

Getting the Story Right: Making Computer-Generated Stories More Entertaining

K. Oinonen, M. Theune, A. Nijholt, and D. Heylen

University of Twente, PO Box 217,
7500 AE Enschede, The Netherlands
{k.oinonen, m.theune, a.nijholt, d.k.j.heylen}@utwente.nl

Abstract. In this paper we describe our efforts to increase the entertainment value of the stories generated by our story generation system, the Virtual Storyteller, at the levels of plot creation, discourse generation and spoken language presentation. We also discuss the construction of a story database that will enable our system to learn from previous stories and user feedback, to create better and more entertaining stories.

1 Introduction

Possibly the most important requirement for a successful story generation is that the stories it generates are compelling. In this paper we describe our efforts to increase the entertainment value of the stories produced by the Virtual Storyteller, an agent-based system that generates simple fairy tales. First, we give an overview of our recent work on the three story generation levels distinguished in the Virtual Storyteller: plot creation, discourse generation and presentation. The remainder of the paper is devoted to our future research on a story database that will enable our system to learn from previous stories and user feedback, to construct stories that are captivating and invoke an emotional response.

2 Entertainment at Different Story Levels

The Virtual Storyteller is a multi-agent framework for story generation at three levels, each handled by specialized agents. The first level is that of *plot creation*: generating a structure of causally linked events forming the content of the story (also called ‘fabula’ in narrative theory). The second level is *discourse generation*: producing a text that expresses the plot in natural language. The final level is that of *presentation*, where the generated discourse is expressed in speech by an embodied agent. We discuss each of these levels below.

Plot Creation. One of many requirements for a plot to be entertaining, is the presence of believable characters. It has been argued that users do not care about the characters in a story if these characters do not exhibit emotions [1]. In the Virtual Storyteller, plot creation is based on the actions of semi-autonomous

character agents. To make these characters believable, we provided them with an emotion model inspired by the Hap architecture [8]. As described in [13], emotions that are running high may give rise to unexpected plot twists. For instance, a prince trying to slay a dragon might become so afraid that he decides to flee instead. Such plot twists help to make the story more realistic and entertaining, but they also pose a challenge for the Plot Monitor: an agent that is responsible for guiding the actions of the characters to achieve a proper storyline. In the example above, it could be essential for the story that the dragon is killed. This means that the Plot Monitor must find a way to make this happen in spite of the hero's fear, for instance, by having the hero find a magic sword that makes him invincible, or having the dragon struck by lightning. To make it possible for the Plot Monitor to come up with such solutions, and decide which of them provides the most entertaining storyline, we will extend it with case-based reasoning capabilities, using a story database as discussed in Section 3.

Discourse Generation. It is not just the plot that makes a story entertaining, but also how it is expressed in natural language. Interesting events that are badly narrated still do not make for a very enjoyable story. As shown by Callaway [3], two essential components of a narrative prose generation system are a discourse history and a revision component. The discourse history is used for generating appropriate references to characters and objects, e.g., by using pronouns. The revision component is responsible for aggregating simple sentences into more complex, fluent prose. In user evaluations, texts generated using these two components were judged to have better style, be more readable and more believable than texts generated without them. For this reason, our focus in developing the discourse generation component (the Narrator agent) for the Virtual Storyteller has focused on performing aggregation / revision while also paying attention to pronominalization [5].

The input for the Narrator consists of a plot representation specifying the setting and listing the basic actions of each of the characters, their emotions and the goals they try to achieve. Rhetorical relations such as temporal sequence, purpose, cause and contrast are used to link these elements together. Based on these rhetorical relations, simple sentences are combined into more complex ones using appropriate cue words ('because', 'so that', etc.). Then, redundant parts of the combined sentences are deleted and pronouns are generated using a very simple algorithm based on gender and recency of mention. Via these steps, a simple story fragment that would have originally been expressed as in (1), is now expressed as in (2). (The examples were translated from Dutch to English.)

