Narrative Generation for Suspense: Modeling and Evaluation

Yun-Gyung Cheong[1] and R. Michael Young[2]

Graphics&OS Group, Software Labs, SAIT, CTO, Samsung Electronics CO., LTD.
Giheung-Gu, Gyeonggi-Do, South Korea[1]
Liquid Narrative Group, Computer Science, North Carolina State University
Raleigh, North Carolina, USA[2]
yuna.cheong AT samsung.com, young AT csc.ncsu.edu

Abstract. Although suspense contributes significantly to the enjoyment of a narrative by its readers, there has been little research on the automated generation of stories that evoke specific cognitive and affective responses in their readers. The goal of this research is to develop and evaluate a system that produces a narrative designed specifically to evoke suspense from the reader. The system takes as input a plan data structure representing the goals of a storyworld's characters and the actions they perform in pursuit of them. Adapting theories developed by cognitive psychologists, the system uses a plan-based model of narrative comprehension to determine the final content of the story in order to manipulate a reader's level of suspense in specific ways. This paper outlines the various components of the system and describes an empirical evaluation. The evaluation provides strong support for the claim that the system is effective in generating suspenseful stories.

Keywords: narrative generation, suspense, cognitive model, AI planning.

1 Introduction

Automated story generation has been extensively studied, with applications ranging from computer games to education and training [7; 16; 21]. While a majority of these studies are on automatic generation of logically flawless content, the emotional aspect of storytelling, which is an essential story element for the reader's enjoyment, has received less attention from the interactive storytelling research community. While the computational models of emotion in relation to individual agents have been explored [13; 18], this paper focuses on a rarely explored emotion—suspense—that the audience would feel.

Suspense is the feeling of excitement or anxiety that audience members feel when they are waiting for something to happen and are uncertain about a significant outcome [17; 25; 28]. The significance of suspense in story appreciation has been supported by several studies [1; 3]. In the Brewer and Lichtenstein's study, the participants reported that suspense is cardinal for discerning a story from a mere series of events [3]. Furthermore, the study of viewers' responses to commercials by Alwitt [1] demonstrates that suspenseful commercials are favored over non-suspenseful commercials. As an effort to explore suspense regarding story structure, Brewer and Lichtenstein [3] claim that affective states in the reader are provoked by arranging the temporal ordering of the events underlying a story world. Their theory explains that suspense could be evoked by presenting the events of a story chronologically to the reader while surprise and curiosity could be caused by hiding a critical fact or event early in the story world and disclosing it later in the text.

This paper presents a computational model of suspense, exploring the concept that a reader's suspense level is affected by the number of solutions available to the problems faced by a narrative's protagonists [2; 5; 6; 10; 11; 12; 28]. The reader's suspense is heightened when undesirable outcomes are likely to happen over preferred outcomes. It is not our intention, however, to deal with the type of suspense that is evoked by visual stimulation such as car chases in film.

Our approach attempts to manipulate the level of suspense experienced by a story's reader by determining what story elements to tell — that can influence the reader's narrative comprehension process. To this end, we make use of a computational model of that comprehension process based on evidence from previous psychological studies [2; 10; 12]. To generate suspenseful stories, we set out a basic approach built on a tripartite model, adapted from narrative theory, that involves the following elements: the *fabula*, the *sjuzhet*, and the discourse [22]. A *fabula* is a story world that includes all the events, characters, and situations in a story. In our approach, the *fabula* is represented as a plan structure generated by Crossbow—a hierarchical, partial-order causal link planner based on the Longbow planning system [26]. A *sjuzhet* is a series of events selected from the *fabula* and an ordering over those events indicating the order in which they are to be presented to readers. The final layer, a discourse, can be thought of as the medium of presentation itself (e.g., text, film). Although not directly discussed in this paper, discourse is important for the effective presentation of a story for the reader [4]. Figure 1 presents our three-stage pipelined architecture for story generation as shown, in which Suspenser is situated as a *sjuzhet* generator.

