Adaptive Storytelling Through User Understanding

Michael Garber-Barron
Cognitive Science Department
Rensselaer Polytechnic Institute
Barrom2@rpi.edu

Mei Si
Cognitive Science Department
Rensselaer Polytechnic Institute
Sim@rpi.edu

Abstract

Storytelling, when happening face to face, is a highly interactive process. A good storyteller attracts the audience by continuously observing their responses and adjusting the storytelling accordingly. The goal of this project is to simulate this process in digital storytellers. We created an automatic storytelling system that periodically estimates the user's preferences and adapts the content in the subsequent storytelling by balancing novelty and topic consistency. We have performed an empirical evaluation on the effectiveness of our approach. The results indicate that gender and age play important roles in affecting one's subjective experience. For younger subjects, stories with mixed amount of novelty and topic consistency are more preferred while for older subjects, larger amounts of variation are preferred. Additionally, in general, women enjoyed the stories more than men.

Introduction

Storytelling is an integrated part of people's lives. The art of storytelling has been studied since the time of Aristotle. With the rapid development of technologies, more media are available for telling stories, e.g. videos, audio books and animated books. However, a problem with non-interactive media is that they cut off the feedback loop between the audience and the storyteller, which is an important component in any performing art. Though the audience cannot affect what happens in the story, storytelling when happening face to face is a highly interactive process. Good storytellers deliberately try to understand their audience's interests and adjust the amount of details or the types of information they provide to the audience.

In this work, we are aiming to simulate human storytelling behaviors in a digital storyteller. More specifically, we want to be able to estimate the user's interests based on their questions and comments, and adapt the subsequent storytelling to the estimated user interests. Reexamining familiar and related material can develop and deepen a listener's preexisting interests (Hidi and Renninger 2006; Arnone et al. 2011). On the other hand, new content can attract a user, leading them to explore different content which can develop alternate interests (Loewenstein 1994), and ultimately make the user feel more immersed and engaged (Wade 1992). In this work, we try to balance novel content with the listener's previous interests when presenting the story.

Empirical evaluations have been conducted on the effectiveness of engaging the user using our digital storyteller. The results indicate that gender and age play important roles in affecting one's subjective experience; for younger subjects, stories with mixed amount of novelty and topic consistency are more preferred while for older subjects, larger amounts of variation are preferred; and the stories are viewed as more enjoyable for women than for men.

Related Work

Digital Storytelling Systems

There is a large body of work in the area of user modeling and interactive narratives in which the user plays a role in the story and interact with other characters controlled by an AI system (Mateas and Stern 2003; Si, Marsella, and Pynadath 2009; El-Nasr 2007; Pizzi et al. 2007; Lee, Mott, and Lester 2010; Riedl, Saretto, and Young 2003; Thue et al. 2007; Rowe et al. 2010; Nelson and Mateas 2005; Weyhrauch 1997; Sharma et al. 2010). This work is towards a different direction. The user is a listener of a story and has no agency in the story world. Instead of taking a specific role in a story, the user's interest may be very different from the characters' interests and may span across several characters.

There are several interesting works along the direction of dynamically telling a story based on the users' feedbacks. The Papous system (Silva, Raimundo, and Paiva 2003) proposes to adapt storytelling based on the listener's emotional experiences. The Virtual Storyteller (Theune et al. 2003) 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. Similar to Papous, the storybook system (SSAU) (Tanenbaum and Tomizu 2008) explored adapting storytelling to the user's mood. Different from all other systems, instead of adjusting the content to be presented to the user, SSAU adjusts the audio and visual presentation, such as darker lighting to signify a more melancholic environment.

In IRST (Nakasone, Prendinger, and Ishizuka 2009), 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.

Our work 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 the user more material along the lines of where he/she has already shown interest and presenting materials that can potentially trigger new interest from the user. In addition, we also developed algorithms for automatically inferring the user's interest from his/her questions and comments.

Curiosity and Interest of the Audience

From a psychological perspective, curiosity have been viewed purely in terms of a need for environmental variation, and novel stimuli (Loewenstein 1994). This later extended to include a more directed view in that the interest and curiosity of an individual can be comprised of a desire for particular knowledge, or information (Berlyne 1954). In general, there is an acceptance of associating interest in relation to exploring information that is novel, or causes uncertainty (Reeve and Nix 1997; Loewenstein 1994; Day 1982).