- | | |
|---|---|
| (1) Diana walked to the desert.
Brutus walked to the desert.
Diana was afraid of Brutus.
Diana walked to the forest. | (2) Diana walked to the desert, and
Brutus walked to the desert too.
Because Diana was afraid of Brutus,
she walked to the forest. |
|---|---|

Version (2) is clearly better than version (1), but further work is required, for instance on lexicalization: choosing contextually appropriate words rather than

always using the same word to express a given concept (e.g., Diana should ‘flee’ rather than ‘walk’ to the forest). Eventually, to create a truly entertaining story, we will also need to address even more difficult tasks such as generating character dialogue, expressing the plot from different perspectives, etc. These are still largely unexplored areas in natural language generation.

Presentation. Finally, the story is presented to the audience using speech synthesis.¹ For a captivating listening experience, we need a more expressive, engaging speaking style than is provided by today’s text-to-speech systems. Therefore we have designed and implemented a set of rules that convert ‘neutral’ speech, as produced by a standard text-to-speech system, into storytelling speech by manipulating the pitch, intensity, and duration of the phonemes in the time domain indicated by tags in the story text. Using this system, we can create a global storytelling speech style which is slower, with longer pauses, than neutral speech and shows much more variation in pitch and loudness. The system can also add prosodic effects to create suspense. Dramatic moments in the story, such as startling revelations, are announced by a steep increase of loudness and pitch. Alternatively, expectation can be built up by a gradual increase in pitch and loudness, accompanied by a decrease in tempo, warning the audience that something exciting is about to happen. A small-scale evaluation of our storytelling speech generation system showed encouraging results, as users found speech fragments generated using our system consistently more exciting than those with neutral prosody. For more details, see [14].

3 Developing a Story Database

For the Virtual Storyteller, we propose to use a database as a memory, to learn from stored information about what makes a good (or a bad) story. The database holds representations of both human authored and computer generated stories and story fragments, annotated with information about their potential effects on the user. These annotations can be used for extensive automated analysis of relevant elements in the story structures. Based on theoretical knowledge about story structure elements that cause certain effects on the receiver [10,11], the system will be able to predict the audience response to new stories that are generated. In addition, users will play an active role during interactive story generation by controlling the actions of one of the characters, thus filling the database with scripts about human behavior in different situations.

Besides information about plots and plot structures [7], the database will contain information about objects and characters, actions (including their conditions and effects), background processes,² and events with different kinds of

¹ Currently, we use a simple animated MSAgent (<http://www.microsoft.com/msagent/>) that uses speech synthesis accompanied with text balloons.

² Background processes simulate the life, physics, biology, chemistry, and evolution in the story world environment.

relations and the users' feedback to these elements. In addition, the formal representation of a story should express the relevant semantic, causal, and temporal relations between various story elements.

Following [2,4,15], we propose to represent stories and story fragments as temporal network models or graphs. The temporal structure of a story is represented as adjacent world states linked via the effects of actions and events, e.g., the causal consequences of the actions performed by the characters. Causal and semantic relations can be divided into different types and may have different strengths [12]. The believability and comprehensibility of a story plot increases if more elements are covered in one coherent and consistent whole by causal and semantic relations, and fewer elements remain uncovered or contradictory [9]. The degree of relevance of an action, a background process, or an event in a story could be computed by combining user feedback annotations with information about (1) the strength and importance of its effect on the story world; (2) the number, strength, and importance of the causally or semantically related elements involved with it; (3) the current goals and motivations of the characters.

Story comprehension is a mental process of detecting character's goals, motivations and plans, which must be integrated into one causal network representation of the story [15]. We need to separate the representation of the characters from the description of the objects in the story world. At least the characters' motivations, goals, plans and properties such as personality and emotional state, and the social and emotional relationships between the characters, including affective and power relations, should be represented in the story database.