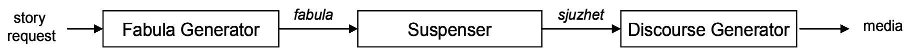

Figure 1. Tripartite suspense generation model

In this paper, we present Suspenser, a framework that determines narrative contents (i.e., *sjuzhet*) from a given story world (i.e., *fabula*) intended to evoke high level of suspense from the reader as illustrated in Figure 1. We assume in the work described here that the stories we deal with all contain conflict. For example, characters' individual goals may be negations of each others', or the plans formed by characters to achieve their goals may interfere with the plans of other characters. While other dramatic devices such as the prolonging of resolution are also useful in crating suspense, we focus here on suspense that arises as a result of users' consideration of these conflicts and their consequence on the protagonist's goals.

2 A Computational Model of Narrative Generation for Suspense

Suspenser takes three elements as input: a *fabula*, a given point *t* in the story plan that corresponds to the point where the reader's suspense is measured, and the story length desired by the system user. The system then determines the *sjuzhet*, the content of the discourse to be used to convey the story up to *t* to a reader, which enables the reader to infer a minimum number of complete plans for the protagonists' goal, following the psychological research on suspense [10; 12]. In addition, we require that the resulting *sjuzhet* shall be read as a coherent story that represents the input *fabula*.

To produce a *sjuzhet* meeting these requirements, the framework composed of two phases: a skeleton building step and an additional story element identification step. In the skeleton building step, Suspenser identifies the *skeleton* of the *fabula*—a partial plan that specifies its plan steps as a set of core story events that cannot be eliminated without harming the understandability of a story—by rating each individual event's importance based on the event's causal relationship to the protagonists' goals. In the second phase, Suspenser finds actions that can harm the protagonist's goals and tests if the addition of these actions intensifies the reader suspense by modeling the reader's inference process and anticipation of the protagonists' success. The core story events together with harmful actions compose the final content of the *sjuzhet*.

In modeling the reader's inference process and anticipation of the protagonists' success, Suspenser uses Crossbow to model the reader's plan-related reasoning processes. Prior work has provided strong evidence that human task reasoning is closely related to partial-order planning algorithms [19] and that *refinement search* [15], the type of plan construction process performed by Crossbow, can be used as an effective model of the plan reasoning process [27].

2.1 Building the Skeleton

The skeleton building phase determines important events based on the user's knowledge. This step first extracts a series of important events of the story, i.e., a skeleton, and then it tests the skeleton to ensure that its content can be understood as an integral story. To generate a candidate skeleton, the system rates the importance of each event based on a method for extracting important actions that are likely to be included in the story recall, devised by Trabasso et al. [23]. Their approach approximates an individual story event's importance by counting the number of causal relationships with other steps in the narrative and by measuring each event's importance by analyzing its role in a series of actions in a story that are causally related. Adapting their approach, the system computes each step's *importance value* by counting the number of the step's incoming and outgoing causal links of the step and taking into account its role in the plan. For instance, the first action in a story plan and actions that establish the goal state are highly eligible for inclusion in the skeleton. Finally, the top *N* (the desired story length) events are selected.

Secondly, the system tests whether the skeleton is coherent from the reader's perspective using an algorithm which is a cycle composed of two phases. The first step uses the reasoning algorithm in the reader model to find complete plans to achieve the protagonist's goals which are consistent with the skeleton candidate. If such a plan is found, the story skeleton is coherent and the program exits. Otherwise, an event in the *fabula* which was not selected as a skeleton with the highest *importance value* is chosen and added to the candidate. Then, the first phase begins again.

Input $<t, k, F, W, G, S_K, T_h, R, P, L>$ where

- *t* is the step where the reader's suspense is measured, *k* is an integer
- An input *fabula* F=(SP, B, CL, O) where SP={s_1, s_2, …, s_l} where s_i is a step, B={b_1, b_2, …, b_n} such that b_i is a tuple of $< s_l, p_1, v_1>$ where $s_l \in$ SP, which means that the plan step s_l binds the parameter p_1 to a literal v_1, CL={c_1, c_2, …, c_n} such that c_i is a causal link information represented as a tuple $<e, s_1, s_2>$ where *e* is a condition, and $s_1 \in$ SP and $s_2 \in$ SP, which means that plan step s_1 enables the precondition *e* of s_2, O={ o_1, o_2, …, o_n } such that o_i is a temporal constraints represented as a tuple of $<s_i\ s_j>$ where $s_i \in$ SP and $s_j \in$ SP
- W = {i_1, i_2, …, i_n} such that i_j is a tuple of $<s_j, w_j>$ where $s_j \in$ SP when w_j is the importance value of $s_{j.}$
- *G* is the protagonist's goals, T_h is a real number that represents a threshold
- S_K is the portion of the skeleton preceding *t*
- R is an integer representing the reader's resource bound
- P is a planning algorithm, L is a plan library denoting the reader's knowledge

