Personality and Preference Modeling for Adaptive Storytelling

Edirlei Soares de Lima
School of Technology, Arts and Communication
Universidade Europeia
Lisbon, Portugal
edirlei.lima@universidadeeuropeia.pt

Bruno Feijó, Antonio L. Furtado, Vinicius M. Gottin
Department of Informatics
Pontifical Catholic University of Rio de Janeiro
Rio de Janeiro, Brazil
{bfeijo, furtado, vgottin}@inf.puc-rio.br

***Abstract*— In almost all forms of storytelling, the background and the current state of mind of the audience members predispose them to experience a given story from a unique personal perspective. However, traditional story writers usually construct their narratives based on an average understanding of the preferences of their audience, which does not guarantee satisfying narrative experiences for its individual members. When a narrative is aimed at providing pleasurable entertainment, having some information about the preferences of the current user for the narrative's content is vital to create satisfying experiences. This paper explores personality modeling and proposes a novel approach to generate individualized interactive narratives based on the preferences of users, which are modeled in terms of the Big Five factors. The paper presents the proposed method and evaluates its precision and real-time performance.**

Keywords – interactive storytelling; personality modeling; preference modeling;

I. Introduction

In most cases, the narrator's success in achieving a desired impact on the audience depends strongly on the ability to accommodate the various audience member's personal preferences. While in traditional forms of storytelling (e.g. books, movies, comics) authors generally have no access to this information, the very nature of interactive storytelling allows the system to automatically obtain this information through the user-interaction process.

Interactive storytelling has been playing a significant role in contemporary media as a trend that has developed alongside with the popularity of video games. Most of the current video games utilize stories simply to create challenges that stimulate gameplay. Despite the current efforts of the gaming industry towards rich story-driven games (e.g. Mass Effect 2 (BioWare, 2010) and Detroit: Become Human (Quantic Dream, 2018)), gaming narrative is still designed for a fixed prototypical player. Here we envisage a next generation of games that aim at creating interactive stories automatically, which are capable of surprising and entertaining a large, varied audience, adding an extra layer of dramatic meaning to their individual experiences.

When a narrative is aimed at providing pleasurable entertainment, having some information about the preferences of the current user for its narrative content is vital to promote a satisfying experience. If the user's preferences are correctly inferred by an interactive storytelling system, well-suited elements of the story can be brought to the viewer's attention or even a completely different storyline can be generated to comply with the viewer's preferences.

The main question that motivates this work is: how to create interactive narrative experiences that satisfy users' preferences? This problem can be initially divided into two sub-problems: (1) how to create a user model to describe the main characteristics of the user in real-time? and (2) how to employ the user model to adapt narratives according to user's preferences? Perhaps, a third even more important question is (3) how user's characteristics can be correlated with narrative preferences?

Some research works demonstrate that reader preferences can influence expectations for future narrative events [1]. Therefore, an important factor to be considered when adapting narratives is the personality of the users, which is known to exert a major influence on their preferences and, consequently, affecting expectations about likely outcomes. Personality is the combination of all characteristics that form a distinctive character, an individual style of thinking, feeling and acting [37]. According to Back and Egloff [38], personality arises from interactions between the situation in which the individual is placed and the processes that take place inside the individual's mind. To assess users' personality, researchers in psychology have developed several psychometric tests and measurement scales. Among the most widely accepted are those that follow the Big Five proposal [2]. Past research in psychology [3] shows that personality of human subjects can be successfully ascertained by asking them a set of well-designed questions.

This paper explores the use of personality modeling to adapt interactive narratives according to users' preferences. We propose a novel approach to create individualized narrative experiences based on the personality of users, which are modeled in terms of the Big Five factors. The main objective of this paper is to present our method and to validate its precision and real-time performance in an interactive storytelling system.

The paper is organized as follows. Section II reviews related work. Section III describes the plot generation method used to generate stories in our system. Section IV presents the proposed personality and preference model. Section V describes how user's personality and preferences can be used to adapt narratives. Section VI describes a technical evaluation of our method. Section VII offers concluding remarks.

II. Related Work

In recent years, user modeling has attracted a lot of attention from academic research [4][5][6]. Among the several application areas, we can find some implementations of user modeling in interactive storytelling systems, which will be reviewed in this section.

Barber and Kudenko [7] present an interactive story generator system that learns the personality of its users by applying predefined increments or decrements to a vector of personality traits, such as honesty and selfishness, in response to the users' decisions. Based on a similar approach, Seif El-Nasr [8] presents an interactive storytelling system where both player behavior and personality are modeled to allow users to participate in a more engaging drama. The system tracks user's actions to adjust a vector of values representing tendencies toward character traits (heroism, violence, self-interestedness, and cowardice) via pre-specified annotations on potential player actions. For example, if a player chose to flee from a presented confrontation, the model's representation of the player's cowardice would increase.

Another interactive storytelling system that models player's personalities using a vector of traits is PaSSAGE [9]. The system uses player modeling to automatically learn a model of the player's preferences through observations of the player in the virtual world, and then uses the model to dynamically select the content of an interactive story. The player is modeled by a vector, where each dimension is the strength of one of the Laws' stereotypes [10]. As the player performs actions, dimensions are increased or decreased in accordance to predefined annotations on potential player actions. Ramirez and Bulitko [11] use this player model with a reward function in such a way that, when several narratives are generated, the one that maximizes this function is automatically selected.

All aforementioned works share a common characteristic: they model users' personalities using vectors of traits, whose values are updated according to predefined annotations on user actions. Although the use of a vector to represent personality traits is widely accepted (even the Big Five dimensions can be represented as a vector), the use of manual annotations on specific actions or events to determine how the personality traits will be updated is problematic. It requires an extra authorial work and extensive studies to correctly measure the impacts of each action in the personality of users.

A different approach is explored by Sharma et al. [12], who use a database of interest-annotated logs of past users to infer the preference of current users. By combining past captured narrative traces and user survey data, they can create user models to dynamically determine the next plot point that is best suited to specific users. Their system uses a database of logs of the experiences that previous users had with its stories and attempts to infer the interests of its current user by matching his/her trajectory through the space of possible story events with the trajectory of similar users. One clear limitation of their proposal is that it requires some user interactions before the system starts to have some information about how the decisions of a current user match the decisions of similar users.

Although this work shares some similarities with the above proposal, the initial assumptions and the strategies adopted to solve the problem are completely different. While it assumes that users with similar gameplay characteristics will have similar narrative interests, we argue that personality is a better descriptor to group users with similar preferences [13]. In addition, we rely on machine learning techniques to find correlations between personality and user's preferences for future narrative events.

