Narrating Stories in Participatory Games

Paula S.L. Rodrigues[1] Bruno Feijó[1] Luiz Velho[2]
Cesar T. Pozzer[3] Ângelo E. M. Ciarlini[4] Antonio L. Furtado[1]

[1] VLab/IGames, Dept. of Informatics, PUC-Rio, Rio de Janeiro, Brazil
[2] IMPA, Institute of Pure and Applied Mathematics, Rio de Janeiro, Brazil
[3] Federal University of Santa Maria, Brazil
[4] Dept. of Applied Informatics, UniRio, Rio de Janeiro, Brazil

Figure 1: Virtual character narrating a dragon-sword story in which the player participates as a co-author.

Abstract

Interaction and participation are in the kernel of the new medium of games. The ultimate goal is the "participatory game", where interactive games and storytelling are merged. One of the most complex forms of this type of game is the narrated ones. This paper presents an architecture that incorporates a virtual narrator, capable of emotional expressions synchronized with speech, to an interactive storytelling system, in order to create a new form of participatory game, in which the player is co-author of the plot.

Keywords: dramatic games, interactive storytelling, virtual narrator, participatory games

Authors' contact:
[1]`{paula,bruno, furtado}@inf.puc-rio.br`
[2]`lvelho@impa.br`
[3]`pozzer@inf.ufsm.br`
[4]`angelo.ciarlini@uniriotec.br`

1. Introduction

People are experiencing the birth of a new medium, which is not fully comprehensible. The situation is similar to that of the early days of cinema, when it was difficult to understand the new concept of kinematics in a reproduction medium. Today people are trying to understand the new concepts of interaction and participation found in the kernel of the new medium of games. By far, in this scenario, participation is the less known concept.

The ultimate goal of gaming is "participation", which goes beyond interaction. The key for participation is "storytelling" and has its roots in the quest of "entering the story" that was dreamed by Brenda Laurel in the early 80's and Janet Murray in the late 90's [Laurel, 1996] [Murray, 1998]. In fact, story reaches inside of us and reveals the world through emotions, surprises, and ongoing experiences. In this paper, we call that ultimate goal of "participatory games", where interactive games and storytelling are merged. Participatory games can be realized in different forms, such as action, adventure, simulation, and drama. Nowadays games are not participatory ones. In current games, elements of games and stories are separated (or weakly coupled), such as the movie-like scenes during game openings and/or between each game mission. On the other side of the spectrum, current storytelling systems are not interactive games.

The concept of participatory games is the same of interactive storytelling presented by [Glassner, 2004]. However, we prefer the term "participatory game" to emphasize the focus on games rather than on storytelling. In this paper, we use the term "interactive storytelling" when the focus is on story generation and dramatization. Participatory games can also be identified in the new concept of alternate reality gaming (ARG), which blends real-world activities and a dramatic storyline [Szulborski, 2005] [Borland, 2005].

One of the most complex forms of participatory game is the narrated ones. In this form, as in traditional live storytelling, interaction and participation are deeply explored. This paper is about virtual narrators in participatory games.

Traditional live storytelling is an interactive performance art form, wherein the teller adjusts the vocalization, wording, physical movements, gestures, and pace of the story to better meet the needs of the responsive audience. Storytelling in its new digital and interactive form combines participation, as occurs in computer games, with automatic story generation and narration. Different storytelling systems have been proposed and implemented with different focus and features [Cavazza et al. 2002] [Mateas, 1997] [Young, 2000] [Spierling et al. 2002] [Sgouros, 1999] [Ciarlini et al. 2005]. Although the presence of a synthetic narrator should be a welcome enhancement to the digital storytelling experience, the existing literature has not duly explored this subject. Research works on digital actors [Thalmann and Thalmann], graphical multimodal user interfaces [Corradini et al. 2005] [Cassell et al. 1999] [Massaro, 2003], and facial animation [Parke and Waters] do not address the question of synthetic narrators in interactive storytelling.

