

The Role of Gender And Age on User Preferences in Narrative Experiences

Michael Garber-Barron and Mei Si

Cognitive Science Department, Rensselaer Polytechnic Institute, Troy, New York
{barrom2, sim}@rpi.edu

Abstract. Storytelling is interactive - storytellers seldom recite the same content, instead, they adapt the narrative to their audience. Storytellers both introduce new topics from time to time, and elaborate on the topics the audience has previously shown interest on. In our previous work, we created a digital storytelling system that simulates this process. The goal of this work is to investigate what is a good balance between the novel topics and the audience's existing interests. We conducted an empirical evaluation using our digital storytelling system. Distinct preferences were observed in relation to both age and gender, with women in general enjoying a more gradual shift in the presentation of novel material.

1 Introduction

Storytelling is an important aspect of our lives. It is both a major form of entertainment, and allows us to obtain new knowledge and communicate beliefs. Unlike playing a digital recording, storytelling when happening face to face is interactive. People adapt how the story is presented to their audience, by answering questions, elaborating on topics the audience has shown interest in before, and introducing new topics from time to time to attract attention. In our previous work [1, 2], we created an automated storytelling system that attempts to engage the audience in a similar way as a human narrator. While telling the story, the system periodically estimates the user's interest and preference based on his/her questions and comments. The system biases its subsequent storytelling towards content related to the user's inferred interests while accounting for the introduction of novel topics and maintaining topic consistency.

There is a variety of work on user modeling in interactive narratives [3–5]. Our system addresses a similar goal, however, in a different context. The listeners have no agency in the story, and they may not identify with a character in the story. Therefore, the listener's models cannot be created based on the models of the characters.

We found several interesting works that concentrate on inferring and adapting storytelling to the audience's experiences. The Papous system [6] proposes to adapt storytelling based on the listener's emotional experiences. The Virtual Storyteller [7] uses autonomous agents to interact with each other and present a story to the audience. The interaction is steered by the positive or negative feedback from the audience members. In IRST [8], the story is developed around the notion of interest. A fixed story is presented to the user, but optional information can be presented to add details about characters or events. As the user views the scenes they can explicitly demonstrate their

interest through an adjustable slider, which is used to indicate which characters interest the user. Suspenser [9] manipulates the presentation and order of a story’s content to optimize suspense.

Primarily, our approach is inspired by IRST. We also try to adapt a story around the interests and preferences of the user. In this work we go further by trying to balance between providing new content that is more along the lines of where the user has already shown interest and presenting novel material that can trigger new interests from the user.

In our previous work [2] we evaluated the efficacy of our storytelling system. In that study, younger individuals (40 and below) identified stories with additional novel content as being more engaging. In this work, we want to further investigate what type of novelty is preferred and whether there are any gender or age preferences. In particular, we compared two approaches for introducing novel content, one that gradually introduces novel content and accounts for the user’s interests, and the other which introduces varied and novel topics as much as possible without incorporating the user’s interests. In the next sections, we will briefly review our storytelling system (see [1, 2] for more details), and then present the empirical study.

2 Story Representation

We represent a story as a set of connected events with preconditions which define what can and cannot happen before the event. Each event is either a physical event that progresses the story or an informational event that describes the characters’ background, beliefs, or motivations.

We modeled a Chinese fantasy story – “The Painted Skin” [10] for this study. The story revolves around a man named Wang, his family, and their interactions with a demon disguised as a girl who eventually murders Wang.

For enabling our system to *understand* the progress of the story and the user’s interests, the events described by the narrator and the statements the user can say to the narrator are tagged using a set of predefined labels. For example, the event “Wang hoped that the priest could find some way of protecting his family” is labeled with [Wife, Fear, Love, Wang].

2.1 User Interest Representation

In our system, the user’s interests are represented as a set of interest *profiles*. Each *profile* represents a possible set of the user’s inferred interest towards the various labels used for modeling the story, for example: [Wife: 0.4, Fear: 0.3, Love: 0.2, Wang: 0.1]. These profiles are randomly initialized, and get updated as the user interacts with the system. When planning for future content, these profiles are combined in order to represent the user’s interests (see [1] for further details).

2.2 Storytelling

The storytelling system plans for the presentation of the future narrative each time the user has made a comment or asked a question, at which point the user’s interest profiles are also updated. For storytelling, we want the user to hear more about the type of content they’ve already shown interest towards as well as to discover new interests. To realize this goal, first, the events that can immediately proceed in the story, in other

words, whose preconditions are satisfied are selected. Then, an interest score is generated for each event. This score is an indicator for not only how well the event itself will interest the user, but also how well the future story trajectory which is likely to follow this event will interest the user. This score is a weighted combination of two factors: Novelty and Topic Consistency.