III. PLOT GENERATION

Our plot generation strategy is based on the reuse of already existing stories that follow the same narrative pattern [14][15][16]. By combining a chosen set of story variants into a network-structured pattern, their coinciding, diverging and converging subsequences of events are conveniently exposed, and we can generate new alternative versions of the story by traversing the network along parts of different variants.

A. Story Network

In our model, *events* are the building blocks of a story. They have the form $e_i(p^i_1, p^i_2, \ldots, p^i_n)$, where e_i denotes a class of event, and the values of the p^i_j parameters serve to characterize different instances of this class. An example of story event is `meet('Little Red Riding Hood', 'Wolf')`, which is an instance of the event class *meet*(*X*,*Y*). An *event sequence S* is a time-ordered set of events, which we represent by the following concise notation: $[e_1, e_2, \ldots, e_m]$. As example of event sequence, consider:

```
[ask_to_take('Mother', 'Little Red Riding Hood',
'cake and butter', 'Grandmother'), go('Little
Red Riding Hood', 'the woods'), meet('Little Red
Riding Hood', 'Wolf'))]
```

involving three classes of events: *ask_to_take*, *go* and *meet*.

The story network is modeled as a directed, connected, labeled graph $G = (N, E, \alpha)$, where N is a finite set of nodes, $E \subseteq N \times N$ is a finite set of edges, $\alpha{:}N \rightarrow \Sigma_N$ is a node labeling mapping, and Σ_N is a set of node labels, such that a node is associated to a story event $e_i(p^i_1, p^i_2, \ldots, p^i_n)$. Sometimes we use e_i only (i.e. with no parameters) as the node label for the sake of simplicity.

In this network, a sequence of events is a graph walk. A *walk S* in a graph G is a sequence $N_1E_1N_2E_2\ldots N_{k-1}E_k$ of nodes and edges in G such that $E_i = N_iN_{i+1}$, for all $1 \le i \le k$. A walk is often written as $N_1N_2\ldots N_k$ and denoted by SN_k, that is:

$$SN_k = N_1N_2\ldots N_k$$

An *event node* N_k is actually a walk SN_k, that is: an event node is represented by the entire series of events (i.e. nodes) starting from the first node N_1.

The first step to create a story network involves the transformation of the chosen repertoire of story variants into event sequences $[e_1, e_2, \ldots, e_m]$. Then, two general border events (called *begin* and *end*) are added to each variant and grouped as a network structure, in which e_i are node labels (Figure 1).

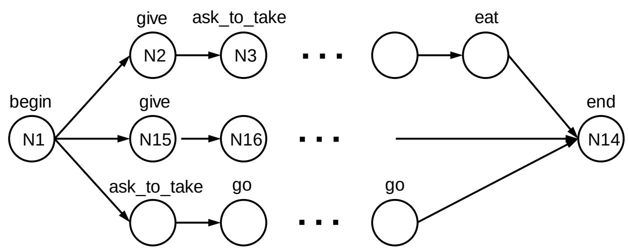

Figure 1. Example of an initial network with 3 story variants.

In the next step, similar event sequences are combined in order to transform the initial network into a reduced form called *condensed network*. Two events are similar if they denote states characterized by similar (not necessary equal) situations. For example, `eat(Wolf, Little Red Riding Hood)` and `eat(Uncle Wolf, Little Girl)` are similar events. In practice, similar events are included in a set of pairs of events that are potentially convergent (named *allowed_convergent_ list*).

The process to generate the condensed network applies two basic cases of condensation repetitively, one by *equality* and another by *similarity*:

Let e_i and e_j denote the last events of the walks S_1N_i and S_2N_j , and e_{i+1} and e_{j+1} denote the last events of the next nodes (Figure 2), that is:

$e_i = last(S_1N_i)$
$e_j = last(S_2N_j)$

and

$e_{i+1} = last(S_1N_{i+1})$
$e_{j+1} = last(S_2N_{j+1})$

then, N_i and N_j are unified if

- Fusion Case (by *equality*): $e_i = e_j$; or
- Condensation Case (by *similarity*):
  (a) $(e_i, e_j) \in$ *allowed_convergent_list*, i.e. e_i and e_i are similar events; and
  (b) $e_{i+1} = e_{j+1}$ or $(e_{i+1}, e_{j+1}) \in$ *allowed_ convergent_list*, i.e. the sequences S_1N_i and S_2N_j lead to the same event or to similar events.

In the condensation case, the condensed node keeps both labels e_i and e_j, and their respective walks S_1 and S_2.

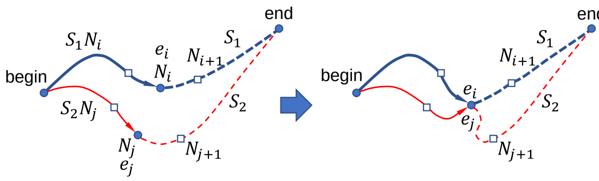

Figure 2. Given two event sequences S_1 and S_2, N_i and N_j can be unified if e_i and e_i are similar events followed by another pair of similar events e_{i+1} and e_{j+1}.

Figure 3 illustrates the process of generating a final condensed network from an initial one. In this example, Nodes `N2`, `N6` and `N4` from Figure 3a are unified as `N2` in Figure 3b because they represent equal events (event `a`). The same happens with nodes `N8` and `N10` (unified as `N6` in Figure 3b), and with nodes `N9` and `N11` (unified as `N7`). Nodes `N4` and `N7` from Figure 3b are unified as `N4` in Figure 3c because the pair (`b,d`) is in *allowed_convergent_list* and both are followed by nodes representing a similar state – indeed, both are followed by `N3`. In other words, the events `b` and `d` produce the same main effects. Also, nodes `N5` and `N1` from Figure 3b are unified as `N1` in Figure 3c because the pair (`e,begin`) is in *allowed_convergent_list* and both are followed by `N6`. In this example, `e` and `begin` are potentially convergent because the event `e` undoes the effect of `a`, and thus, if the occurrence of `a` leads from `N1` to `N2`, `e` would revert to `N1` (the `begin` event), thereby introducing a loop in the network. The final network (Figure 3c) comprises 5 nodes, of which `N1` and `N2` are fork nodes (also called branching points or interaction points in this paper).