This paper describes the incorporation of a virtual narrator, capable of emotional expressions synchronized with speech, to the LOGTELL storytelling system [Ciarlini et al., 2005], in order to create a new form of participatory game. In this form of game, the player is co-author of the plot. Despite the rudimentary implementation of the prototype, this paper presents an innovative approach to gaming. In the proposed system, speech and facial animation techniques were combined with plot generation, user interaction and 3D dramatization, in order to better communicate the story, increase dramatic potential and help user interaction.

2. Related Work

In the literature, there is no work where interaction, storytelling, and virtual narrators are treated together towards a new form of digital entertainment. The works that are most related to the present paper are discussed below.

The main focus of [Silva et al., 2001] is on automatic plot creation, but without any kind of user interaction. That paper describes a virtual storyteller framework, where plots are not predefined but created by the actions of the characters, under the guidance of a virtual director. The virtual director is a separate agent who has general knowledge about plot structure. Both the characters (or "actors") and the director are implemented as intelligent agents, capable of reasoning within their own domain of knowledge. The characters can make plans to achieve their personal goals using story-world knowledge, *i.e.* knowledge about their virtual environment and the actions they can perform. The director is able to judge whether the intended action of a character fits into the plot structure, using both story-world knowledge and general knowledge about what constitutes a "good" plot. However, the virtual director is not endowed with any kind of speech; it uses text balloons to present the narrative. Another limitation is the absence of emotional facial expressions, since the narrator always presents the same behavior and attitude.

[Theune et al., 2003] describe a storytelling system with semi-autonomous agents where a synthetic virtual narrator character reads an input text enriched with control tags. Those tags allow the storywriter to control the character's emotional state and the behavior of the surrounding environment. That work has several drawbacks: the 3D scenarios are rather limited; the system is not interactive; and the narrator is a presentation agent implemented as a Microsoft MSAgent [Microsoft, 2003] with no emotional facial expressions and no lip synchronization.

Finally, there is a group of interesting facial animation systems that are not associated with any kind of storytelling system [Zhang et al., 2003] [Bui et al., 2004] [Pandzic and Forcheimer]. They are correctly defined as facial animation tools, but some of them [Zhang et al., 2003] provide no speech treatment and the characters have limited emotional expressions. Avatars based on the MPEG-4 standard [Pandzic and Forcheimer] have the potential of being used as virtual narrators in storytelling systems and participatory games. Unfortunately the MPEG-4 facial animation framework suffers from a limitation: the system loses portability and platform independence, because the framework requires an encoder and a decoder to propagate (send and receive) streams of MPEG-4 facial animation. [Bui et al., 2004] present a 3D talking head system with speech synchronization. Their work seems to have the necessary infrastructure to become a virtual narrator for a storytelling system, with features similar to those presented in our paper, but it is still under development. Furthermore, that work has no such a robust interactive plot generation module like the LOGTELL module.

3. Storytelling and Face Simulation

A key point in the implementation of a storytelling system is whether it should be "character-based" or "plot-based". In a character-based approach, as described in [Cavazza et al. 2002], the storyline usually results from the real-time interaction, at any time, between virtual autonomous agents and the user. Although powerful in terms of interaction and variety of stories, such an extreme interference level may lead the plot to unexpected situations or miss essential predefined events. In contrast, in a plot-based approach, as in [Spierling et al. 2002], characters should follow rigid rules specified by a plot. By doing this, the coherence of the story might be guaranteed but the level of interaction is reduced.

LOGTELL [Ciarlini et al., 2005] combines both plot- and character-based features. It is based on the logic specification of a model of the chosen story genre, where possible actions and goals of the characters are described. Plot generation and 3D dramatization are integrated but separately treated. During dramatization, virtual actors perform the events in the plot without user (player) interference. Nevertheless, the user can alternate phases of plot generation, in which intervention is possible, and phases of dramatization. In this way, the two requirements are met: the generated stories are always coherent and we are not limited to a small group of predefined alternatives. This is a new form of gaming, reacher than standard Role Playing Games, where the player is co-author and experiences a glimpse of what Interactive TV/Cinema should be in future.