1 Initialization

- Set $S_T = S_K$, *BSP* = the portion of *SP* preceding *t*
- Set $S = BSP - S_T$, NZ = {}, PZ = {}

2 (Construct S_T.) Repeat this step for *k* times

2.1 If *S* is empty, return S_T and exit, otherwise

- o Pick the action e_S in *S* generating the highest $h(e_S, F)$. If several candidates are found, non-deterministically select an action with the greatest importance value.
  - o If $h(e_{S,}, F) < T_h$, return S_T and exit, otherwise
  - o Remove e_S from *S*.
- o Pick the action e_K in S_T with the lowest importance value. If several candidates are found, non-deterministically select an action.
- o Replace e_K with e_S
  - o If $h(e_K, F) < h(e_S, F)$
    - ▪ Construct a partial plan *NZ* that only contains information in F which has s where s$\in$ ($S_T - e_K + e_S$).
    - ▪ Construct a partial plan *PZ* that only contains information in F which has s where s$\in$ S_T.
    - ▪ If $sl(G, NZ, L, P, R) > sl(G, PZ, L, P, R)$
      - ▪ $S_T = S_T - e_K + e_S$
      - ▪ *Add* e_K to *S*
      - ▪ *Goto* step 2.1
  - o Add e_S to *S*.
  - o Goto step 2.1

Figure 2. Content selection for suspenseful stories

2.2 Finding the Additional Story Elements for Suspense

The additional story element identification step constructs the *sjuzhet* (content) evoking the intended suspense level from the reader at *t*, the target step when the reader's suspense level is measured. The algorithm in Figure 2 first selects $e_{S,}$ the action with the greatest *potential suspense*, from the events in the input *fabula* that are not included in the current S_T, where S_T is a series of events to be presented to the reader. If the *potential suspense* of e_S is lower than a predefined threshold T_h, then the program exits and creates a partial plan *P* composed of the steps in S_T. If the *potential suspense* of e_S passes the threshold, the system chooses the least important action e_K

in S_T, and replaces it with the action e_S. Then the system computes the *suspense level* of the newly updated *sjuzhet*. If this substitution lowers the *suspense level* produced from the previous *sjuzhet*, the system brings back the previous value of S_T; otherwise, the update is maintained. This process repeats until there is no candidate is found or for a specified times. When it terminates, the system specifies the content of the output *sjuzhet* as S_T. The next two subsections describe a heuristic function that computes the *suspense level* of a given partial plan, and two heuristics that together compute the *potential suspense* of an action in a plan.

2.2.1 Measuring Suspense Level

In measuring the suspense level on the reader's part, the system follows the notion articulated by Gerrig and Bernardo [12], in which they view an audience as problem-solvers: an audience will feel an increased measure of suspense as the number of options for the protagonist's successful outcome(s) decreases. Adopting these models, we devise Heuristic Function 1 for measuring the *level of suspense*; the function computes the reader's suspense level as the inverse of the number of planned solutions for the protagonists' goal using her reasoning algorithm and her plan library within her reasoning limit. The function sets the minimum level of suspense when no usable solutions are found in her plan space.

Heuristic Function 1 (*Level of suspense*) In the Suspense level function $SL(G, Z, L, P, R)$, G is a set of literals representing the goal of a narrative's protagonist, Z is a partial plan, L is a plan library, P is a planning algorithm, R is an integer representing a reasoning bound, and $success(G, Z, L, P, R)$ returns the number of paths to make G true with given Z and R. $SL(G, Z, L, P, R)$ is set to $(1/success(G, Z, L, P, R))$ when $success(G, Z, L, P, R)$ returns a non-zero value and zero when $success(G, Z, L, P, R)$ returns 0.