More recently, Arnone et al. has proposed a functional definition where interest and curiosity are viewed as mediating factors towards the development of user engagement (Arnone et al. 2011). As our preference for particular content develops, we examine related material and information pertaining to those interests. Certain content that initially sparks the user's curiosity can be further explored leading to a deeper sense of interest (Hidi and Renninger 2006). This in turn can lead to a greater degree of immersion and engagement with the content. In contrast, failing to effectively address the user's interests may lead to a sense of withdrawal and boredom (Arnone et al. 2011).

Example Domain

Our example domain is a Chinese fantasy story – "The Painted Skin" (Sung-Ling 1916). The story starts as a young man – Wang – meets an attractive young lady on the road. She was actually a demon, and she approached young men like Wang in order to obtain energy from them. Wang will eventually die when his energy is drained. However, Wang did not know this, and quite enjoyed the demon's company until one day he met a priest on the street who warned him about the demon. The story continues with the demon accidentally revealing her true identity to Wang, killing Wang, and the priest helping Wang's family to defeat the demon and resurrect Wang.

We created two versions of the story. One is a 3D interactive environment created in the Unity game engine. The other is a text version with associated pictures. For the purpose of conducting our evaluation as an online survey, the text version was utilized. In both representations, as the story progresses, the user selects various opinionated statements to direct towards the storyteller. A snapshot of how we present the story and the user's options is shown in Figure 1.

Figure 1: User Interface for Evaluation

Automated Storytelling System

We want to engage the user by letting him/her hearing more about the type of content they've already shown interest towards as well as to discover new interests. To realize this goal, our system is comprised of two components: an interest profiling module that is used to infer what aspects of the story interest the user and an event selection module which tells the story. The event selection module determines the content to present to the user after taking into account his/her interests and the progression of the story.

Story Representation

We represent a story as a set of connected events. Each event is either a phyical event that progresses the story or an informational event that describes the characters' background, beliefs, or motivations. Both kinds of events are mapped to a line of text that is displayed or spoken by the narrator. Similar to other interactive narrative work that use a graph of plot-points (Nelson and Mateas 2005; Weyhrauch 1997), to ensure that the story can only progresses in a logical and coherent manner, events are modeled with preconditions, which are defined as a combination of events that should or should not have already happened in the story. Further, while both physical and informational events are treated algorithmically the same, in the case of physical events, they are structured to progress in the same relative order. In contrast, informational events can appear variably and optionally in terms of presentation. For example, in our story, both the physical event of Wang meeting the demon and his meeting of the priest will always occur in the story, and in that order of occurrence. However, an informational event such as Wang discussing his attitude about the demon may optionally occur, but, it can only appear after Wang's meeting with the demon.

For enabling our system to *understand* the progression 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 with 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]. The labels are determined by the designer of the interactive story. For modeling the "Painted Skin" story, twenty-four labels were used.

Interest Profiles

In our system, the user's interests are represented as a vector of interest *profiles*. Each *profile* is a distinct representation of the user's inferred interest towards the various labels the designer used for modeling the story. More specifically, each profile includes a vector of values that specify the user's inferred interest towards each of the labels, for example: [Wife: 0.4, Fear: 0.3, Love: 0.2, Wang: 0.1].

A *profile* can be thought of as a distribution of the user's disposition towards the labels and the associated events. The objective of developing multiple *profiles* is to distinguish between possible focuses of interest based on limited input. For example, as the story is progressing, the user has demonstrated a high preference towards both content labeled with "Love" and "Fear". However, the user has not shown an interest towards contents that contain both labels. To account for this difference, we need at least two different profiles. One profile has a large preference value for "Love", and the other has a large preference value for "Fear". Then, when the system is trying to determine new content to present, events with the dual labels of "Love" and "Fear" will be considered less interesting to the user over events that only contain one of the labels. The more profiles there are, the more fine distinctions the system can make about the user's interests. Of course, this brings up the question of the necessary number of different profiles. Our proposed algorithms for event selection and profile update will operate in the same fashion as long as there is more than one profile. The exact number of profiles is determined by the designer. For a short story like this, we set the number as two.