The concepts we want to formally represent in the story database are similar to Gratch and Marsella's [6] list of functions that have to be integrated in a single architecture in order to simulate the generality of human emotional capabilities, such as coping and appraisal strategies. Human authored stories based on the general ontology of a single story world description will serve as initial examples for the computational story generation system, as they will be used for case-based plot management by the Plot Monitor, and case-based action planning by the characters during automatic story generation. A case-based approach to story generation has been successfully used before in the MINSTREL system [16].

4 Conclusion

We have described our recent efforts to make the stories generated by the Virtual Storyteller more engaging at different story levels, by providing the characters with personalities and emotions, using aggregation and pronominalization to create more fluent discourse, and by converting neutral speech synthesis to storytelling speech, including suspense effects. We are currently working on a new implementation of our system that will integrate all these components as well as a, still to be developed, story database containing both human authored and system generated stories, annotated with user feedback. This will require addressing some important issues in story representation, as well as adding new

forms of interactivity to the system. The ultimate goal is to enable the system to learn by example to construct better and more entertaining stories.

Acknowledgements. We thank Sander Faas, Sander Rensen, Koen Meijs, Feikje Hielkema, Ivo Swartjes, and Jasper Uijlings for their contributions to the work reported here. The authors participate in the EU Network of Excellence HUMAINE (Human-Machine Interaction Network on Emotion).

References

1. J. Bates. The role of emotion in believable agents. *Communications of the ACM* 37(7), 122-125, 1994.
2. P. van den Broek and R.F. Lorch. Network representations of causal relations in memory for narrative texts: Evidence from primed recognition. *Discourse Processes* 16, 75-98, 1993.
3. C. Callaway. *Narrative Prose Generation*. PhD thesis, North Carolina State University, 2000.
4. R. Fincher-Kiefer. The role of predictive inferences in situation model construction. *Discourse Processes* 16, 99-124, 1993.
5. F. Hielkema, M. Theune and P. Hendriks. Generating ellipsis using discourse structures. *ESSLLI Workshop on Cross-Modular Approaches to Ellipsis*, Edinburgh, 2005.
6. J. Gratch and S. Marsella. A domain-independent framework for modelling emotion. *Journal of Cognitive Systems Research* 5(4), 269-306, 2004.
7. W. Lehnert. Plot units and narrative summarization, *Cognitive Science* 4, 293-333, 1981.
8. B. Loyall. *Believable Agents: Building Interactive Personalities*. Ph.D. thesis CMU-CS-97-123, Carnegie Mellon University, 1997.
9. N. Pennington, R. Hastie. The story model for juror decision making. In R. Hastie (Ed.). *Inside the Juror: The Psychology of Juror Decision Making*, Cambridge University Press, 192-221, 1993.
10. N. Szilas. IDTension: A narrative engine for interactive drama. *Proceedings of the Technologies for Interactive Digital Storytelling and Entertainment (TIDSE) Conference*, 187-203, 2003.
11. N. Szilas. A simulation of the narrative. In *Proceedings of COSIGN-2003*, 2003.
12. I. Tapiero, van den Broek and Quintama. The mental representation of narrative texts as networks: The role of necessity and sufficiency in the causal detection of different types of causal relations, *Discourse Processes* 34(3), 237-258, 2002.
13. M. Theune, S. Rensen, R. op den Akker, D. Heylen and A. Nijholt. Emotional characters for automatic plot creation. In S. Göbel et al. (Eds.), *Proceedings TIDSE 2004: Technologies for Interactive Digital Storytelling and Entertainment*, Lecture Notes in Computer Science 3105, Springer-Verlag Berlin Heidelberg, 95-100, 2004.
14. M. Theune, K. Meijs, D. Heylen and R. Ordelman. Generating expressive speech for storytelling applications. Submitted to *IEEE Transactions on Speech and Audio Processing, Special Issue on Expressive Speech Synthesis*.
15. T. Trabasso, P. van den Broek and S. Suh. Logical necessity and transitivity of causal relations in stories. *Discourse Processes* 12, 1-25, 1989.
16. S. Turner. *The Creative Process: A Computer Model of Storytelling*. Lawrence Erlbaum Associates, 1994.