2.2.2 Measuring Potential Suspense for an Action

In computing the potential suspense of an action's effect, we consider the action's possible causal relationship to accomplishing the protagonist's goal from the reader's point of view. For example, in a scene in the film *Back to the Future* directed by Robert Zemeckis, the protagonist Marty McFly who just came back to 1985 from 1955 saw Dr. Brown being shot by terrorists. A moment later, however, it was revealed that Dr. Brown was still alive because he was wearing a bullet-proof vest. Although Dr. Brown survived from the shooting after all, the viewers would experience suspense in the shooting scene because they are ignorant of the bullet-proof vest.

In a similar fashion, Heuristic Function 2 computes the *potential suspense* for an action by counting the number of its effects that negate the protagonist's goal and the number of its effects that unify the goal under the assumption of the audience's partial knowledge. As an illustration, Figure 3 shows that the action A has an effect $\sim g$, which is the negation of the goal literal g. We call this type of a temporary threat as a *threatening link*, referring to an action's effect negating another step's precondition in the plan. In contrast, the suspense creator establishes a *supporting link* when an effect of an action unifies with a precondition of an action in the plan. One effect can have multiple threatening links or supporting links in a single plan. The *potential suspense* of an effect is computed as the supporting link summation subtracted from the threatening link summation as formalized in Heuristic Function 3.

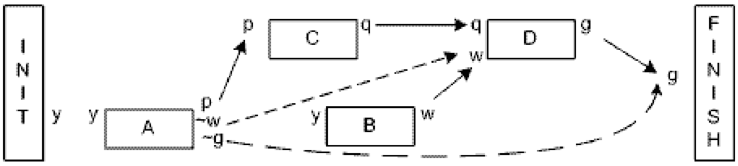

Figure 3. Threatening links in a story plan. A box represents an action, with its preconditions on the left and effects on the right. Solid arrows denote causal links. Dotted arrows are threatening links which represent an action's effect negates a precondition of other actions.

Heuristic Function 2 (Potential Suspense of an action) *h(a, p)* returns the summation of *ps(e, a, p)* where *ps(e, a, p)* is the potential suspense of the effect *e* of the action *a* in the plan *p*.

$$h(a, p) = \sum_{e \in effects(a)} ps(e, a, p)$$

Heuristic Function 3 (Potential Suspense of an effect) *ps(e, a, p)* returns potential suspense of an effect *e* of an action *a* in a plan *p*, which is the summation of the potential threat of the *e*'s supporting links subtracted from the summation of the potential threat of the *e*'s threatening links as formalized as the following equation.

$$ps(e, a, p) = \sum_{t \in Tlink(e)} \frac{w_t}{dist(d_t, p)} - \sum_{s \in Slink(e)} \frac{w_s}{dist(d_s, p)}$$

Where *Tlink(e)* returns all the threatening links of an effect *e*, *Slink(e)* returns all the supporting links of *e*, w_t and w_s are coefficients, d_l denotes the destination step of the link *l*, and *dist(s, p)* returns a value associated with the causal distance between step *a* and the goal step of plan *p*. All scaling factors are constrained to be nonnegative real numbers. In this study, *Dist(a, p)* returned $d \times (d + 1)$ where *d* denotes the distance from an action to the goal (i.e., the minimum number of causal links that relate an action to the goal in a plan). The scaling factors w_t and w_s were assigned 7.0 and 2.0 respectively, and the value of the predefined threshold T_h in the algorithm shown in Figure 2 is assigned 0.07 in this study, determined empirically from some informal experiments.

3 Evaluation

A number of informal experiments and pilot studies have been carried out to evaluate partial implementations of the Suspenser framework [8; 9]. This section describes the experiment that we carried out to evaluate the effectiveness of stories that a complete implementation of Suspenser produces in terms of suspense. The hypothesis for our study was to test if there was any association between the story generator type (independent variable) and the suspense level of the stories (dependent variable). To test this hypothesis, the suspense levels among the stories produced by a) Suspenser in high-suspense mode, b) a human author intended to create high suspense, and c) a human author intended to create low suspense were compared to detect a significant difference among them.