The initial values of the labels in each profile can either be set by the designer or by default configured randomly, which indicates that the system has no prior knowledge of the user's interests. These profiles are updated as the user makes choices in the storytelling process, and the system gets additional information about the user's preferences. The more the user interacts with the system, the more likely the profiles will converge and reflect the user's true interests.

As the user selects statements, the system calculates confidence/preference values associated with each of the user profiles. These values indicate how closely each profile is representative of the user's interests so far. Similar to the values of the labels in each profile, the initial preference values can either be randomly assigned or set by the designer. These values are adjusted as the user makes more choices. These preference values determine the impact that each profile will have during the **Event Selection**.

Profile Update Procedure

The profiles are updated as the story progresses based on the choices of the user. When the user selects a statement it is compared to each profile, and used to update the most *similar* profile, which is determined by comparing the labels of the statement with the label values in each profile. Specifically, the Euclidean distance between the two profile arrays are used, and the profile with the smallest distance to the statement is selected as the most similar profile. The score of this profile is then incremented to indicate that the system views the profile as relatively more representative of the user's interests. After this, the label values within this profile are updated by calculating the average occurrence of each label relative to the number of times the profile has been chosen by the system (see (Garber-Barron and Si 2012) for details). This way, labels that are persistently chosen are valued more highly.

Event Selection

Algorithm 1 details how the system selects the next informational or physical event to be presented to the user. An overall score is generated for each event that can immediately proceed (judged by their preconditions) in the story. These scores are estimates for how the events will interest the user immediately, as well as lead to future content and events that will continue to engage the user.

In order to generate these scores and determine which event to present to the user, Algorithm 1 is applied to determine what to select. The scores for each event are calculated by using a weighted sum of the user's inferred preference towards each interest profile along with the profile scores and consistency values the system had generated in Algorithm 2 (see Algorithm 1, Line 6 and Line 8) which consider both the topic's novelty and consistency.

Once a score for each event is calculated, the system chooses the event with the highest calculated score to present to the user. The overall score for each event (Algorithm 1, Line 8) represents a combined measure of how well each event fits the profiles of the user (their preferences and interests), the consistency of the topics of interest that can appear in relation to that event, and the degree of novel and new labels that can potentially stem from it. In doing so, we are acknowledging the perceived importance for the system of each profile, and the notions of interest, novelty, and the consistency of those interests. Algorithm 2 details the process of how the system evaluates each individual event. Events are scored by determining the event trajectories that can follow them in the story – a sequence of events through the story that best account for the user's interests. This is done by recursively exploring the various branching paths through the story stemming from each possible event.

A trajectory generates a cumulative score that accounts for how likely the system believes each of the possible sequences of events will engage the user while accounting for novelty and topic consistency. Both the original event and

Algorithm 1 EventSelection($EventOptions$, $Profiles$)

1: #$AvgStoryPath$: The average length to traverse through a story
2: #$Preferences$: An array containing the inferred user preference towards each profile
3: $Preferences$ = **Normalize**($Preferences$)
4: #$ProfileScore, Consistency$: Arrays with all their elements initialized to zero
5: **for each** $Event$ in $EventOptions$ **do**
6: $\quad ProfileScore, Consistency$ = **EventEvaluation**($Event, Event, ProfileScore, Consistency, Profiles$)
7: $\quad$ **for each** $Profile$ in $Profiles$ **do**
8: $\quad\quad EventScore$+= $Preferences[Profile] * ProfileScore[Profile] * ((Consistency[Profile])/AvgStoryPath)$
9: $\quad$ **if** $EventScore > PreferredScore$ **then**
10: $\quad\quad PreferredEvent = Event$
11: $\quad\quad PreferredScore = EventScore$
12: **Return** $PreferredEvent$

each of the subsequent events that can follow are compared to the user's profiles and summed in an attempt to contribute to the original event's overall score - an estimate of its perceived importance towards leading the user along an engaging story.

Importantly, the scores generated along these event trajectories are comprised of a sum of sub-scores that are equal to the number of profiles. This allows the system to evaluate the importance of an event to the user in relation to how well it matches each profile, as well as all events that can causally follow. By doing this, the system incorporates the relevance of the immediate event and future content that can arise from it in relation to all profiles of the user.