In this new form of digital entertainment, every intended plot alternative can be obtained using a combination of simulation and user interaction, provided that it is in accordance with the logics of the genre. The use of a virtual narrator in such an environment provides a very convenient way to explain the chaining of events, entailed by the conventions of the genre, and to transmit the emotion associated with each event.

In this paper, the virtual narrator is an expressive talking head implemented by a facial animation module ETH (Expressive Talking Head). This module receives markup-texts containing story fragments and produces, on the fly, a facial animation that gives voice to this input text. The speech is automatically generated using text-to-speech (TtS) mechanisms. The ETH module controls the lip synchronization and the emotional expressions, which are obtained through the text markup parameters.

The proposed environment has the capability of building, presenting and narrating different stories from different genres. The example shown in this paper is based on a Swords-and-Dragons context, where heroes, victims, and villains interact in a 3D scenario occupied by castles and churches. The narrator's primary duty is to tell, from a third person perspective and with the appropriate emotion, each scene of the story. The story can be changed at any moment through the intervention of the user (player). The prototype has problems with real-time performance when deep changes require extensive planning operations. However, the real-time issue is part of the ongoing research by the authors.

3.1 The Interactive Storytelling Module

LOGTELL [Ciarlini et al., 2005] is based on modeling and simulation. The idea behind LOGTELL is to capture the logics of a genre through a temporal logic model, which is able to guide the generation of story plots by simulation combined with user intervention. Therefore we focus not simply on different ways of telling stories, but on the dynamic creation of plots. The model is composed by typical events and goal-inference rules.

Our stories are told from a third-person viewpoint and the user intervention is always indirect. During the simulation, the user (player) can intervene either passively (weak intervention), just allowing the partially-generated plots that seem interesting to be continued, or, in a more active way (strong

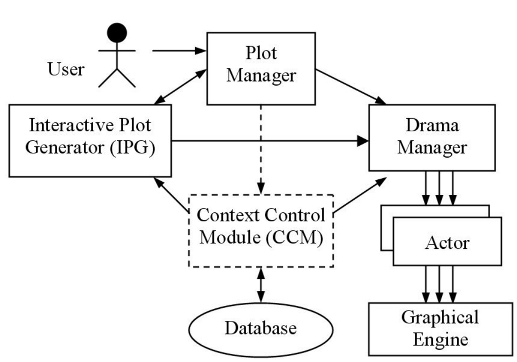

Figure 2: LOGTELL's architecture (arrows represent dataflow)

intervention), trying to force the occurrence of desired events and situations. These are rejected by the system whenever there is no way of changing the story to accommodate the intervention. Plot dramatization can be activated for the exhibition of the final plot or the partial plots. During dramatization, characters are represented by actors in a 3D-world. During the performance of an event, low-level planning is used to detail the tasks involved in each event. In order to integrate dramatization and plot generation, we decided to implement our own graphical engine, so that we could guarantee the compatibility between the logical model of our plots and the corresponding graphical dramatization.

LOGTELL comprises a number of distinct modules to provide support for generation, interaction (management), and 3D dramatization of interactive plots, as shown in Figure 2.

The Interactive Plot Generator (IPG) of Figure 2, implemented in Prolog, performs simulations using the context specified by the user. The context contains: (a) the possible types of events, associated with operations which are specified in terms of their pre- and post-conditions; (b) a set of goal-inference rules, formulated in a temporal modal logic, to specify the goals the characters will want to pursue when certain situations are detected during the plot; and (c) the initial configuration, describing the various characters at the beginning of the story. The generation of a plot starts by inferring one or more goals motivating the characters at the initial configuration. Given this input, the system uses a non-linear planner that inserts the first events in the plot, in order to enable the characters to try to fulfill their goals. When the planner detects that all goals have been either achieved or abandoned, the first stage of the process is finished. If the user does not like the partial plot, IPG can be asked to generate another alternative. If the user accepts the partial plot, the process continues by inferring new goals from the situations holding as a result of this first stage. The process alternates goal-inference, planning, and user interference, until the user decides to stop or a state is reached wherein no new goal is inferred.

