

Generation of Dilemma-based Narratives: Method and Turing Test Evaluation

Heather Barber and Daniel Kudenko

University of York, Heslington, York, YO10 5DD, England
{hbarber, kudenko}@cs.york.ac.uk

Abstract. This paper describes the analysis of the narrative generation aspect of the Generator of Adaptive Dilemma-based Interactive Narratives (GADIN). This system automatically generates narratives that are focused around dilemmas to create dramatic tension. The dilemmas are based on those clichés encountered in many of today’s soap operas. A story planner creates sequences of actions that lead to dilemmas for storyworld characters. A narrative generated by GADIN is compared to that which took place in a television soap opera. The stories were subjected to a Turing-style test, with 127 participants, which showed that the readers could not determine which narrative had been written by a computer.

1 Introduction

In this paper we analyse a system that automatically generates narratives. This is one aspect of the Generator of Adaptive Dilemma-based Interactive Narratives (GADIN). Dramatic tension is created by incorporating dilemmas for storyworld characters. These dilemmas are based on the clichés found in many contemporary soap operas, such as the trade-off between personal gain and loyalty to a friend. Overarching stories connect these dilemmas within a coherent plotline that is dynamically created.

Story generation systems [1–4] created thus far have written stories which have clearly not had a human author. The stories generated by the GADIN system are more consistent and have a level of dramatic interest equivalent to that in human-authored stories. This is demonstrated here through a Turing-style test, in which participants were unable to distinguish the GADIN narrative from one which had occurred in a television soap.

Soap operas (or soaps) are a popular means of entertainment. They consist a story which never ends, shown in episodes which see the continuation, foundation or conclusion of some of the ongoing storylines. Soaps are commonly known to consist of very similar and clichéd storylines. The clichéd storylines utilised in soaps are usually found to be essentially conflicts (or dilemmas). Writers utilise these dilemmas in the creation of soap stories in which the story is built around the cliché. A general form of each such clichéd dilemma can be determined, and GADIN can generate a soap narrative around these. The soaps which are particularly concentrated on in the discussed version of GADIN are *Hollyoaks*, *Eastenders*, *Coronation Street*, *Emmerdale* and *The Archers*.

2 GADIN

Many storytelling genres make frequent use of clichéd storylines which are created around dilemmas to storyworld characters. These dilemmas can be generalised and the GADIN architecture uses planning to achieve such dilemmas, the combination of plan and dilemma constitute a dramatically interesting sub-story of the generated narrative. Characters act and experience these dilemmas in the course of a GADIN narrative, making decisions and action choices depending on their individual properties and state.

More details on narrative generation in GADIN for soap operas can be found in [5]. The application of GADIN to children’s dinosaur adventures is discussed in [6].

3 Story Quality Evaluation

The Turing Test [7] can be interpreted as requiring a human reader to be unable to tell whether a narrative was written by a computer or human author. In the test described here the reader must decide which of two narratives they think was written by a computer.

To ensure a fair comparison both narrative creations began with the same storyworld state. The characters were limited to only those included in the selected story, but the actions and dilemmas included all of those available to the GADIN system. Subjects who were familiar with the television soap narrative selected were asked to not answer the survey. The names were changed to be anonymous and the same in both narratives.

To make the narratives more readable they were adapted slightly. For example, the actions of a character moving between locations were removed in both versions – this is relevant in the experience but not in the subsequent telling. For the television soap it was necessary to transcribe the events in the form output by GADIN. Using this style of writing is an obvious limitation of the narrative quality, but in this evaluation only the core components of the narrative and its structure are required to be compared and thus this is sufficient. In both versions it was essential not to impose any reasoning on the characters in the wording.

The human-authored narrative used in this experiment has been taken from a television soap opera. The GADIN-authored narrative was not specially selected but was the first to be generated from the given storyworld state. It was ended when at the same length as the television soap narrative, as although the narrative generation would continue indefinitely from this point the later narrative content was not relevant to this comparison.

4 The Narratives

The narratives shown to the participants are given here.

STORY 1: Jane and Tom are in a relationship. Jane becomes pregnant. She decides not to keep the baby. Tom proposes but Jane rejects him, ending the

relationship. Nick and Sally go to the shop. Nick flirts with Tom. Sally flirts with Tom. Tom must choose whether to partner Sally or Nick. He decides to go out with Sally. Nick no longer fancies Tom. Nick starts to go out with Rich. Sally wants to start an affair with Rich but he chooses not to. Tom expresses his disapproval of Richs relationship with Nick, but Rich decides to continue the relationship anyway. Jane and Sally stop liking one another. Tom starts an affair with Jane.