Consistency To prevent the oscillation between topics, the current system emphasizes trajectories that fully explore one topic – events in succession that are similar to one associated profile and their corresponding label values – followed in succession by alternative topics related to other profiles (see Algorithm 2, Line 13). Events along a trajectory that are similar to the same profile in succession cause an increase in the consistency value and thus the overall score. The generated value is later recombined with the novelty and interest values when they are used to calculate the original event's score (Algorithm 1, Line 8).

Novelty When calculating the score of an event, novelty is applied by increasing the impact of labels that occur less frequently in the selection process. This is achieved by exploring possible future trajectories of events from the original and giving a greater weight in the scoring process to less frequently occurring labels. As can be seen in Algorithm 2, Line 11, as future events are examined farther ahead in the story, those that have fewer labels and topic similarity will have an increasingly high positive impact on the overall score of the event.

Evaluation

To evaluate how well our system can engage its users, a preliminary evaluation was conducted. We want to investigate not only the overall hypothesis of whether the system makes the storytelling process more engaging, but also how the subcomponents of the system, i.e. keeping topic consistency, and incorporating novel content affect the user's experiences.

Algorithm 2 EventEvaluation($OriginalEvent$, $NextEvent$, $ProfileConsistency$, $ProfileScore$, $Profiles$)

1: #$ProfileConsistency$: An array containing the average number of events that are traveled in succession pertaining to each profile
2: #$ProfileScore$: An array containing the calculated scores for $OriginalEvent$ as it pertains to each profile
3: $FutureEvents$ = **GetFutureEvents**($NextEvent$)
4: #$MaxDepth$: The maximum depth that events can still be evaluated
5: **if** $FutureEvents$ is Empty or at Max Depth **then**
6: $\quad$ **Return** $ProfileConsistency, ProfileScore$
7: **for each** $Event$ in $FutureEvents$ **do**
8: $\quad Profile$ = **GetClosest**($Profiles, Event$)
9: $\quad PreviousProfile$ = **GetClosest**($Profiles, NextEvent$)
10: $\quad Distance$ = **GetDistance**($Profile, Event$)
11: $\quad ProfileScore[Profile]$+= $\left|1 - \frac{LengthOfPath}{MaxPathLength} - Distance\right|$
12: $\quad$ **if** $PreviousProfile = Profile$ **then**
13: $\quad\quad ProfileConsistency[Profile]$+= 1
14: $\quad$ **EventEvaluation**($OriginalEvent, Event, ProfileConsistency, ProfileScore, Profile$)
15: **Return** $ProfileConsistency, ProfileScore$
16:
17: #**GetFutureEvents**($Event$): Returns a list of future events that can directly follow $Event$

Because of the nature of interactive storytelling, each user of the system will hear a slightly different story based on their responses, and the system's estimation of what content would best interest them. As such, these variations of the story, though generated using the same algorithms, can lead to fairly variable content in terms of informational events. Instead, for this initial evaluation, we want to have a very controlled study. Instead of having the subjects directly interact with the system, we manually simulated two different types of 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. Out of the huge collection of possible user profiles, we picked two prototypical ones – one presenting a positive user and the other presenting a negative user. They are described in the next section.

Experimental Design

A 2 by 2 between subjects design was used. The first independent variable is storytelling style, i.e. whether the system is responding to the user's expressed interests while telling the story. This independent variable has two levels: non-random (generating the story based on the estimated user interests) and random (generating the story based on a random and fixed user profile). We expect the non-random condition to be judged as more engaging because in this condition novelty and topic consistency are more balanced. In the random condition, the system will introduce a large degree of variability as it does not algorithmically account for or balance between novelty and topic consistency.

The second independent variable is the type of the user that interacted with the system.We want to evaluate whether our proposed storytelling algorithms work for different types of users. In this preliminary evaluation, we test on two very distinct user profiles. The positive user is modeled as having a preference for positive events and relationships such as care, love, protection, etc. He/she will be more likely to select statements/questions such as: "Whoa, that is really nice of Wang to do for her." The negative user is modeled as having a preference for negative events and relationships such as violence, deceit, lust, etc. He/she will be more likely to select statements/questions such as: "Wang broke his promise pretty quickly, huh?" We do not have a specific hypothesis regarding the effect of the user's interaction style on the story's engagement.