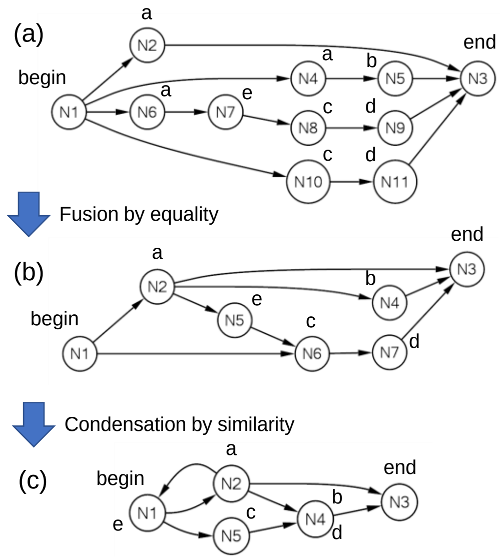

Figure 3. Example of event unification. First the network is transformed by equality and then it is condensed by similarity.

More details about our plot generation method are presented in [17].

B. Story Domain

The story domain chosen to our interactive storytelling system is based on the Little Red Riding Hood folktale (LRRH for short). After a brief survey of the literature [33][34][35][36], we selected four strikingly divergent variants that tell the little girl's story with quite different outcomes, which produces a network with a good number of branching nodes (opportunities for user interaction).

The first chosen variant is the classic Le Petit Chaperon Rouge (Little Red Riding Hood), composed in France in 1697 by Charles Perrault [33]. In this variant, the little girl, going through the woods to see her grandmother, is accosted by the wolf who reaches the grandmother's house ahead of her. The wolf kills the grandmother and takes her place in bed. When the girl arrives, she is astonished at the "grandmother"'s large, ears, large eyes, etc., until she finally asks about the long teeth, whereat the wolf gobbles her up.

The second variant, perhaps even more influential, is that of the brothers Grimm (Jacob and Wilhelm), written in German and entitled Rotkäppchen (Little Red Cap) [34], first published in 1812. It complements the Perrault variant with a rescue episode, which is performed by a hunter, who finds the wolf sleeping and cuts his belly, allowing the girl and her grandmother to escape. The wolf has his belly filled with heavy stones fetched by the girl, wakes up, tries to run away and falls dead, unable to carry the weight.

The third chosen variant is the Conte de la Mère-grand (The Story of Grandmother), collected by folklorist Achille Millien in the French province of Nivernais, 1870, and later published by Paul Delarue [35]. In this variant, the villain is a "bzou" (a werewolf). After killing and partly devouring the grandmother's body, he stores some of her flesh and fills a bottle with her blood. When the girl comes in, he directs her to eat and drink from these ghastly remains. Then he tells her to undress and lie down on the bed. When the question about the "big mouth" is asked, the Bzou gives the conventional reply: "All the better to eat you with, my child!" – but this time the action does not immediately follow the words. What happens instead is that the girl asks permission to go out to relieve herself, which is a ruse whereby she ends up outsmarting the villain and safely going back to her mother's home.

The last variant used to compose our story network is the Uncle Wolf [36], collected by Italo Calvino. In this variant, the girl does not resist the temptation to eat and drink all that her mother was sending to Uncle Wolf in return for the loan of a skillet, offering him instead an ugly mess composed of donkey manure, dirty water and lime. As revenge, Uncle Wolf sneaks into the little girl's house and eats her.

The network generated for the selected LRRH variants is composed of 47 nodes (story events), 5 fork nodes (branching points or interaction points), and 5 join nodes. The total number of plots that can be generated by traversing all possible paths across the network is 378, but, after grouping together the plots that, despite minor differences, comprise the same sequence of basic episodes, the number of essentially distinct alternatives is reduced to 13. The full network can be seen at the bottom partition of Figure 5 and a full resolution image is available in a separate online document.[1]

IV. Personality and Preference Modeling

Personality plays an important role in influencing individual preferences for game genres [18], heroic roles in games [19], and cultural participation [38].[2] While traditional forms of storytelling lack the ability of creating individual and personalized experiences, an *interactive* storytelling system can take advantage of its nature to obtain the personality of its users and use this information to adapt stories according to their individual preferences.

The personality of an individual can be determined through a variety of tests and measurement scales. Among the most widely accepted are those that follow the Five Factor Model (also known as "Big Five") [2]. Big Five is a dimensional representation of human personality structure, which claims that, by using five personality traits, it can suitably account for personality diversity. The Big Five factors are:

1. **Openness**: those who are high on this factor are imaginative, curious and receptive to new ideas. In contrast, those who score low on this factor are indifferent and uninterested;
2. **Conscientiousness**: the ones that display high degree of this factor are meticulous, efficient and systematic. Who scores low is careless, chaotic and disorderly;
3. **Extraversion**: high scorers are characterized by enjoying social activities. On the opposite side, low scorers are reserved and shy.
4. **Agreeableness**: a high score on this factor characterizes helpful, cooperative and friendly people. In contrast, low score characterizes selfish and hostile people.
5. **Neuroticism**: those who score high on this factor are emotionally unstable, anxious and aggressive. In contrast, those who score low are well-adjusted and calm.

The five dimensions of the human personality structure are supported by several questionnaires, inventories, and adjective rating scales designed to measure each dimension (e.g.: [24][25][26]). Personality classification is then achieved by assigning five numerical scores (one per dimension) that account for how well each factor describes the person. The attribution of the scores is typically performed with questionnaires that consider observable behavior and characteristics of the individual.

In psychology research, the Big Five dimensions are usually assessed through long questionnaires (60 or 44 items). However, forcing users to answer such long questionnaires in interactive storytelling applications certainly produces negative effects in the general user experience. Therefore, a better solution is to adopt simplified questionnaires, such as the BFI-10 [20], which is one of the shortest questionnaires that, as its denomination implies, measures the scores of the Big Five factors with only 10 questions.

In BFI-10, the subject answers the following 10 questions "I see myself as someone who …": (1) is reserved; (2) is generally trusting; (3) tends to be lazy; (4) is relaxed, handles stress well; (5) has few artistic interests; (6) is outgoing, sociable; (7) tends to find fault with others; (8) does a thorough job; (9) gets nervous easily; (10) has an active imagination. The answers (L values) are given in a five-point Likert scale: 1 (disagree strongly), 2 (disagree a little), 3 (neither agree nor disagree), 4 (agree a little), and 5 (agree strongly).

