streamed to play over the Web for performance analysis and critique via teacher and peer review, as well as to analyze the story content. Professional storytellers are recorded in studio as exemplars for excellence in storytelling and these performances are linked within 13 modules classified by themes. These professional performances and their story content are similarly analyzed in module discussions and in online synchronous chat sessions with guest artists.

The Digital Storytelling Project and the VISTA Project were also developed to address the issues involved in capturing live communication events (storytelling performances) in digital forms, establishing methodologies to collect and preserve this digital matter and their related narrative transcripts, developing metadata standards to make the digital and transcribed matter indexable and retrievable, developing Web-based platforms to distribute the digital communication (performances), and developing new tools to make the performances and the narratives that emerge from them more interactive and educationally beneficial. A streaming media server is used to support over 400 digitally recorded storytelling performances that are donated to the collection by students and professional storytellers. This paper will describe the architecture for these agents and how virtual interactive storytelling agents serve as an educational tool to help an individual or group deeply explore the narratives that are generated from multimedia content as an information access and retrieval system for learning (feedback loop).
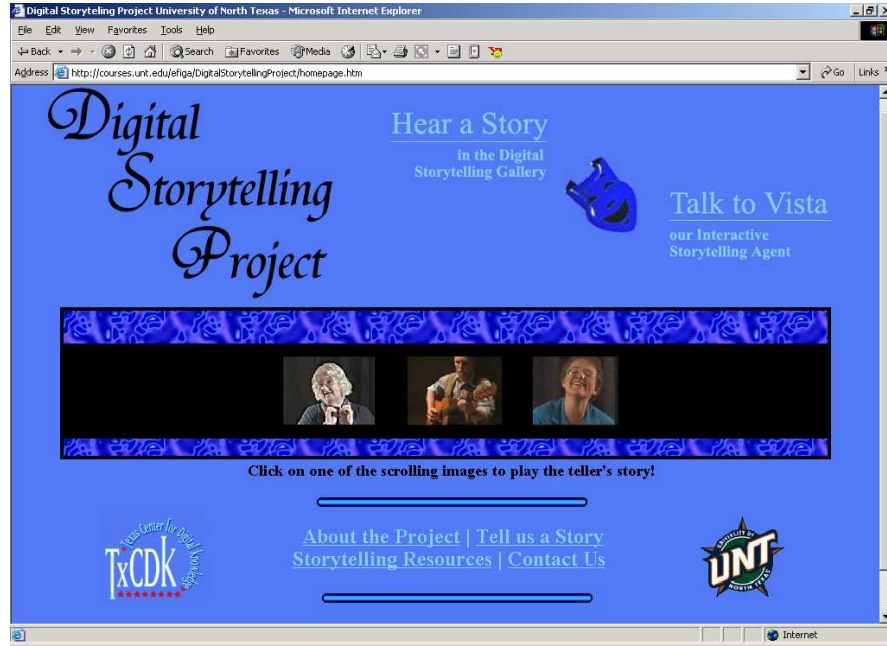
Fig. 1. The University of North Texas Digital Storytelling Project

1.2 Virtual Storytelling Agents

Our interest in Virtual Storytelling Agents has emerged from the development
of information access and retrieval systems built to ensure the preservation of
the cultural heritage represented by stories. Consider that many of the most
prominent professional storytellers in the world are aging and because the na-
ture of their performance work is ephemeral and not replicable, artifacts of world
culture are being lost. In other types of work settings, corporate memory and
organizational knowledge is being similarly lost or not exploited for its optimal
use. Using virtually reality technologies for storytelling has been shown to "bring
to life" the collaborative computer-centered work environments necessary to sus-
tain and make thrive the work of distributed teams, groups or academic classes
working in virtual environments [4]. Developers approach this in a number of
ways through new technologies [5], applications [6], authoring tools [7], virtual
characters [8], and models for narrative construction [9]. The core constructs of
these developments are stories, techniques to share stories, and methods to make
meaning of them for purposes of entertainment or to support work functions or
learning.

Virtual Interactive Story Telling Agents (VISTAs) are coded as a combina-
tion of AIML [10, 11] scripts (a subset of XML specialized to support AI appli-
cations) and rules in a Prolog knowledge base. The AIML scripts (query/answer
exchanges about a story narrative) are generated from transcripts of the live

performance digital video/audio recordings of storytellers, synchronous online query/answering chat sessions (which are automatically transcripted) about a story's content, and indexed information about the storytelling event, the storyteller, etc.