The dependent variables are the subjects' rating of the following 5 items using a 7 points Likert scale:

- $[ENG\_USR]$: Based on the user's responses how engaging do you think the story was for the user?
- $[REP\_INT]$: What degree do you think the story was representative of the user's interests?
- $[ENG\_CON]$: What degree do you think the story was engaging for the user because it was consistent with the user's interest?
- $[ENG\_NOV]$: What degree do you think the story was engaging for the user because of the novelty of the content presented?
- $[ENG\_SEL]$: How engaging did you find the story?

Procedure

We conducted this experiment as an anonymous survey on Amazon's Mechanical Turk. A total of 252 subjects participated in this study over four days' time. 117 of them are women and 132 are men. The subjects' ages range from 19 to 61. For data analysis purpose, we partitioned subjects' ages into four groups: 30 and below; 30 to 40; 40 to 50; and 50 and older. There are 82, 91, 39 and 39 subjects in each of the groups respectively, and 3 subjects did not specify their genders.

Each subject was randomly assigned to one of the four experimental conditions. In each condition, the subjects read an interaction history consisting of 8 segments that were generated using the corresponding story generation algorithms and user profiles. In each segment, as shown in Figure 1, a picture that represents the development of the story is displayed, followed by the narrator's description of the story, the choices the user can pick from, the choice picked by the user and the narrator's response in sequence. The subjects are informed that the user's choice not only affects the narrator's response, but also the subsequent descriptions of the story.

At the end, the subjects were asked to rate the 5 items listed in the previous section. In addition, we presented to the subjects 15 of the labels used in this story. They were asked to rate how well each of the labels represented the user's interest using a 7 point Likert scale. We also collected data on the subjects' gender, age, and location.

Data Analysis and Results

There are five questions in the questionnaire. We first performed a Pearson's bivariate correlation analysis, which indicates a high correlation among their answers ($p<.01$ for all the comparisons). Therefore, we decided to perform a MANOVA test instead of individual ANOVA tests. Further, we suspected that in addition to the independent variables, the subjects' gender and age may influence their judgments. Our MANOVA test has four independent variables which are the storytelling style, user interaction type, gender and age. The Roy's Largest Root test was used for significance testing.

The MANOVA test revealed a significant main effect of gender ($p<.01$), female subjects consistently rated each question higher than male subjects (M=0.21). Further, a marginally significant main effect of age ($p<.01$), a significant interaction effect between storytelling style and age ($p<.01$), a significant interaction effect between gender and age ($p<.01$), a marginally significant interaction effect between user type and age ($p<.01$), and a marginally significant interaction effect between storytelling style, gender and age ($p<.01$) were observed.

This analysis is performed using SPSS, and the covariances among the dependent variables (the 5 questions) has been taken into account.

Further, the tests of between-subjects effects indicates a significant interaction effect between storytelling style and age on $[ENG\_NOV]$ ($p<.01$) and $[ENG\_SEL]$ ($p<.05$) and a marginally significant interaction effect between gender and user type on $[ENG\_SEL]$ ($p<.05$). For controlling the overall type I error, we adjust α to $.05/5 = .01$ for these tests. Figure 2 show the interaction effects.

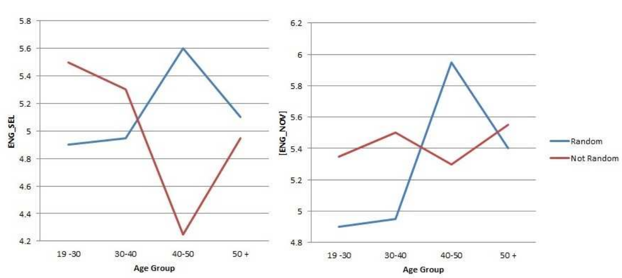

Figure 2: Interaction Effect Between Storytelling Style and Age on $[ENG\_SEL]$ and $[ENG\_NOV]$

We can observe a common phenomenon that the age groups below 40 and above 40 exhibit different, fairly contrasting behavioral patterns. As shown in Figure 2, there are different trends of how the storytelling style affects the subjects' own experiences of engagement $[ENG\_SEL]$ and their expectation of the story being engaging to the user (listener) because of the novel content integrated into the story

$[ENG\_NOV]$. For $[ENG\_SEL]$, when the system considers the user's input – the non-random condition, as age goes up, the subjects found the story less engaging. The opposite happens when the storytelling does not account for the user's input – the random condition. As the subjects' ages increase, so did their sense of engagement with the random condition. For $[ENG\_NOV]$, as the subjects' ages increase, the subjects tend to find the story more engaging because of the novel content in it. Not surprisingly, when the story was generated without considering the user's input, this trend is far greater.