For each Big Five dimension, BFI-10 calculates the average score of two poles, which correspond to a true-scored item and a false-scored item respectively. The false-scored item must be reverse scored before calculations are made, so that the values 1, 2, 3, 4, and 5 become 5, 4, 3, 2, and 1 respectively, i.e.:

$$reversed\ score = 6 - score \tag{1}$$

For instance, for the Neuroticism dimension, BFI-10 evaluates how much one "gets nervous easily" (say, L = 4 for question 9) and "is relaxed, handles stress well" (say, L = 5 for question 4), that is, Neuroticism would be in this case (4 + 1)/2 = 2.5 (where 1 is the reversed score of 5). The scored items for each dimension are defined as follows (where R indicates a reversed-scored item): Extraversion (6 and 1R), Agreeableness (2 and 7R), Conscientiousness (8 and 3R), Neuroticism (9 and 4R), and Openness (10 and 5R).

Another well-known short measurement scale is the *Ten-Item Personality Inventory* (TIPI) proposed by Gosling et al. [21], which also uses two items associated with each personality dimension. However, we favored BFI-10 over TIPI because of the following reasons: (1) BFI-10 uses a five-point Likert scale rather than the seven-step scale of TIPI – which makes BFI-10 simpler and slightly faster (both take about a minute to complete); (2) BFI-10 uses statements representing two extremes of the same dimension clearly, which are more aligned with actions and

[1] http://www.icad.puc-rio.br/~logtell/fullnetwork1.pdf

[2] For example: TV programs, book reading, attending museums and concerts. Games are not mentioned in [38], but we can speculate on similarities with exciting or unconventional cultural activities.

attitudes than the more generic opposite adjectives of TIPI; (3) the BFI-10's authors (cf. [20]) have shown that BFI-10 is psychometrically superior to TIPI; (4) BFI-10 was successfully tested in more than one idiom, besides the original version in English and German [22][23] – which suggests that BFI-10 might be particularly adequate for multi-language storytelling applications.

There are two common methods to integrate Big Five questionnaires into interactive storytelling applications: (1) integrating the questionnaire statements into the narrative through story-related dialog choices; and (2) directly asking users to answer a questionnaire when they begin to interact with the system. In a previous work [27], we explored the first approach by creating an introductory narrative with 10 story-related scenes followed by decision-making points (one for each BFI-10 question), where users make decisions that are equivalent to answering BFI-10 questions. Each scene creates a situation that stimulates users to react in a way that makes evident their answer to the BFI-10 question that defined the scene. Although this approach favors the overall user experience, the simplicity of the narrative domain used in the present work reduces the possibilities of designing coherent story-related scenes to represent all the BFI-10 questions. For this reason, we opted here for a more straightforward solution based on the second method. That is, we directly integrate the BFI-10 questionnaire into our system to assess the user's personality.

To compute the final scores of the Big Five dimensions, we normalize the score bf_i of the i-th dimension in the interval [0, 1] instead of [1, 5], i.e.:

$$bf_i = \frac{\overline{bf_i} - 1}{4} \quad i = 1{,}5 \tag{2}$$

where

$$\overline{bf}_1 = \overline{bf}_{extraversion} = \frac{L_6 + L_1^R}{2} \tag{3a}$$

$$\overline{bf}_2 = \overline{bf}_{agreeableness} = \frac{L_2 + L_7^R}{2} \tag{3b}$$

$$\overline{bf}_3 = \overline{bf}_{conscientiousness} = \frac{L_8 + L_3^R}{2} \tag{3c}$$

$$\overline{bf}_4 = \overline{bf}_{neuroticism} = \frac{L_9 + L_4^R}{2} \tag{3d}$$

$$\overline{bf}_5 = \overline{bf}_{openness} = \frac{L_{10} + L_5^R}{2} \tag{3e}$$

and L_j and L_k^R are the Likert scale values of the true-scored item and the reversed-scored item of each dimension respectively.

Although the scores of the Big Five factors can be directly used to describe the personality of users, an important question remains: how personality can be related with narrative preferences? The proposed solution to solve this problem involves the use of machine learning techniques to ascertain the preferences for narrative content of past users based on their personality. This knowledge can then be used to predict the preferences of future users.

Considering that narrative preferences can be described as personal predilections for entire narratives or for specific scenes, there are two possible ways to formulate this problem as a machine learning problem. First, we could create a model to classify user preferences for entire plots, that is, by using each possible plot as a class in a classification problem. The second formulation takes a more in-depth approach by creating several models to classify user preferences for specific story decisions. That is, each model represents the predilections of users for the possible choices of a branching point in the story network. Again, we have a classification problem, but using the branching points' choices as classes. As will be described in Section VI, the second formulation proved to be far superior to the first one, so it is the approach that we adopted in our system.

The proposed model to map users' personalities to narrative preferences is illustrated on Figure 4. For each branching point in the story network (user decision points), we use an artificial neural network trained to predict the best choice to satisfy user's preferences. Distinct neural networks are necessary because each decision point involves completely different choices. For example, in the LRRH domain, the first branching point involves the decision of which path the girl should take to go to her grandmother's house (crossroad, forest, or uncle wolf's house), while the second branching point refers to the reaction of the girl when she arrives at grandmother's house (lay down on the bed, question the villain, or eat something). Therefore, each neural network is trained to recognize how the personality of users affects their preferences for the choices presented at each branching point.

The proposed model uses single hidden layer neural networks trained by a standard back-propagation learning algorithm using a sigmoidal activation function. The input for all neural networks comprises the five scores of the Big Five factors. Their output is defined by the possible choices available for their respective branching points. For example, the first branching point of the LRRH domain (decision of which path the girl should take to go to her grandmother's house) as three possible choices: (1) crossroad; (2) forest; or (3) uncle wolf's house. Therefore, the neural network for this branching point has three neurons in the output layer.

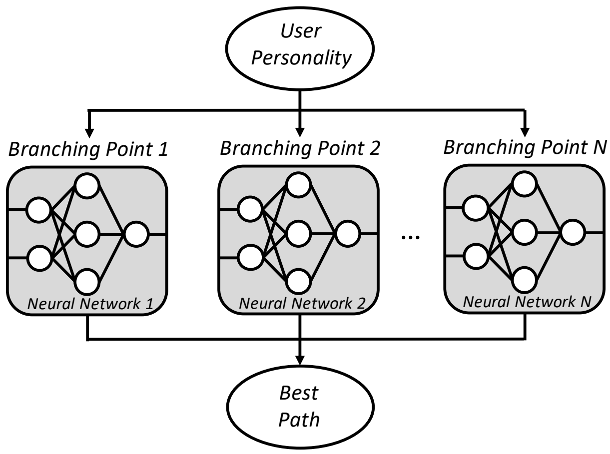

Figure 4. Model to map users' personalities to narrative preferences.