The agents interact with the user through natural language query/answer patterns to establish dialogue about narrative content while a Jinni 2002 based Prolog knowledge server [12] provides them inferential capabilities, *blackboard* based multi-agent coordination, and meta-search over XML and HTML Web-documents. These powerful inferential capabilities allow the agents to interact with and answer questions about the story in dialogic exchanges with users.

Such interactions allow a user to learn about the content of a story by asking the questions he/she personally needs or wants to ask. As a generic design pattern, reusable in the context of other forms of cooperative work and learning environments with a fixed ontology, our agent architecture provides interactive information retrieval for various forms of multimedia content.

The components of the system include a web client with a video/audio agent interface, an HTTP and media server, a story database (access points to the stories via an index in a library of stories) [13, 14], AIML Programs, and Prolog Knowledge Base Programs executed by the Jinni system.

2   The VISTA Agent Architecture

The VISTA Project components allow users to search the story database by title, topic, performer name, etc., for a digital video/audio storytelling performance, select the one they want to play, view the performance via a streaming media player, simultaneously scroll the narrative transcript of the storytelling, and/or interact with VISTA in a dialogue about the narrative. AIML scripts are programmed to answer questions about themes, motifs, [15, 16] characters, and story content. The query/answer format supports individual learning styles and information needs.

The objective is to integrate access to the stream of information sources related to a given story into a natural metaphor – a virtual story telling agent which is modeled after what people ask and answer about the story – while being aware of the ontology and the context of the story, modeled as hierarchy of classes.

The VISTA agent architecture is centered around a public domain AIML script processor [17] and the Jinni 2002 Prolog compiler [18], both written in Java and able to run as extensions of a Web server. The AIML patterns closely mimic the online chat session's query/answer correlations although they also consult the Jinni 2002 Prolog knowledge base. The knowledge base creates an agent instance based on the class to which the story is known to belong and provides inferences about related stories and default assumptions, which are used for queries not covered by the pattern extracted from the online chat sessions.
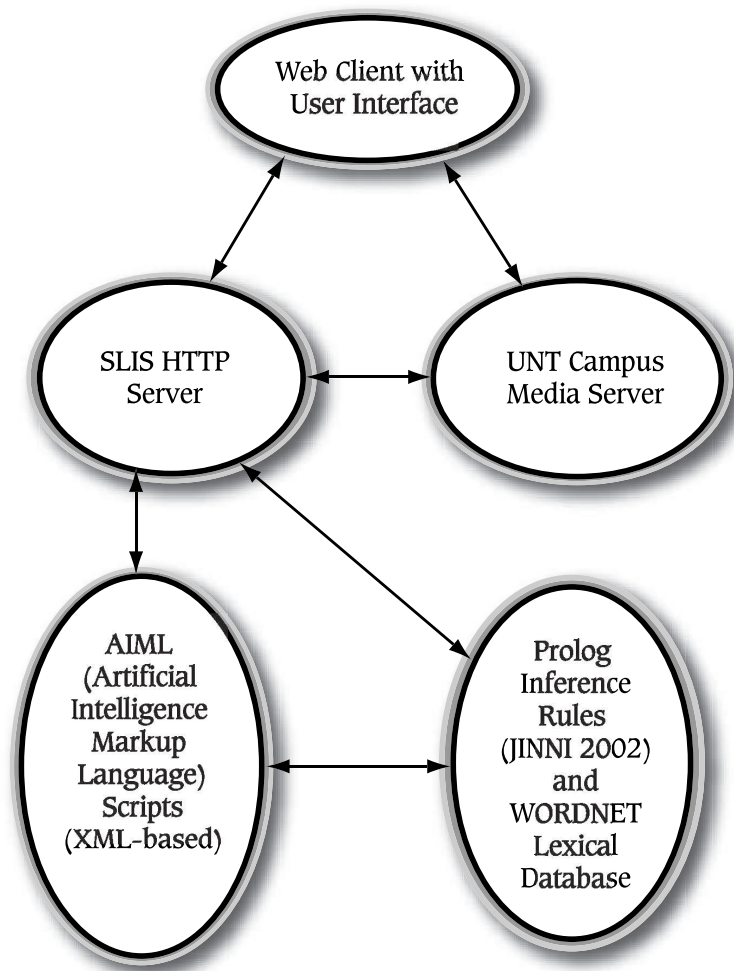
4

Fig. 2. The VISTA Agent Architecture

2.1 Object Oriented Story Hierarchies

The query/answering process about a given story is modeled as a combination of story specific AIML query/answer patterns, a generic AIML pattern library, and a set of Jinni 2002 classes implementing the underlying storytelling ontology that emerges from classifying the stories by themes, motifs, genres, and other indexing schemes [15].