In both figures in Figure 2, the subjects in the younger age groups seem to behave more consistent with our primary hypothesis – we hypothesized that the story will be more engaging to the user if the user's input is taken into consideration. To test this effect, we ran the same MANOVA again with only the first two age groups, that is subjects who are younger than 40. This time, the storytelling style becomes a significant factor that affects $[ENG\_NOV]$ ($p<0.05$). Though it does not affect other variables significantly, because of the very high correlation among the dependent variables, we take this as a positive sign of the storytelling style affecting how the user experiences the story. We did not find other significant main effect or interaction effect.

Finally, we evaluated whether the subjects perceive the user as having the same preference if the user makes the same selections. If the subjects observed the same user responses, but provided a significantly different set of labels and label ratings about the user when presented with a different story progression (different informational events), their ratings must be influenced by differences in the story being described. To perform this analysis, the subjects' ratings of how much he/she thinks the user cares about each label is first normalized. Then we calculated the Euclidean distances between the subjects' ratings and the actual user profiles we used to generate the user's selections. The shorter the distance, the closer the subject's rating is to the user profile. T-tests were performed for evaluating if the subjects viewed the same user selections response differently when exposed to different storytelling styles (random versus non-random). No significant difference was found for stories that were generated using the negative profile. However, a significant difference was found for stories that were generated using the positive profile ($p<.05$).

Discussion and Future Work

Our hypothesis is partially confirmed. For subjects younger than 40, they preferred the story generated when the user's input is taken into consideration. However, for subjects in the higher age range, a different relationship was observed. This suggests that older adults may have certain expectations for stories or fables that younger individuals do not. For 40 to 50 years old subjects there is a clear preference towards the random condition. In contrast, in every other age group, this effect is not observed. It may also be the case that older audiences require a longer story to effectively notice the topic consistency demonstrated through the informational events in the non-random condition. Our current stories only consist of 8 steps. In future evaluations, we plan to use longer stories.

The propensity to prefer the random condition as more novel in older individuals may also relate to a potentially emblematic problem in subject's interpretation of the user – participants naturally use their own level of engagement, interest, and novelty as that of the user's. The significant covariation between each of the dependent variables, particularly in regard to subject engagement and the impact of novelty for the user are biasing their decision. This is also indicated through the positive user selection conditions. There was a significant difference in the labels selected to represent the user's preferences, even though the user's responses were identical. In this way, their sense of the user's experience were actually reflecting their own preferences for the random stories.

For our next step, the strong correlation among the dependent variables over all age groups in this study suggest that a retooling of the questions are necessary for future studies. Likewise, the correlation between subject engagement and questions pertaining to the user's perception of the story indicate the necessity of having the subjects actually interact with the system as opposed to evaluating another user.

Ongoing work is also being directed towards refining several aspects of the system as well. Currently, the label values in each profile are initialized randomly. This can be problematic, as initially the system may predict the user prefers materials that do not actually interest them. To reduce this possible complication, these values can be initialized according to the results of future user's interactions and their corresponding decisions in a subsequent study.

Conclusion

In this work we have proposed a computational approach for active and adaptive storytelling. Our storytelling system adapts the narrative of the story based on the user's feedback, and also consider aspects of novelty, exploration, and consistency of topics during storytelling. A preliminary evaluation was performed and confirmed that the more adaptive stories are preferred in younger age groups. This evaluation also revealed that people in older age groups may prefer stories with greater variability, and females typically found the story both more engaging and relative to males expected subjects to feel the same way. These results suggest several interesting directions for us to improve the storyteller and create a more personal experience for users in the future.

References

Arnone, M.; Small, R.; Chauncey, S.; and McKenna, P. 2011. Curiosity, interest and engagement in technology-pervasive learning environments: a new research agenda. *Educational Technology Research and Development* 59:181–198.

Berlyne, D. 1954. A theory of human curiosity. *Motivation and Emotion* 45:256–265.

Day, H. 1982. Curiosity and the interested explorer. *Intrinsic motivation* 53:109–132.