Novelty measures how similar/different the labels in the trajectory are from the user's current interest profiles. The smaller the similarity is, e.g. the trajectory contains a label which has been rarely utilized or the user has showed no interest towards previously, the more novel the trajectory is (see [1] for further details.)

For this study, we compared two approaches for incorporating novelty in storytelling: gradual novelty and immediate novelty. For achieving gradual novelty, we want the storytelling algorithm to gradually transition among topics. Gradual Novelty is modeled as $\left|1 - \frac{1}{\text{distance from current step}} - \text{similarity}\right|$. Thus, for the same amount of similarity/difference, the further away the event is, the more (gradually) novel it is considered. Using immediate novelty, the storytelling algorithm continually tries to find the most varied and differing content, and novelty is simply defined as, $|1 - \text{similarity}|$.

Topic Consistency measures whether events in succession are similar in terms of the labels they address, independent from whether the trajectory is novel to the user (see [1] for further details.) This factor is incorporated to prevent the system from oscillating between topics. This procedure remains the same for both models of novelty. When used together with the gradual model of novelty, the storytelling system will switch topics less abruptly compared to when used with the immediate model of novelty.

2.3 Experimental Design

A between subjects design was used with three independent variables: gender, age and the type of novelty. We partitioned subjects' ages into three groups: group 1, 30 and below; group 2, 30 to 40; group 3, 40 and older. As previously described, the type of novelty has two levels: gradual novelty and immediate novelty. In both conditions, the topic consistency aspect of the algorithm was untouched, so that the story presented to the user is intended to always be reasonable.

For this evaluation, instead of having the subjects directly interact with the system, we manually simulated users interacting with the system, and showed the histories of these interactions to the subjects. The subjects were asked about what they think of the (simulated) users' experiences of interacting with the system.

We simulated a positive user, which is modeled as having a preference for positive events and relationships such as care, love, protection, etc. We simulated this user so as to clearly distinguish the user's interests from the mood of the overall story which has a darker tone.

After reading the interaction between the simulated user and the system, the participants were asked to rate the following questions using a 7 point Likert scale with 7 being the strongest.

Questions: [*INT*]: "To what degree did the story seem interesting to the user?", [*ENG*]: "Given the user's responses to the narrator, to what degree do you think the story was engaging for the user because it was consistent with the user's personality?", [*APP*]: "Given the user's responses to the narrator, to what degree do you think the

elements of the story were appropriate for the user’s personality?”, [NOV]: “Given the user’s personality and responses to the narrator, how novel did the elements of the story presented feel to the user?”, [RAND]: “Given the user’s personality and responses to the narrator, how random did the elements of the story presented feel to the user?”, [PLOT_APP]: “How well did the plot elements seem appropriate to the user’s personality?”, [APP_YOU]: “How well did the plot elements seem appropriate to your personality?”, [ENG_USR]: “How engaging did the user find the story?”.

These questions are intended to examine the subjects’ impressions on engagement ([INT, ENG, ENG_USR]), the appropriateness of the content ([APP, PLOT_APP, APP_YOU]), and how novel or random the content seems ([NOV], [RAND]).

2.4 Procedure

We conducted this experiment as an anonymous survey on Amazon’s Mechanical Turk. A total of 185 subjects participated, 80 of them were women and 105 were men. The subjects’ ages range from 18 to 66. There were 115 in group 1, 33 in group 2 and 37 subjects in group 3, and 3 subjects did not specify their gender and were excluded from the analysis.

Each subject was randomly assigned to a novelty condition. In each condition, the subjects read an interaction history consisting of eight segments that corresponded to their condition. In each segment, as shown in Figure 1, a picture that represented the development of the story is displayed, followed by the narrator’s description of the story, the choices the user could pick from, the choice picked by the user and the narrator’s response in sequence. At the end, the subjects were asked to rate the items listed in the previous section.



Story described by Narrator:

Her parents had sold her to a cruel master and she had just fled... Wang offered to let her stay at his residence.

Options the user can direct to the narrator:

1. Whoa, that is really nice of Wang to do for her.
2. Wang is way too trusting. Nowadays you would be more cautious, she could be a serial killer.

Option chosen by user: Statement 1

Narrator’s response: Indeed, it does appear that way.

Fig. 1: Text Version

2.5 Data Analysis and Results

Because participants were asked to evaluate 8 questions in this study, we first performed a Pearson’s bivariate correlation analysis, which indicates a high correlation among the variables ($p < .01$ for all the comparisons except *[RAND]* which had a relatively neutral relationship with most of the other variables). Due to the significant covariance, we decided to perform a three-way MANOVA test instead of individual ANOVA tests. The Roy’s Largest Root test was used for significance testing.