Background

Sykes is the owner of the Hollywood Theater, which was once prosperous but has now become dilapidated and is in need of major renovations. Sykes has accrued a sizable gambling debt, and with his theater in shambles, he has no means with which to pay it back. He is constantly threatened by his crooked debtors. Janet is a famous actress with dreams of winning an Oscar, an acting award. She is jealous of the actress Agatha, who is her contender for the Oscar this year and also is well-known for her active involvement in charity. Janet knows a number of scoundrels including a guy named Kent, a bomb dealer, and the theater owner Sykes. Agatha is in love with Bill, who serves as a lieutenant in the Los Angeles Police Department's Serious Crime squad. Janet knows that Agatha is planning to go to the Charity Bazaar for the Poor to be held in Hollywood Theater. To ensure that she will win the Oscar, Janet plans to kill Agatha during the charity event.

Storywriter's selection for high suspense effect

Janet and Sykes plan to burn down Sykes' theater to get the insurance money and kill Agatha during the charity bazaar. Janet gives Kent's contact information to Sykes and informs him of Kent's expertise with firebombs. Kent takes a bomb to the Hollywood Theater and meets with Sykes. Sykes purchases the firebomb. Sykes installs the firebomb. Kent informs Bill that Sykes is planning to firebomb his own theater during the charity event. Agatha goes to the theater for the charity event. Sykes sets the timer of the firebomb to explode during the charity event. Sykes switches on the firebomb. Bill searches for the firebomb in the theater. *Bill defuses the firebomb.*

The system's selection for high suspense effect

Kent takes a bomb to the Hollywood Theater and meets with Sykes. Sykes purchases the firebomb. Sykes installs the firebomb. Bill arrests Kent. Kent informs Bill that Sykes is planning to firebomb his own theater during the charity event. Bill releases Kent for his cooperation. Agatha goes to the theater for the charity event. Sykes sets the timer of the firebomb to explode during the charity event. Sykes switches on the firebomb. Bill searches for the firebomb in the theater. *Bill defuses the firebomb. Agatha participates in the charity event.*

Storywriter's selection for low suspense effect

Janet convinces Sykes to participate in her plan to kill Agatha by convincing him that if he participates, he will be able pay off his gambling debts. Sykes borrows some money from the bank by mortgaging his theater. Sykes buys insurance to cover his loss in case of a fire. Janet gives Kent's contact information to Sykes and informs him of Kent's expertise with firebombs. Kent takes a bomb to the Hollywood Theater and meets with Sykes. Sykes purchases the firebomb. The lieutenant, Bill, issues a warrant permitting the arrest of Kent for his illegal weapons dealing. Bill coaxes Kent to give information in exchange for releasing him. Bill releases Kent for his cooperation. *Agatha participates in the charity event.*

Figure 4. Three *Sjuzhets* produced from *Fabula* C: Italicized sentences are the portion after suspense was measured.

3.1 Method

Participants and design

A total of 98 unpaid subjects voluntarily took part in the experiment, ranging in age from 20 to more than 50 years old (42 males, 57 females, and one no response): 72 recruited from NCSU communities including recently graduated under/graduate students across different departments and 26 from internet female technical communities (e.g., Systers.org). All subjects were native-speakers of English. The study utilized a repeated measured between group design: subjects were randomly assigned to one of nine subject groups. These groups were arranged according to a 3 × 3 Latin Square design to counter-balance the interference from different orderings of stories. From this design, a subject was shown one version of each of the three *fabulas*.

Materials and apparatus

To obtain an input to Suspenser, we ran Crossbow to plan three *fabulas*. The resulting plans consisted of partially ordered 18-20 steps which were manually linearized, and each plan was realized as text using a simple template-matching technique that mapped one plan step into a single sentence. For the study, we prepared a total of nine *sjuzhets*, by generating three *sjuzhets* for each *fabula*—one by Suspenser in high suspense mode and two stories by a human author. One of the *fabulas* and its three *sjuzhets* used in this study are shown in Figure 4. To obtain human generated stories, we recruited one Master's student majoring in English at North Carolina State University, a freelancer writer who had her short story published in a local newspaper. She was presented with texts on sheets that describe the three *fabulas* and their corresponding measurement points. She then was asked to select two series of sentences for each *fabula*: one to arouse high suspense and the other to arouse low suspense from the reader when his suspense level would be measured at a given point in the story. For this study we did not constrain the number of sentences that she selected. As a result, her two versions of a story differed in size within a margin of 2.