Jinni supports multiple cyclic inheritance [18] allowing stories to be organized based on multiple classification criteria, very much as if they were related Web pages linked to each other. Jinni's Object Oriented Prolog layer is built as a natural extension to ISO Prolog. Classes are simply Prolog files with *include* declarations. As the dispatching of method calls is handled at compile time and instances are lightweight, Jinni 2002's Prolog Objects are extremely efficient. Prolog class files can be located at arbitrary URLs on the Web and can inherit predicate definitions from each other. When story instances are created, the object constructor receives the URLs to the locations of the multimedia (digital video/audio) recording of the storytelling event, the story transcript, and the log of the story-related query/answering chat session.

2.2 Story Specific AIML Scripts and Question Answering Beyond AIML

An online chat transcript is used for establishing the AIML query/answer patterns through an example driven learner implemented in Prolog. The query/answer correlations are used to generate an AIML script specific to the story, as well as a number of action rules in Prolog, allowing the agent to play the story over the Web or to retrieve and display specific sections of the transcript.

AIML has good handling of individual patterns but has limited generalization and inference capabilities. We extend AIML-based pattern processing with a logic-based engine. The engine consists of a natural language parser, a common sense database, and a lexical disambiguation module, as well as a set of transformation rules mapping surface structures to semantic skeletons in a way similar to the language processor described in Tarau et al. [19]. The inference engine uses a dynamic knowledge base, which accumulates facts related to the context of the interaction. Such facts can be used for future inferences. This dynamic knowledge base works as a short-term memory similar to the one implicit in human dialogue and provides means to disambiguate anaphoric references.

2.3 The Knowledge Processor

Logic programming (the programming paradigm on which Jinni is based) provides well-understood, resolution-based inference mechanisms. *Unification* – its key selection mechanism – provides generalized parameter passing and associative search of data objects on Jinni's blackboards, allowing agent components to interoperate effectively.

The system organizes agent components in a library of roles, behaviors, and event processors. Storytelling agents are built by gluing together, through multiple inheritance mechanisms such roles, behaviors, and reactions to events. Traditional inheritance has been confined to trees (simple inheritance) or lattices (multiple inheritance). This contrasts with the dominant information-sharing model – the World Wide Web – which has an arbitrary directed graph structure. Intuitively, this allows story classes to be aware only of a small set of similar "neighbors" and be able to safely import roles and behaviors without being aware of the complete story class library.

3 Extending and Automating the Capabilities of the Storytelling Agents

The agent technology involved in the automation of the interactive chat query /answer patterns needs both a story analysis and a story generation component. The analysis capabilities are needed to understand the question and the generation capabilities are needed to construct the answer.

Agents use two orthogonal techniques to answer questions. The first technique uses transcripts from human chat sessions as analogical sources to replicate what humans do directly; from questions we build question patterns and we map them to answers on a one-by-one basis. The question patterns do not invoke inference rules – their responses are mostly reactive and context-free. These query/answering rules put more emphasis on giving the user the illusion of a natural dialog more than on precision of the content of the dialogic interaction. The second technique is inferential/deductive, it tries to identify, at least partially, what the focus of interest is in the question and consults the story classification hierarchy and related dialogue patterns to handle unknown situations. This technique involves story analysis through the development of *abstract story traces* and *ontology driven story projections*.

3.1 Story Analysis and Understanding via Abstract Story Traces

Stories are fairly intricate fragments of natural language with strong intertextual connections. A story contains complex speech acts sometimes expressing emotions or requests for participation as well as metalevel judgments involving the complete understanding of the whole interaction context. Understanding them to the point where one can provide answers to a human user about the higher-level abstractions of a story's content requires advanced natural language disambiguation and concept extraction.