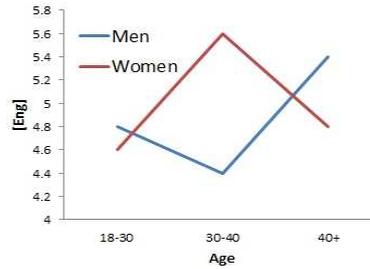


Fig. 2: Age and Gender Interaction

The MANOVA test revealed a significant interaction effect of gender and age ($p < .05$). The follow up ANOVA tests on each independent variable only revealed significant interaction effects between gender and age groups on *[ENG]* – the participant’s view of the user’s level of engagement with the story ($p < .05$).

As seen in Figure 2, the relationship of age and gender correlate to participants’ perception of the user’s engagement in each story. Male and female participants demonstrated different preferences with men in the 30-40 age bracket rating the *[ENG]* question the lowest (Mean=4.6) comparing to other age groups, and women rating it the highest comparing to other age groups (Mean=5.6).

	<i>[INT]</i>	<i>[ENG]</i>	<i>[APP]</i>	<i>[PLOT_APP]</i>	<i>[APP_YOU]</i>	<i>[ENG_USR]</i>	MEAN
Immediate	5.4	4.9	4.8	4.8	4.3	5.4	5.9
Gradual	5.7	4.9	5.1	4.7	4.9	5.4	6.1

Table 1: Female Ratings

Women appeared to have a greater preference towards the gradual novelty condition. Table 1 lists the average rating of female subjects for each question. We performed the same MANOVA test with only female subjects. This test revealed a marginally significant main effect of novelty type ($p = .078$). No other significant effects were observed.

3 Discussion and Future Work

The large difference in ratings in the 30-40 age range for women and men suggests that there are significant population trends. This result suggests that while we seek to

model each individual's preferences, accounting for each user's demographic information would allow for a better initial estimate of what should be presented. Further, the differences in preferences for women (Table 1), and the marginally significant effect suggest that novelty type may be more relevant for women. Nevertheless, our primary evaluation of novelty type did not reveal a significant difference between the two conditions. However, given the short time frame (Mean=4.5 minutes), it may be the case that there is simply not enough content for participants to observe the gradual novel condition as clearly distinctive from the immediate novel condition.

Ongoing work is being directed towards several goals including repeat the experiment with the subjects directly interacting with the storytelling system, and modeling a longer version of the story with more details. We are also interested at exploring automating the initialization of the profiling component.

4 Conclusion

In this work, we evaluated the subjects' preferences over two different approaches for incorporating novel topics in storytelling using our automated storytelling system. A marginally significant effect was observed for women. Women subjects seem to enjoy stories that switch topics gradually more than those that switch abruptly. Further, significant interaction effects were observed between gender and age. Future work has been planned on further investigating these differences and improving the initialization process of the storytelling system.

References

1. Garber-Barron, M., Si, M.: Towards Interest and Engagement: A Framework For Adaptive Storytelling. In: Proceedings of the 5th Workshop on Intelligent Narrative Technologies Collocated with AIIDE. (2012)
2. Garber-Barron, M., Si, M.: Adaptive Storytelling Through User Understanding. In: Proceedings of AIIDE. (2013)
3. Si, M., Marsella, S., Pynadath, D.: Proactive Authoring for Interactive Drama: An Authors Assistant. In: Proceedings of IVA. (2007) 225–237
4. Yu, H., Riedl, M.: A Sequential Recommendation Approach For Interactive Personalized Story Generation. In: Proceedings of AAMAS. (2012) 71–78
5. Thue, D., Bulitko, V.: Procedural Game Adaptation: Framing Experience Management as Changing an MDP. In: Proceedings of AIIDE. (2012)
6. Silva, A., Raimundo, G., Paiva, A.: Tell Me That Bit Again... Bringing Interactivity To A Virtual Storyteller. In: Proceedings of the International Conference on Virtual Storytelling. (2003) 146–154
7. Theune, M., Faas, S., Heylen, D.K.J., Nijholt, A.: The Virtual Storyteller: Story Creation By Intelligent Agents. In: Proceedings of the Technologies for Interactive Digital Storytelling and Entertainment (TIDSE) Conference. (March 2003) 204–215
8. Nakasone, A., Prendinger, H., Ishizuka, M.: ISRST: Generating Interesting Multimedia Stories on the Web. *Applied Artificial Intelligence* **23**(7) (2009) 633–679
9. Yun-Gyung, C., Michael, Y.: Narrative Generation For Suspense: Modeling And Evaluation. In: Proceedings of ICIDS, Springer (2008) 144–155
10. Sung-Ling, P.: *The Painted Skin*. In: *Strange Stories from A Chinese Studio*. University Press of the Pacific (1916) 47–51