Procedure

Each subject individually participated in the study by accessing a web site. Each subject was presented with three stories and was asked to rate his suspense levels at one point in his reading each of the stories. Each story was presented to the subject sentence by sentence; one page contained only one sentence and a button click led to the next page. After reading the portion preceding the measurement point displayed on separate pages, the subject was asked to describe his suspense level on a five-point scale basis ranging from "no suspense" to "extremely suspenseful." After responding to the question, the subject was presented with the second part of the story sentence by sentence, followed by a page asking generic questions about story coherence and enjoyment on a five-scale basis ranging from "not at all" to "strongly agree."

3.2 Results

The collected data contained 294 responses from 98 subjects. To detect a significant difference between three story generators, we performed a one-way ANOVA on the collected data using SAS version 9.1.3 SP4. In this analysis, two main effects were examined: the story generator type and the *fabula* type. Each type has three levels.

As shown in Table 1, the data indicated that the story generator type had an effect on the suspense level ($F(2, 285)=4.27$, p value=0.015). The result also shows that the *fabula* type had no effect on suspense. No interaction effect was found between the *fabula* type and the story generator type ($F(4, 285)=0.66$, p value=0.622). Despite the short sample stories, the subjects rated their experience of suspense is "moderate"

(Mean=2.571/5.0, SD=1.059) on a five-point Likert scale. The system performance was superior to the other story generators in the categories of *fabula* B (Mean=2.727, SD=1.126) and *fabula* C (Mean=2.939, SD=1.088).

A series of standard one-tailed t-tests were used to compare the performance of the three story generators. The results indicate that the stories produced by the system (Mean=2.704) and the human author intended for high suspense (Mean=2.694) were rated as more suspenseful than the version produced by the human author intended for low suspense (Mean=2.316) with a 99% of confidence (Suspenser vs. Human-LS $t(194)=2.56$, p value=0.006; Human-HS vs. Human-LS $t(194)=2.50$, p value=0.007).

Table.1. Data for Suspense

Means and standard deviations for suspense in each story generator type (N=98)

Suspenser in the high-suspense mode		Human author for high suspense		Human author for low suspense	
M	SD	M	SD	M	SD
2.704	1.057	2.694	1.049	2.316	1.061

Means and standard deviations for suspense in each story and story generator

Story	Generator	N	M	SD
Fabula A	Suspenser	32	2.438	0.914
	Human-HS	33	2.667	0.890
	Human-LS	33	2.303	1.104
Fabula B	Suspenser	33	2.727	1.126
	Human-HS	32	2.656	1.096
	Human-LS	33	2.394	1.144
Fabula C	Suspenser	33	2.939	1.088
	Human-HS	33	2.758	1.173
	Human-LS	32	2.250	0.950

NOTE: Human-HS denotes the human author’s selection intended to create high suspense and Human-LS denotes the human author’s selection intended to create low suspense.

ANOVA summary table for Suspense

Source	**DF**	**SS**	**Mean Square**	**F Value**	**Pr > F**
Fabula	2	1.712	0.857	0.76	0.467
Generator	2	9.571	4.786	4.27	0.015
*Fabula**Generator	4	2.954	0.738	0.66	0.622
Error	285	319.760	1.122		

3.3 Discussion

The data clearly show that the story generators had an influence on the amount of suspense that the subjects felt. In particular, the stories produced by Suspenser created stories comparable in suspense to those produced by human authors intended for high suspense effect (Suspenser Mean = 2.704; Human author intended for high suspense Mean = 2.694). The results also show that the difference between the suspense levels felt by the subjects from Suspenser's story for high-suspense and the human author's story for low-suspense was significant with a 99% of confidence.

To test if Suspenser selects appropriate content for the effect of suspense, we investigated the contents of the six *sjuzhets*, and the result indicates that the set of stories for high suspense effect differed in content from the set for low suspense effect. The story created by the system overlapped that created by the human author intended for high suspense in 50%-80% of the total number of story sentences (*fabula* A 50%, *fabula* B 60%, *fabula* C 80%). In contrast, the stories created for high suspense overlapped the story created by the author intended for low suspense in 20%-30% (*fabula* A 20%, *fabula* B 20%, *fabula* C 30%). This means that the story event sets targeting high suspense and the set intended for low-suspense tend to be mutually exclusive. The story events that the author selected for low suspense were not related to the protagonist's goals.