Since our method employs a supervised machine learning technique, training samples are necessary to teach the neural networks how the personality of users is related with their preferences for narrative choices. Although the process to recognize user's preferences must be executed in real-time, the training procedure can be performed offline.

The proposed method to collect training samples uses the authoring tool described in [17], which was originally designed to assist professional and non-professional writers to compose narrative variants interactively. As illustrated in Figure 5, the tool displays the network for the story domain in the lower partition of the screen. On the upper left partition, the user can create his plot by making decisions at branching points. The resulting path is automatically drawn in colors over the network. As soon as the plot is ready, a storyboard is shown on the upper right partition.

The procedure to collect training samples is divided into two steps. First, participants answer the BFI-10 questionnaire, which is integrated in the authoring tool and is displayed as soon as they start the application. Then, they use the authoring tool to create a story that they like. Participants are allowed to freely explore all possible storylines and take the time they need to find the one that best suits their personal preferences. After composing the story, the system automatically computes the scores of the Big Five factors based on the participant's answers to the BFI-10 questionnaire. The final scores and the composed plot are then stored in a text file.

After collecting the data, the training dataset for each neural network is created by separating the participants' scores of the Big Five factors and its respective decisions for each branching point. Then, this information is added to the dataset of its respective neural network as a training sample. That is, each training sample comprises 5 numerical values (scores for the Big Five factors) and a class (a number representing a choice at a branching point). They are assigned to a training dataset that encompasses all participants' decisions for a specific branching point. Not all plots include decisions for all branching points (some branching points may occur only as a consequence of specific decisions in previous branching points), therefore, a different number of training samples are expected for each dataset.

In our experiments, the training procedure was conducted with 58 computer science students with ages ranging from 17 to 26 years (mean of 19.1), who created a total of 8 different plots using the authoring tool. As the story network for the LRRH domain has 5 branching points, 5 different datasets were created to represent the participants decisions for each interaction point. The numbers of samples of the datasets are (we named each dataset to express the decision that users make at its respective branching point): (1) girl's path – 58 samples; (1) girl's reaction when arriving at grandmother's house – 38 samples; (1) girl's reaction to the disguised wolf – 38 samples; (1) wolf's action after eating the girl – 35 samples; and (1) girl's action after escaping – 50 samples.

After creating the datasets, the neural networks can be trained offline and then used to predict the narrative preferences of new users in real-time. An evaluation of the precision and performance of the neural networks is presented in Section VI.

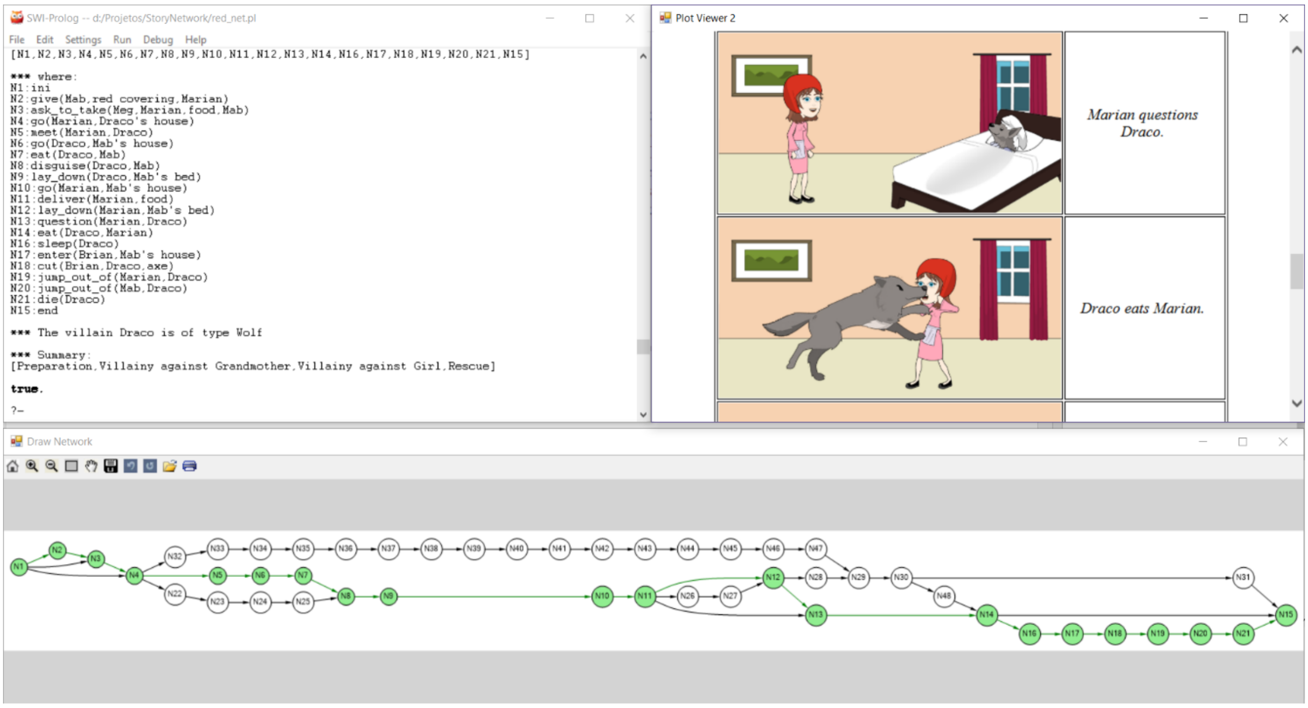

Figure 5. Authoring tool used to collect training samples: plot creation (top left), storyboard dramatization (top right), and the full story network for the LRRH domain (bottom).

V. Adaptive Storytelling

A personality and preference model can be used to adapt interactive narratives in a variety of ways depending on how users interact with the story. In interactive storytelling systems based on active user interactions, such as Façade [27] and PaSSAGE [9], wherein the user is constantly interacting with the story (in a game-like manner), the preference model can be used to change how characters react to specific situations (considering how the user prefer they react) or to bring well-suited elements of the story to the user's attention (i.e. by focusing on specific events of the story being told). For systems based on object-oriented interactions [29][30], in which the user interacts with the story indirectly by giving objects to characters or by manipulating the virtual world, the preference model can be used to adapt the effects that objects and world changes have over the story (considering the user's preference for a set of possible effects).

