Ontological Support for Creative Writing

Darina Dicheva*, Christo Dichev*, Mariana Damova#

* Department of Computer Science, Winston-Salem State University, USA
601 Martin Luther King, Jr. Drive, Winston Salem, N.C. 27110
Ontotext AD, Tsarigradsko Chosse 135, Sofia 1784, Bulgaria
{dichevad, dichevc}@wssu.edu, mariana.damova@ontotext.com

Abstract

In this paper we propose an ontological framework for tools facilitating creative writing and story reading. It is based on an ontology implemented as a topic map and employs linguistic analysis methods for discovering conceptual entities in the text.

Introduction

Using their knowledge, people are able to combine objects and events in meaningful semantic sequences when narrating their experiences. Hence, to attain a reasonable level of expressiveness, creative writing systems require a body of facts about the relationships between things and events. Over the past decade, the research on intelligent narrative technologies has demonstrated success in several areas related to education (Rowe et al, 2008, Nakasone et al, 2009). It has been shown that they have a significant potential for enhancing students' learning experiences and of being pedagogically compelling by increasing students' engagement. A set of applications were motivated by the research on planning and plot and character development experienced by writers. Systems such as StoryMill, Celtx, yWriter, were created to keep the bits and pieces of story characters and locations together. Semantic Web technologies provide means to access the wealth of information on the web. Authoring tools like Zemanta provide contextual access to textual and visual resources harvested from the web. However, such software applications are not more than text mining tools based on word processing. Efficient story telling implies deeper accounts for story plans, protagonists' participation, relationships and detailed story lines. In this paper we propose an ontological framework for building tools facilitating story creation and story consumption. To validate it, we are building a Topic Maps based prototype.

The Approach

Creative writing in educational settings requires support for modifying, revising and expanding story elements, such as events, characters, relationships, sequencing, etc., from which different kind of stories can emerge. We propose an ontology for building the conceptual model of a story and drawing the relations and the different stand points from which the story is told. We partition the concept of a story at three levels: conceptual, structural and presentational. For representing the ontology, we have chosen the Topic Maps standard (Park & Hunting, 2002). With its roots in the back-of-the-book index, Topic Maps are a Semantic Web technology, which is appropriate to represent the conceptual structure of a written text. A topic map annotating specific text document can be used for semantic search, navigation, and visualization of the text.

Linguistic analysis and processing are relied upon to recognize language expressions that refer to people and events talked about in the text, their relations, and the temporal profile of the story and transform them into conceptual entities representing the story. The semantic vectors approach (Widows & Ferraro, 2008), which has been successfully used for tasks of automated ontology acquisition, is being employed for this.

The language expressions are dynamically discovered and classified into ontological entities in the process of writing or reading. This enables assisting both the authors in their creative writing and the readers in comprehending the text, by providing a quick view of the current content organization. Such a tool can be very useful in educational context for creating synopses of stories and "story maps".

Ontology for Story Representation

Our focus is on designing an ontology which could support light semantic annotation performed by the writer on-the-fly, that is, in the actual process of writing. Some of the existing narrative ontologies are based on Mieke Bal' narrative theory (Bal, 1997), according to which, narrative texts can be considered at three layers called *fabula*, *story*,

and *text*. While the fabula is a sequence of events experienced by the story characters, the story is a specific presentation of the fabula. The text layer refers to the final "physical" presentation of a story, e.g. the actual words and sentences written by the author. Similarly, we consider three narrative layers: *text*, *fabula* and *entity* layers. The first two can be paralleled to the Bal's text and fabula layers.

Thus, in our approach, the text layer of a story consists of chapters and a chapter consists of scenes. Scenes are distinct parts of the text that represent the setting and the events that occur. A setting, as part of the story, includes the context, facts and conditions in which an event takes place. At the fabula layer, we represent story lines. A story line is a sequence of events. An event is something that happens at a given place and time. Finally, at the story *entity* layer, we represent the entities involved in the events and scenes: the story characters, locations/places, and other (physical or abstract) things, important for the story.

Consequently, at the highest level the proposed ontology consists of:

- Concepts, including Chapters and Scenes (Setting, Events), Story Lines and Entities (Characters, Locations, Items),
- Relationships, including Circumstantial and General relationships, Temporal relationships, and Relationships between story characters.

This ontology is a core component of the proposed framework, enabling the transformation of textual items into conceptual structures, so as to help story writing and understanding of written content.

Architecture

The proposed ontology-based framework for supporting story writing consists of four modules: a text editor, a Topic Map tool, an NLP tool, and a visualization tool.

The Topic Map module consists of a topic map engine and a graphical user interface. The Topic Map interface enables the writer to create topics *on-the-go*, that is, in the process of text writing. The topics can be entered as soon as the entities (characters, places, objects, etc) that they reify are introduced, or at the completion of the current section. The interface allows easy addition of new topics and associations to the pre-defined ontology categories, as well as removing, renaming and editing of topics and associations.

The NLP tool provides means for the automatic creation of some ontological constructs through natural language analysis, such as detecting the boundaries of scenes, recognizing relationships between story characters, and revealing temporal relations between events in the light of interval temporal logic (Allen & Fergusson, 1994), in order to help the writer in ontology building.

The visualization module provides visualization of the constructed topic map or its parts parallel to the running text, and a number of other visualizations, including story lines, the timeline, etc.

A prototype of the proposed architecture - TM4Book - is being implemented. Fig. 1 presents a screenshot from the TM4Book interface. The incorporated text editor saves the text in HTML format. The topic map window displays the ontology based on the loaded text. The developed tool can be deployed in web context as well as in a desktop environment.

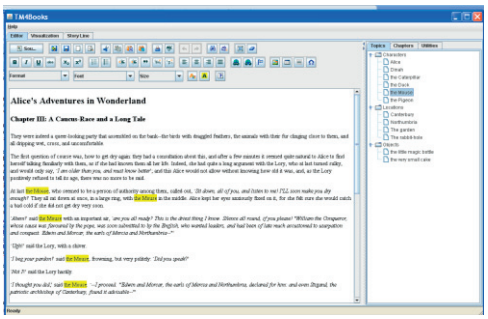

Fig. 1. Screenshot from the TM4Book interface

Conclusion

We presented an approach to building a creative writing and consumption tool based on ontologies implemented as topic maps. We believe that a tool based on this approach can support the integration of the participants in the creative process and can be an efficient companion in any educational process based on text.

Acknowledgement. This research is supported partly by the Bulgarian National Science Fund Project D002-111/15.12.2008 "SmartBook" and by the National Science Foundation under Grant No. DUE-1044224.

References

Bal, M. 1997. *Introduction to the theory of Narrative*. 2nd Ed. University of Toronto Press, Toronto: Canada.

Widdows D., Ferraro K. 2008. *Semantic Vectors: A Scalable Open Source Package and Online Technology Management Application*, LREC.

Allen, J., Fergusson, G. 1994. *Actions and Events in Interval Temporal Logic*. Technical Report, University of Rochester.

Nakasone A., Prendinger H., Ishizuka M. 2009. Isrst: Generating Interesting Multimedia Stories on the Web. *Applied Artificial Intelligence* 23(7): 633-679.

Park, J., Hunting, S. 2002. *XML Topic Maps: Creating and Using Topic Maps for the Web*, Addison-Wesley.

Rowe, J., Mott, B., McQuiggan, S., Robinson, J., Lee, S., Lester, J. 2009 Crystal Island: A Narrative-Centered Learning Environment for Eighth Grade Microbiology. *Workshop on Intelligent Educational Games*, U.K , 11–20.

Zemanta, 2010. http://www.zemanta.com