To test if the text quality affected the reader's story comprehension, the subjects' responses to story coherency were also analyzed. The data suggest that the text quality was good enough for the subjects to understand the stories. The participants evaluated the given stories as relatively coherent (Mean=2.938/5.0, SD=1.031).

4 Conclusion and Future Work

The generation of stories by computers has been the focus of research by computational linguists and AI researchers for several decades. Although a number of approaches have shown promise in their ability to generate narrative, there has been little research on creating stories for an intended emotion.

To address this problem, we present a computational model that takes a complete story world and elaborates a story structure—content—that can manipulate reader suspense at a specific point in its telling. In constructing the story structure, this approach gauges the suspense level that a reader would feel by modeling the reader's narrative comprehension using a planning technique. This approach takes as input a partial plan indicating the portion of a story that has been conveyed so far and computes the reader's anticipated suspense level based on the inverse of the number of solution plans that can be found to the protagonist's goals in the space of plans she can consider within her reasoning resources.

To generate a partial plan that maximizes the reader's suspense, the system takes a plan as input and selects a set of core events that have high causal connectivity and that also play an important role in the story as basic building blocks. The partial plan then is supplemented by harmful actions (e.g., those that conflict with the protagonist's goals) that intensify the reader's suspense level. The model has been implemented and formally evaluated. The data from the experiments have shown this system to be successful in selecting content that elicits high suspense. In particular, the data show that, in the context of our experiments, this model was as effective as a human author in generating suspenseful stories.

While the results of this study show that Suspenser was effective in generating suspenseful stories, the design of the experiment does not allow us to point conclusively at single reason for its effectiveness. For instance, the plan representation used in this study did not allow a plan to have conflicting goals; a plan structure used in this research was considered a sound solution plan only when it contains no conflicts. In order to create conflicting situations—critical conditions for suspense—the characters' goals were manually specified to foster a compelling story. As a result, protagonist's and antagonist's plans were often related via causal relationships. A redesign of the experiment to use a more conflict-expressive plan representation is needed to better characterize the contribution of the system in the readers' level of suspense.

We plan to extend this model to interactive environments by expanding previous related work on narrative replanning techniques [14; 20]. Our future work also includes bidirectional interactions among the *fabula*, *sjuzhet*, and discourse layers. For example, the technique of postponing story resolution has often been employed for the effect of suspense in human-authored narratives. This is also computerized in the MINSTREL system by inserting additional events that detail the protagonist's struggles in between the story's climax and its resolution [24]. With a bidirectional interaction model, Suspenser could revise the *fabula* to include auxiliary events that situate the protagonist in a seemingly dangerous position upon the request from the *sjuzhet* generator. Likewise, the *fabula* and the *sjuzhet* could be adjusted upon request from the discourse generator. For instance, in filming a scene, the discourse generator may find no spots to capture a specified shot of characters due to the physical setting that the current *fabula* provides. Bidirectional interaction would allow the *fabula* to be replaced with a new *fabula* to fix this problem.

Acknowledgments

Partial support for the work described in this paper has been provided by National Science Foundation CAREER award # 0092586 and award 0414722.

References

1. Alwitt, L. F. Suspense and Advertising Response. Journal of Consumer Psychology, 12 (1), pp. 35-49 (2002)
2. Brewer, W. F. The Nature of Narrative Suspense and the Problem of Rereading. In P. Vorderer, H. J. Wulff, & M. Friedrichsen (Eds.), Suspense: Conceptualizations, Theoretical Analyses, and Empirical Explorations, pp. 107-127, Mahwah, NJ: Erlbaum (1996)
3. Brewer, W. F. and Lichtenstein, E.H. Stories Are to Entertain: A Structural-Affect Theory of Stories, Journal of Pragmatics, 6 (5-6), pp. 437-86 (1982)
4. Callaway, C. B. and Lester, J. C. Narrative Prose Generation. Artificial Intelligence, 139, pp. 213-252 (2002)
5. Carroll, N. Toward a Theory of Film Suspense. Persistence of Vision, 1, pp. 65-89 (1984)
6. Carroll, N. The Paradox of Suspense. In P. Vorderer, H. J. Wulff, & M. Friedrichsen (Eds.), Suspense: Conceptualizations, theoretical analyses and empirical explorations, pp. 71-92, Mahwah, NJ: Lawrence Erlbaum Associates, Inc. (1996).