Perhaps an even more intuitive application of the personality and preference model is in interactive storytelling system based on passive user interactions, such as Logtell [31] and Pinter [32], where users are allowed to assume a more passive behavior without being forced to interact with the story. In this type of system, an entire storyline can be generated based on the user's preferences, which allows the user to enjoy a personalized narrative without being distracted with the interaction process.

In order to validate the proposed model, we developed a simple web-based interactive storytelling system that offers a storyboard-like comic strip representation for the generated stories, where each event gains graphical illustrations and speech balloons. The system runs on a web browser and allows users to freely scroll the horizontal comic strip to see and read the narrative. Comic panels that represent events situated at branching points in the story network include interactive thought balloons (Figure 6), where users can interact and interfere in the story by choosing the decisions made by virtual characters (indirectly selecting different branching paths to follow). As a result of user interaction, the plot and the visual representation of its events are automatically updated to reflect the new storyline. The system was implemented in Lua using the Löve 2D framework[3].

The BFI-10 questionnaire to assess the personality of users is integrated into the web page of the system. As shown on Figure 7, the questionnaire is displayed when the user accesses the web page for the first time, which allows the system to establish the personality model of the user before generating the initial plot for the story.

After establishing the personality model, the system uses the neural networks of the preference model to compute the initial decisions for all branching points, that is, it selects the best branching paths based on the output of the model. Then, the initial plot for the story is generated by traversing through the story network and using the predicted decisions to define the path to follow in the branching points. In this way, the default story presented to users when they first access the system will be the one that best matches their personal preferences for narrative content. Still, they are freely allowed to interact and explore other possible storylines when desired.

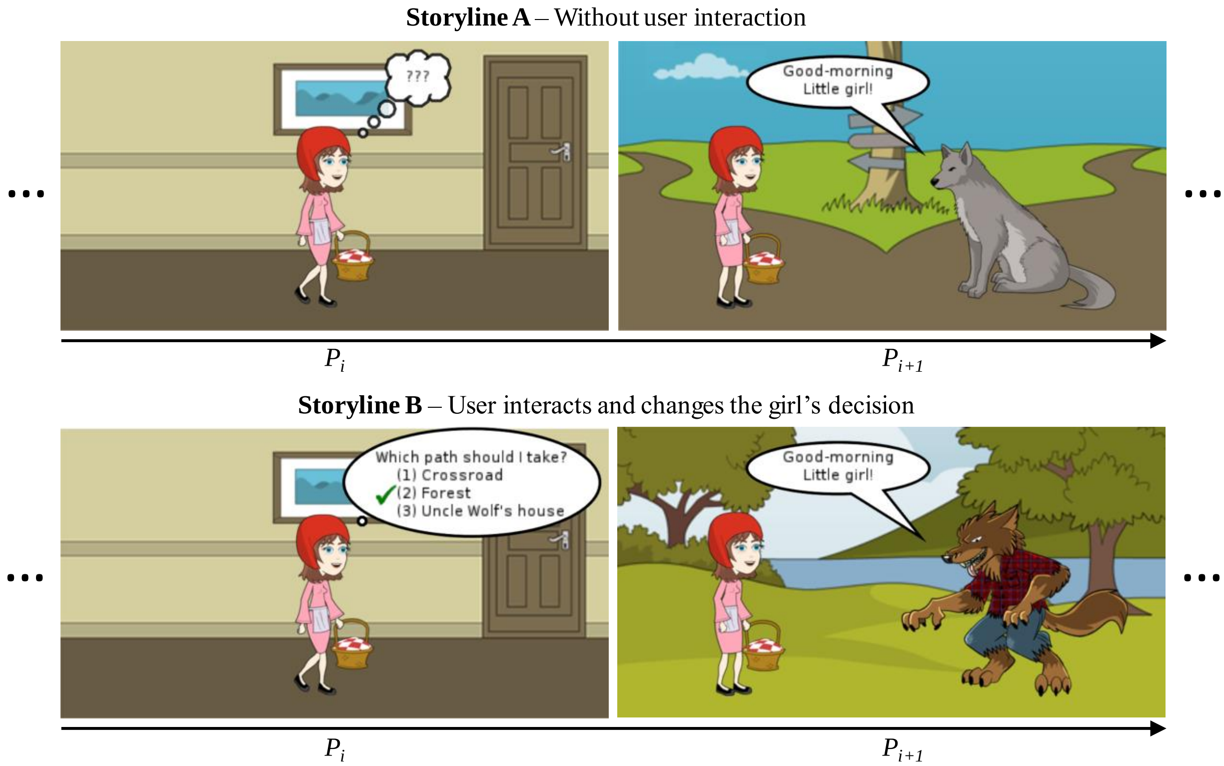

Figure 6. User interaction process: by clicking on the thought balloon, the user can change the decision made the by the virtual character.

[3] https://love2d.org/

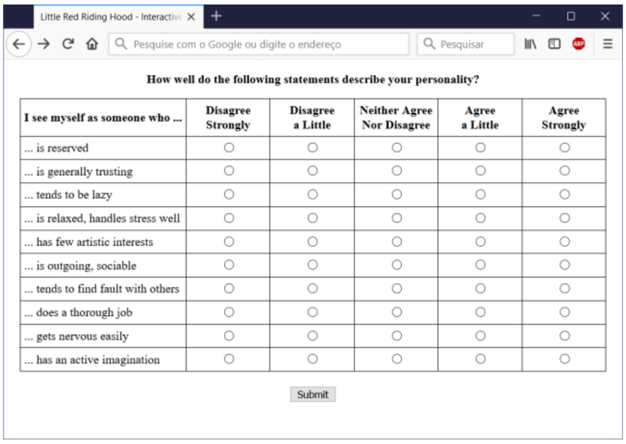

Figure 7. BFI-10 questionnaire in the web page of our system.

VI. Evaluation

As described in section IV, there are two possible approaches to create a model to map the personality of users into preferences for narrative events: (1) by classifying user's preferences for entire plots; or (2) by classifying user's preferences for specific story decisions. In order to evaluate and compare both approaches, we performed two tests: (1) a precision test to check the accuracy of the model; and (2) a performance test to evaluate the real-time performance of the neural networks used by the model. For both models, we used the data collected from 58 users, which were obtained according to the procedure to collect training samples described in section IV.

For Model 1 (to classify user's preferences for entire plots), we grouped the samples of users that created equal plots, that is, we assigned a unique class label to all samples that represent the same plot. In our experiment, the 58 users created a total of 8 different plots. Thus the dataset for this model comprises 58 samples of 8 classes, where each sample includes the user's scores for the Big Five factors (feature vector) and a class representing the plot created by the user.