STORY 2: Tom and Jane are in a relationship. Rich and Sally flirt with each other, and start going out. Rich wants to start an affair with Tom but he chooses not to. Rich splits up with Sally. Tom stops liking Rich. Rich and Sally start going out with each other again. Rich ends the relationship with Sally. Rich starts a relationship with Nick. Tom expresses his disapproval of this relationship but Rich ignores him. Rich and Tom start having an affair. Tom proposes to Jane and she accepts. Rich tells Jane about his affair with Tom. Jane ends her relationship with Tom. Tom and Rich start a relationship.¹

5 The Results

The survey was divided into two groups of participants. One consisted of those who regularly view soaps (although not the soap in question). These participants were targeted through posts on 4 English soap forums.² Given the focus on English soaps throughout it was important that these forums were used only by an English audience, as the style and content of soaps does vary to some extent between nationalities. The second group contained those who regularly play computer games and are thus more familiar with computer-based stories. For this the survey was posted on 2 English games forums and 1 international.³

It is generally considered that a result which has a more than 5% chance of occurring if random choices are made has successfully passed the Turing test. The results of this test were as follows:

- 42 soap viewers participated in the survey. Of these, 24 – or 57.1% – thought that story 1 was the one which had been written by a computer. This will occur 22% of the time with random choices.
- 85 games players took part in the survey. In this case 49 participants, or 57.6%, thought that story 1 was written by a computer. This will occur 10% of the time if random choices are made.
- Overall there were 127 surveys completed. 57.5% thought that story 1 had been written by a computer. This will occur more than 5% of the time if random choices are made.

¹ Story 1 was written by the GADIN system and Story 2 is from Hollyoaks.

² The soap forums utilised were at: www.soapforum.co.uk; www.digitalspy.co.uk; www.soapboards.co.uk; www.tvforum.co.uk

³ The games forums used were: www.enemydown.co.uk; www.na-clan.co.uk; www.wcreplays.com

From these results it can be concluded that the GADIN system is capable of generating a narrative which is not discernibly different from an outline of a story occurring in a television soap opera.

6 Conclusions

We have shown that this story generation method generates narratives which are indistinguishable from those in standard soap operas.

In this experiment only a single storyline has been evaluated. The infinite nature of soaps means that there will be a number of such storylines. However since this was chosen randomly the implication is that other storylines are likely to also be indistinguishable from those in television soaps.

This experiment is focused on the very generalised story development aspect and on the creation of the core component of dramatic interest and the underlying structure of the narrative. This will inevitably affect users opinions of the system and their experience but hopefully the fundamentals of the narrative maintain sufficient dramatic interest for this to not too adversely affect the results.

It would potentially be possible to use GADIN in collaboration with script writers to create televised soaps. Future experiments could involve the GADIN story outline being written as a script which would be performed by human actors. This could then be compared to the television soap script being acted by the same characters. This would compare the whole viewing experience as opposed to the core components of the narrative as was the focus in this experiment.

7 Acknowledgements

Thanks are due to everyone who participated in the survey, and in particular to Christian Wood who coordinated the games survey.

References

1. Meehan, J.R.: *The Metanovel: Writing Stories by Computer*. Garland Publishing, Inc., New York and London (1980)
2. Bringsjord, S., Ferrucci, D.A.: *Artificial Intelligence and Literary Creativity: Inside the Mind of BRUTUS, a Storytelling Machine*. Lawrence Erlbaum Associates, Mahwah, NJ (2000)
3. Lebowitz, M.: *Planning stories*. In: *Ninth Annual Conference of the Cognitive Science Society*, Seattle WA (1987)
4. Turner, S.R.: *The Creative Process: A computer model of storytelling and creativity*. Lawrence Erlbaum Associates (1994)
5. Barber, H., Kudenko, D.: *Dynamic generation of dilemma-based interactive narratives*. In: *AIIDE'07*, Stanford, California (2007)
6. Barber, H., Kudenko, D.: *Generation of dilemma-based interactive narratives with a changeable story goal*. In: *2nd International Conference on Intelligent Technologies for Interactive Entertainment*, Cancun, Mexico (2008)
7. Turing, A.: *Computing machinery and intelligence*. *Mind* **59** (1950)