7. Cavazza, M., Charles, F., and Mead, S. J. Character-Based Interactive Storytelling. IEEE Intelligent Systems, 17(4), pp. 17-24. (2002)
8. Cheong, Y. and Young, R.M. A Computational Model of Narrative Generation for Suspense. In the AAAI 2006 Workshop on Computational Aesthetics, pp. 8-15 (2006)
9. Cheong, Y. and Young, R.M. A Framework for Summarizing Game Experiences as Narratives. In Proc. of AIIDE 2006, pp. 106-108 (2006)
10. Comisky, P., and Bryant, J. Factors Involved in Generating Suspense. Human Communication Research 9(1), pp. 49-58 (1982)
11. de Wied, M. The Role of Temporal Expectancies in the Production of Film Suspense. Poetics, 23, pp. 107-123 (1994)
12. Gerrig, R., and Bernardo, D. Readers as Problem-solvers in the Experience of Suspense. Poetics, 22, pp. 459-472 (1994)
13. Gratch, J., Marsella, S.: Tears and Fears: Modeling Emotions and Emotional Behaviors in Synthetic Agents. In Proceedings of the Fifth International Conference on Autonomous Agents 2001, pp. 278-285 (2001).
14. Harris, J. and Young, R. M. Proactive Mediation in Plan-Based Narrative Environments. In Proc. of IVA 2005, pp. 292-304 (2005)
15. Kambhampati, S., Knoblock, C. A., and Yang, Q. Planning as Refinement Search: A Unified Framework for Evaluating Design Tradeoffs in Partial-Order Planning. Artificial Intelligence, 76(1-2), pp. 167-238 (1995)
16. Mateas, M. and Stern, A. Façade: An Experiment in Building a Fully-Realized Interactive Drama. In Game Developer's Conference: Game Design Track, San Jose, California (2003)
17. Ortony, A., Clore, G., and Collins, A. (1988). The Cognitive Structure of Emotions. New York: Cambridge University Press.
18. Pizzi, D., Charles, F., Lugrin J.-L., and Cavazza, M. Interactive Storytelling with Literary Feelings. In Proc. of ACII 2007, Lisbon, Portugal, September (2007).
19. Rattermann, M. J., Spector, L., Grafman, J., Levin, H. and Harward, H. Partial and Total-order Planning: Evidence from Normal and Prefrontally Damaged Populations, Cognitive Science, 25(6), pp. 941-975 (2002)
20. Riedl, M., Saretto, C. J. and Young, R. M. Managing Interaction between Users and Agents in a Multiagent Storytelling Environment. In Proc. of AAMAS 2003, pp. 741-748 (2003)
21. Riedl, M. O. and Young, R. M. An Intent-driven Planner for Multi-agent Story Generation. In Proc. of AAMAS 2004, pp. 186-193 (2004)
22. Rimmon-Kenan, S. Narrative Fiction: Contemporary Poetics. New York: Metheun, Routledge (2002)
23. Trabasso, T. and Sperry, L. L. Causal Relatedness and Importance of Story Events. Journal of Memory and Language, 24, pp. 595-611 (1985)
24. Turner, S. The Creative Process: A Computer Model of Storytelling and Creativity. Hillsdale, NJ: Lawrence Erlbaum Associates (1994)
25. Vorderer, P. Toward a Psychological Theory of Suspense. In P. Vorderer, H. J. Wulff, and M. Friedrichsen (Eds.), Suspense: Conceptualizations, Theoretical Analyses, and Empirical Explorations, 233-254, Mahwah, NJ: Lawrence Erlbaum Associates (1996).
26. Young, R.M., Pollack, M.E., and Moore, J.D. Decomposition and causality in partial-order planning. In Proceedings of AIPS 1994, pp. 188-193 (1994)
27. Young, R. M. Using Grice's Maxim of Quantity to Select the Content of Plan Descriptions, Artificial Intelligence, 115, pp. 215-256 (1999)
28. Zillmann, D. The Psychology of Suspense in Dramatic Exposition. In P. Vorderer, H. J. Wulff, & M. Friedrichsen (Eds.), Suspense: Conceptualizations, theoretical analyses and empirical explorations, pp. 199-232, Mahwah, NJ: Lawrence Erlbaum Associates, Inc. (1996)