For Model 2 (to classify user's preferences for specific story decisions), we separated the users' scores of the Big Five factors and the users' decisions for each branching point, and then stored this information in separate datasets (one for each branching point). Therefore, each sample comprises the user's scores for the Big Five factors (feature vector) and a class that represents a choice at a branching point. As the story network for the LRRH domain contains 5 branching points, 5 different datasets were created for Model 2. The number of samples of each dataset was presented in section IV.

In order to evaluate the precision of the models, we used the datasets of both models to train and test the neural networks. First, we divided the datasets into training and testing sets (66% of the samples were used for training and the remaining samples were used for testing). Next, we trained the neural networks of both models with the training sets and used them to predict the plot for the samples of their respective testing sets. Then, we compared the entire plots and the individual decisions at branching points generated by both approaches (for Model 2, we combined the output of all neural networks to compose the full plot; and for Model 1, we divided the full plots into individual decisions at all branching points). Following a 10-fold cross-validation strategy, this process was repeated 10 times (varying the samples used for training and testing) and then the average accuracy was calculated.

The results of the precision test are shown in Figure 8, where the bars represent the average accuracy of the models (Model 1 and Model 2) for each branching point of the story network (BP 1 to BP 5). The results indicate that Model 2 is far superior to Model 1, being able to correctly recognize the preferences of users for all story decisions in most of the cases (general average accuracy of 91.9%). Similar results are obtained when comparing the full plots produced by the combination of all story decisions: average accuracy of 39.3% for Model 1 and 81.3% for Model 2.

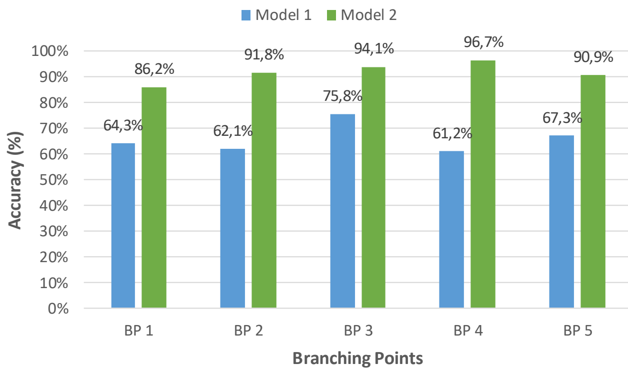

Figure 8. Average accuracy of the models (Model 1 and Model 2) for each branching point of the story network (BP 1 to BP 5).

To evaluate the performance of Model 2, we computed the average time for three tests: Test 1 to train the neural networks without trying any optimization; Test 2 to train the neural networks using precomputed models (i.e. loading the data of the neural networks from files that were created offline); and Test 3 to predict a story decision using the neural networks. The computer used to run the experiments was an Intel Core i7 7820HK, 2.9 GHZ CPU, 16 GB of RAM using a single core to process the neural networks. Each test was performed 100 times for each neural network and then the average time was calculated. Table I shows the results of the tests, which confirm the applicability of the proposed method in interactive storytelling systems without noticeable delays, especially when training the neural networks offline.

TABLE I. Average time to: train the neural networks without any optimization (Test 1); train the neural networks using precomputed models (Test 2); and predict a story decision using the neural networks (Test 3).

	Test 1	**Test 2**	**Test 3**
Time (*ms*)	663.48	1.03	0.01

VII. Concluding Remarks

In this paper we present a new approach to create personalized narrative experiences based on the personality of individual users. The proposed method can adapt interactive narratives according to the users' preferences for narrative events, which scales up the system's ability to deliver customized narrative experiences. In addition, it also caters for passive users, who are allowed to enjoy their favorite narrative without being forced to make decisions during the story.

Although the proposed method presented good results in the LRRH domain, some limitations must be pointed out. First, considering only the experiments conducted in the course of this work, we cannot guarantee that our proposed model can achieve the same success when trying to correlate users' personalities with all possible interaction choices in all story domains – therefore our next duty is to extend the experiments to other contexts. Secondly, it is important to consider that the training data used in our experiments were obtained from a small group of 58 participants with similar backgrounds, which may have increased the occurrence of subjects with similar preferences. Thus, once again, further studies are necessary to evaluate the proposed model with a significantly larger variety of subjects.

As another research objective, to be pursued while exploring other story domains, we plan to integrate a fair diversity of story-related scenes into our system, so as to avoid direct questionnaires to measure the personality of users (as done in our previous work [27]). Another envisaged future effort involves a promising task, still in the scope of the present work, which is the automatic extraction of the Big Five factors from the user's decisions in the story. Furthermore, we consider that more extensive user studies are needed to evaluate how our method affects the overall user experience of the system, which is a paramount commitment in our current research agenda.

REFERENCES

[1] D. N. Rapp, and R. J. Gerrig, "Predilections for narrative outcomes: The impact of story contexts and reader preferences," Journal of Memory and Language, vol. 54, pp. 54-67, 2006.

[2] L. R. Goldberg, "An alternative description of personality: The Big-Five factor structure," Journal of Personality and Social Psychology, vol. 59 (6), pp. 1216-1229, 1999.

[3] B. Rammstedt, and O. P. Johnb, "Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German," Journal of Research in Personality, vol. 41 (1), pp. 203-212, 2007.

[4] G. Fischer, "User Modeling in Human–Computer Interaction," User Modeling and User-Adapted Interaction, vol. 11, (1–2), pp. 65-86, 2001.

[5] A. Kobsa, "Generic User Modeling Systems," User Modeling and User-Adapted Interaction, vol. 11, (1–2), pp. 49-63, 2001.

[6] G. I. Webb, M. J. Pazzani, and D. Billsus, "Machine Learning for User Modeling," User Modeling and User-Adapted Interaction, vol. 11, (1–2), pp. 19-29, 2001.

[7] H. Barber, and D. Kudenko, "A User Model for the Generation of Dilemma-Based Interactive Narratives," Proceedings of the AIIDE 2007 Workshop on Optimizing Player Satisfaction, Stanford, California, AAAI, 2007, pp. 13-18.

[8] M. Seif El-Nasr, "Interaction, Narrative, and Drama Creating an Adaptive Interactive Narrative Using Performance Arts Theories," Interaction Studies, vol. 8 (2) pp. 209-240, 2007.

[9] D. Thue, V. Bulitko, M. Spetch, and E. Wasylishen, "Interactive Storytelling: A Player Modelling Approach," Proceedings of the 3rd Artificial Intelligence and Interactive Digital Entertainment Conference, Stanford, United States, 2007, pp. 43-48.