Answering *what-is-this-about* questions is relatively easy – by extraction of dominant nouns and noun phrases from each story. However, creating a dialogue to get at the deeper hermeneutics of the story or the impact of a storytelling performance narrative upon an individual is harder. Different people will select a different trace in a story to chat about. A story trace is a sequence of meanings extracted from the lexical material of a story to which one or more meaning

transformations are applied. The semantic ambiguity coming from the polysemy (multiple meanings) of the lexical material is intensified by the pragmatic ambiguity of the listener's personal experience, the parameters of the storytelling performance, and the nature of the multimedia experience (seeing video, hearing audio tracks, reading a transcript, listening to a musical story, etc.)

Our *abstract story traces* involve the use of the Wordnet lexical database [20], a dictionary providing morphology, syntactic categories, as well as a many-to-many word-to-meaning mapping. Through the use of Wordnet, abstractions can be traced to help determine what a given story and its parts are about. Wordnet contains semantic links to allow the users to navigate on a network of meaning-to-meaning relationships, like synonyms, antonyms, hypernyms, and meronyms [1]. Meaning elements obtained by navigating Wordnet concept hierarchies naturally generalize the meaning of individual sentences. By starting from a story's lexical material and working upward in word meaning hierarchies to understand higher level indexing terms, story similarities and differences can be compared and query/answer patterns can be automatically extracted.

The first mechanism we are exploring is *abstraction* of story traces by navigating upward in the hypernym hierarchy. Semantic ambiguity is reduced when one moves up toward a more general concept. This narrows the sequence of sentences in a story to a few abstract patterns forming the skeleton – *the bones* – of a story. The second mechanism is projection. These *"ontology-driven"* story projection operations consist of selecting only sentences subsumed by a set of concepts (meanings associated to nouns) or predicates (meanings associated to verbs). We fix the ontology to a set of concepts and predicates and extract a story trace restricted to their hyponyms and meronyms.

The abstract story traces inform the agent architecture in a number of ways. Traces help construct a simplified story skeleton and can help classify stories based upon a number of predetermined story types. Traces improve the query/answering mechanism because knowledge about similar stories can, to some extent, be reused in the absence of specific knowledge of a given story. Lastly, if a story trace matches a question pattern, it can generate the appropriate answer pattern, which then can be filled in with the concrete lexical material from the interaction context and the story content.

4 Future Work

VISTA agents are specialized to their context: teaching storytelling online and providing a methodology for users to explore the meaning of a story based upon their individual interests. The agent technology is part of a computer supported cooperative work architecture that supports online human interaction, the delivery of storytelling performances via the Web, and the development of the audience-storyteller feedback loop. Future research emerging from University of

---

[1] Hyponyms are subsumed concepts and meronyms are concepts referring to parts of an entity – for example, "tiger" is a subsumed concept of "animal"; "wheel" is a meronym of "car".

North Texas Digital Storytelling and VISTA Projects will focus on the following topics:

- How virtual interactive storytelling agents may impact the work of individuals or groups as they more deeply explore the narratives generated from multimedia and other forms of narrative content.
- How this architecture, and its generic design, may be transferable to other types of interactive chat settings.
- Improving the methods by which agent scripts are extracted from interactive query/answering transcripts.
- The development of metadata standards for classifying and indexing digital storytelling performances to represent their multi-dimensional nature as a form of knowledge that can be captured for use in multiple contexts.
- The development of automated generation of metadata via the extraction of key terms.
- Further work with the story traces as a method to deeply analyze story content for abstraction and ontology-driven story projections to match question patterns and generate answer patterns.

5 Conclusion

This paper has described the VISTA Project, an agent architecture for virtual interactive storytelling and the technical constructs used to "build" the interactive agents. The VISTA project uses an XML-based Web Interface, object-oriented story hierarchies, AIML scripts, a natural language and a knowledge processor provided by the Jinni 2002 Prolog compiler, and techniques to produce abstract story traces and ontology-driven story projections for the VISTA query/answer patterns and transcripts. The VISTA agent architecture merges these technologies to build an information system that allows users to select and view digital recordings of storytelling performances, read narrative transcripts from performances, and chat with a storytelling agent about the content of the story. The interactivity of the digital storytelling experience provides a dual sensorial and cognitive experience for a user – entertainment pleasure via the video/audio component of the storytelling and enhanced cognitive development about the content of a story via the ability to read/study the transcripted narrative and participate in an interactive chat with a storytelling agent. The agent technology has applications for online teaching and in building shared virtual environments to support learning and understanding from narratives generated from multimedia or other forms of narrative content. The overall project has applications in comparative story analysis, the development of new story skeletons, and story generation. Finally, more and more streaming media is being made available via the Web and results of this project offer methods to automate the indexing of metadata for the narrative content of multimedia for improved information access and retrieval.