El-Nasr, M. S. 2007. Interaction, narrative, and drama:

Creating an adaptive interactive narrative using performance arts theories. *Interaction Studies* 8(2):209–240.

Garber-Barron, M., and Si, M. 2012. Towards interest and engagement: A framework for adaptive storytelling. In *Proceedings of the 5th Workshop on Intelligent Narrative Technologies Co-located with 8th AAAI Conference on Artificial Intelligence in Interactive Digital Entertainment*.

Hidi, S., and Renninger, K. A. 2006. The four-phase model of interest development. *Educational Psychologist* 41(2):111–127.

Lee, S.; Mott, B.; and Lester, J. 2010. Optimizing story-based learning: An investigation of student narrative profiles. In Aleven, V.; Kay, J.; and Mostow, J., eds., *Intelligent Tutoring Systems*, 155–165. Springer Berlin / Heidelberg.

Loewenstein, G. 1994. The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin* 116:75–98.

Mateas, M., and Stern, A. 2003. Façade: An experiment in building a fully-realized interactive drama. In *proceedings of 2003 Game Developers Conference*.

Nakasone, A.; Prendinger, H.; and Ishizuka, M. 2009. Isrst: Generating interesting multimedia stories on the web. *Applied Artificial Intelligence* 23(7):633–679.

Nelson, M. J., and Mateas, M. 2005. Search-based drama management in the interactive fiction anchorhead. In *Proceedings of Artificial Intelligence For Interactive Media and Games*, 204–215.

Pizzi, D.; Charles, F.; Lugrin, J.-L.; and Cavazza, M. 2007. Interactive storytelling with literary feelings. In *Proceedings of the 2nd international conference on Affective Computing and Intelligent Interaction*, ACII ’07, 630–641. Springer-Verlag.

Reeve, J., and Nix, G. 1997. Expressing intrinsic motivation through acts of exploration and facial displays of interest. *Motivation and Emotion* 21:237–250.

Riedl, M.; Saretto, C. J.; and Young, R. M. 2003. Managing interaction between users and agents in a multi-agent storytelling environment. In *Proceedings of the second international joint conference on Autonomous agents and multiagent systems*, AAMAS ’03, 741–748. New York, NY, USA: ACM.

Rowe, J. P.; Shores, L. R.; Mott, B. W.; and Lester, J. C. 2010. A framework for narrative adaptation in interactive story-based learning environments. In *Proceedings of the Intelligent Narrative Technologies III Workshop*, INT3 ’10, 14:1–14:8. New York, NY, USA: ACM.

Sharma, M.; Ontan, S.; Mehta, M.; and Ram, A. 2010. Drama management and player modeling for interactive fiction games. *Computational Intelligence* 26(2):183–211.

Si, M.; Marsella, S.; and Pynadath, D. 2009. Directorial control in a decision-theoretic framework for interactive narrative. In *Proceedings of the International Conference on Interactive Digital Storytelling*, 221–233.

Silva, A.; Raimundo, G.; and Paiva, A. 2003. Tell me that bit again... bringing interactivity to a virtual storyteller. In *Proceedings of the 2nd International Conference on Virtual Storytelling*, 146–154.

Sung-Ling, P. 1916. The painted skin. In *Strange Stories from A Chinese Studio*. University Press of the Pacific. 47–51.

Tanenbaum, J., and Tomizu, A. 2008. Narrative meaning creation in interactive storytelling. *International Journal of Computational Science* 3–20.

Theune, M.; Faas, S.; Heylen, D. K. J.; and Nijholt, A. 2003. The virtual storyteller: story creation by intelligent agents. In *Proceedings of 2003 Technologies for Interactive Digital Storytelling and Entertainment*, 204–215.

Thue, D.; Bulitko, V.; Spetch, M.; and Wasylishen, E. 2007. Interactive storytelling: A player modelling approach. In *Proceedings of the 3rd Conference on Articial Intelligence for Interactive Digital Entertainment Conference*.

Wade, S. 1992. How interest affects learning from text. In Renninger, A.; Hidi, S.; and Krapp, A., eds., *The role of interest in learning and development*. 281–296.

Weyhrauch, P. W. 1997. *Guiding interactive drama*. Ph.D. Dissertation, Pittsburgh, PA, USA.