[10] R. Laws, Robin's Laws of Good Game Mastering. Austin: Steve Jackson Games, 2002.

[11] A. Ramirez, and V. Bulitko, "Automated planning and player modeling for interactive storytelling," IEEE Transactions on Computational Intelligence and AI in Games, vol. 7 (4), pp. 375-386, 2015.

[12] M. Sharma, S. Ontañón, M. Mehta, and A. Ram, "Drama Management and Player Modeling for Interactive Fiction Games," Computational Intelligence, vol. 26 (2), pp. 183-211, 2010.

[13] R. P. Karumur, T. T. Nguyen, and J. A. Konstan, "Personality, User Preferences and Behavior in Recommender systems," Information Systems Frontiers, 2017.

[14] E. S. Lima, A. L. Furtado, B. Feijó, and M. A. Casanova, "Towards Reactive Failure-Recovery Gameplaying: The Fall and Rise of the Grail Hero," Proceedings of the XV Brazilian Symposium on Computer Games and Digital Entertainment (SBGames 2016), São Paulo, Brazil, Sept. 2016, pp. 262-271.

[15] E. S. Lima, A. L. Furtado, B. Feijó, "Storytelling Variants: The Case of Little Red Riding Hood," Proceedings of the 14th International Conference on Entertainment Computing (ICEC 2015), Trondheim, Norway, 2015, pp. 286-300.

[16] E. S. Lima, B. Feijó, M. A. Casanova, A. L. Furtado, "Storytelling Variants Based on Semiotic Relations," Entertainment Computing, vol. 17, pp. 31-44, 2016.

[17] E. S. Lima, V. M. Gottin, B. Feijó, and A. L. Furtado, "Network Traversal as an Aid to Plot Analysis and Composition," Proceedings of the XVI Brazilian Symposium on Computer Games and Digital Entertainment (SBGames 2017), Curitiba, Brazil, 2017, pp. 418-427.

[18] N. Peever, D. Johnson, and J. Gardner, "Personality & Video Game Genre Preferences," Proceedings of The 8th Australasian Conference on Interactive Entertainment, Auckland, New Zealand, 2012, Article No. 20.

[19] S. H. Hsu, C. H. Kao, and M. C. Wu, "Factors influencing player preferences for heroic roles in role-playing games," Cyberpsychology & Behavior, vol. 10 (2), pp. 293-295, 2007.

[20] B. Rammstedt, and O. P. Johnb, "Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German," Journal of Research in Personality, vol. 41 (1), pp. 203-212, 2007.

[21] S. D. Gosling, P. J. Rentfrow, and W. B. Swann Jr, "A very brief measure of the Big-Five personality domains," Journal of Research in Personality, vol. 37 vol. 504-528, 2003.

[22] G. Guido, A. M. Peluso, M. Capestro, and M. Miglietta, "An Italian version of the 10-item Big Five Inventory: An application to hedonic and utilitarian shopping values," Personality and Individual Differences, vol. 76, pp. 135-140, 2015.

[23] R. Carciofo, J. Yang, N. Song, F. Du, and K. Zhang, "Psychometric Evaluation of Chinese-Language 44-Item and 10-Item Big Five Personality Inventories, Including Correlations with Chronotype, Mindfulness and Mind Wandering," PLoS ONE, vol. 11 (2), pp. 1-26, 2016.

[24] B. de Raad, and M. Perugini, Big Five assessment, Boston: Hogrefe & Huber Publishers, 2002.

[25] P. T. Costa, and R. R. McCrae, "Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual," Psychological Assessment Resources, 1992.

[26] S. D. Gosling, P. J. Rentfrow, and W. B. Swann, "A very brief measure of the Big-Five personality domains," Journal of Research in Personality, vol. 37 (6), pp. 504-528, 2003.

[27] E. S. Lima, B. Feijó, and A. L. Furtado, "Player Behavior and Personality Modeling for Interactive Storytelling in Games," Entertainment Computing, vol. 28, pp 32-48, 2018.

[28] M. Mateas, "Interactive Drama, Art, and Artificial Intelligence," Ph.D. Thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, 2002.

[29] E. S. Lima, B. Feijo, S. D. J. Barbosa, A. L. Furtado, A. E. M. Ciarlini, and C. T. Pozzer, "Draw Your Own Story: Paper and Pencil Interactive Storytelling," Entertainment Computing, vol. 5 (1), pp. 33-41, 2014.

[30] E. S. Lima, B. Feijo, A. L. Furtado, S. D. J. Barbosa, C. T. Pozzer, and A. E. M. Ciarlini, "Non-Branching Interactive Comics," Proceedings of the International Conference on Advances in Computer Entertainment Technology (ACE 2013), Enschede, Netherlands, 2013, pp. 230-245.

[31] M. M. Camanho, A. E. M. Ciarlini, A. L. Furtado, C. T. Pozzer, and B. Feijó, "A Model for Interactive TV Storytelling," Proceedings of the VIII Brazilian Symposium on Digital Games and Entertainment, Rio de Janeiro, Brazil, 2009, pp. 197-206.

[32] A. Gilroy, J. Porteous, F. Charles, M. Cavazza, "Exploring passive user interaction for adaptive narratives," Proceedings of the 2012

ACM international conference on Intelligent User Interfaces, Lisbon, Portugal, 2012, pp. 119-128.

[33] C. Perrault, "Little Red Riding Hood," In: Beauties, Beasts and Enchantment - Classic French Fairy Tales. J. R. Planché and J. Zipes (trans.). England: Meridian, 1991.

[34] J. Grimm, W. Grimm, The Complete Grimm's FairyTales. M. Hunt and J. Stern (trans.), New York: Pantheon, 1972.

[35] P. Delarue, "The Story of Grandmother," In Little Red Riding Hood: A Casebook. Dundes, A. (ed). Wisconsin: University of Wisconsin Press, 1989.

[36] I. Calvino, Italian Folktales. Boston: Mariner Books, 1992.

[37] R. McCrae, and P. Costa, "Personality trait structure as a human universal," The American Psychologist, vol. 52 (5), pp. 509-16, 1997.

[38] M. Back, and B. Egloff, "Yes we can! A plea for direct behavioral observation in personality research," European Journal of Personality, vol. 23, pp. 403-405, 2009.

[39] G. Kraaykamp, and K. V. Eijck, "Personality, media preferences, and cultural participation. Personality and Individual Differences," vol. 38 (7), pp. 1675-1688, 2005.