References

1. B. Bettleheim. *The Uses of Enchantment.* Alfred A. Knopf, New York.

2. O. Balet, G. Subsol, and P. Torget. Preface. Virtual Storytelling: Using Virtual Reality Technologies for Storytelling. In *Proceedings of the International Conference on Virtual Storytelling*, Avignon, France, 2001.

3. Elizabeth Figa. SLIS5440 STORYTELLING. Technical report, University of North Texas, 2002. Available at http://courses.unt.edu/efiga/FACULTYPAGE/SLIS5440SyllabusInformationPage.htm.

4. *Proceedings of the International Conference on Virtual Storytelling*, Avignon, France, September 2001.

5. O. Balet, P. Kafno, F. Jordan, and T. Polichroniadis. The VISIONS Project. In *Proceedings of the International Conference on Virtual Storytelling*, Avignon, France, 2001.

6. M. Rousseau. The Interplay between Form, Story and History. In *Proceedings of the International Conference on Virtual Storytelling*, Avignon, France, 2001.

7. M. Zancanaro, A. Cappelletti, and C. Signorino. Interactive Storytelling: People, Stories, and Games. In *Proceedings of the International Conference on Virtual Storytelling*, Avignon, France, 2001.

8. M. Cavassa, F. Charles, and S. Mead. Characters in Search of an Author: A.I. Based Virtual Storytelling. In *Proceedings of the International Conference on Virtual Storytelling*, Avignon, France, 2001.

9. C. Fencott. Virtual Storytelling as Narrative Potential: Towards an Ecology of Narrative. In *Proceedings of the International Conference on Virtual Storytelling*, Avignon, France, 2001.

10. A.L.I.C.E. AI foundation. Artificial Intelligence Markup Language (AIML). Technical report, A.L.I.C.E. AI foundation, 2001. Available at http://alice.sunlitsurf.com/TR/2001/WD-aiml/.

11. Richard Wallace. AIML Pattern Matching Simplified. Technical report, A.L.I.C.E. AI foundation, 2002. Available at http://alice.sunlitsurf.com/documentation/matching.html.

12. Paul Tarau. Inference and Computation Mobility with Jinni. In K.R. Apt, V.W. Marek, and M. Truszczynski, editors, *The Logic Programming Paradigm: a 25 Year Perspective*, pages 33–48. Springer, 1999. ISBN 3-540-65463-1.

13. A. Aarne. *The Types of Folktales A Classification and Bibliography. Translated and revised by Stith Thompson.* Helsinki: Academia Scientiraum Fennica, 1961.

14. D. Ashliman. *A Guide to Folktales in the English Language: Based on the Aarne-Thompson Classification System.* Westport, CT: Greenwood Press, 1987.

15. S. Thompson. *Motif-Index of Folk Literature: A Classification of Narrative Elements in Folktales, Ballads, Myths, Fables, Medieval Romance, Exampla, Fabliaux, Jest-Books, and Local Legends. 2 vols.* Westport, CT: Greenwood Press, 1956.

16. M. MacDonald. *Storytellers Sourcebook: A Subject, Title, and Motif Index to Folklore Collections for Children, 1983-1999.*

17. A.L.I.C.E. AI foundation. Alicebot Program D - reference implementation. Technical report, A.L.I.C.E. AI foundation, 2002. Available at http://alice.sunlitsurf.com/downloads.

18. BinNet Corporation. Jinni 2002 A High Performance Java and .NET based Prolog for Object and Agent Oriented Internet Programming. Technical report, BinNet Corp., 2002. Available at http://www.binnetcorp.com/download/jinnidemo/JinniUserGuide.html.

19. Paul Tarau, Koen De Boschere, Veronica Dahl, and Stephen Rochefort. LogiMOO: an Extensible Multi-User Virtual World with Natural Language Control. *Journal of Logic Programming*, 38(3):331–353, March 1999.

20. George Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine Miller. Five papers on WordNet. CSL Report 43, Cognitive Science Laboratory, Princeton University, July 1990.