The Plot Manager of Figure 2 comprises the user graphical interface (in Java), through which the user can participate in the choice of the events that will figure in the plot and decide on their final sequence. The selection of alternative compositions of events and the choice of a particular sequence (according to which the events will be exhibited) correspond to the "weak" way of user intervention. We should notice that IPG generates events in a partial order, determined by the chaining of pre- and post-conditions, but the current version of the 3D dramatization expects a totally ordered sequence. Stronger ways of interventions are also possible. The Plot Manager has commands to force the insertion of events and situations, seen by IPG as additional goals. Such strong interactions are however subject to validation by IPG, which tries to conciliate user interventions and the logic requirements of the genre (what is not currently possible to be done in real time). At any intervention phase, the user can:

(a) order IPG to continue the generation process; (b) query IPG to obtain details about the situation of the story, such as the state of specific characters when a certain event occurs; or (c) order the Drama Manager to dramatize the events inserted so far.

The Drama Manager (Figure 2) is responsible for the dramatization of the plot. The Drama Manager translates symbolic events into fully realized 3D visual graphical animations, guaranteeing the synchronism and logical coherence between the intended world and its graphical representation. As received from IPG, the plot is organized as a sequence of events, each one associated with a discrete time instant, and their effects are supposed to occur instantaneously. For the purposes of visualization, a different concept of time is used. The simulation occurs in continuous real-time and the duration of an event rendering task is not previously known. The values of certain variables change as the event is dramatized. In order to conciliate logical and graphical representations, the values of such variables before the dramatization of each event must conform to the pre-conditions of the event, and the values at the end to its post-conditions. During graphical representation, all control of actions each actor is supposed to perform is done by the Drama Manager. It acts as a director, responsible for coordinating the sequences of linear or parallel actions performed by the whole cast. It continuously monitors the representation process, activating new tasks whenever the previous ones have been finished. As a director, it also controls the positioning of the (virtual) camera, which an option of LOGTELL permits to be transferred to the user.

3.2 The Emotional Facial Animation Module

The Expressive Talking Head (ETH) module is a facial animation system which, upon receiving an input text with some special markups, is able to generate a real-time character facial animation to say this text. The ETH module is developed to provide a framework for applications wherein a talking head unit may be desirable. Some applications have already been developed using this framework, such as the integration with a hypermedia presentation system [Rodrigues et al., 2004].

The ETH module is composed by three major submodules: Input Synthesis, Face Management, and Synchronization, as illustrated in Figure 3 (at the end of this paper).

The Input Synthesis submodule is responsible for two tasks: (a) to capture and treat the input text, which have the following information: some markup elements conveying information about the character's emotion, the voice gender (masculine or feminine), and the speech language (American English or British English); (b) to generate an output that is a data structure containing the fundamental units (phonemes, duration, emotion, etc.) needed to build the facial animation corresponding to the input text. The Input Synthesis module has two secondary submodules: a parser, responsible for separating the speech content itself (text without markups) from the speech and animation markups, and the TtS (Text to Speech) submodule.

The parser interacts with the TtS submodule to build the facial animation and lip-sync data structures sending to the TtS each fragment of the input markup text. The TtS submodule is made up of two independent subsystems: Festival [Black and Taylor] and MBROLA [Dutoit and Pagel], as shown in Figure 3. In this blend of two synthesizers, Festival works as the Natural Language Processing unit (NLP), being responsible for the speech phonetic description creation (list of phoneme entries, each one containing the phoneme label, duration and pitch); while MBROLA works as the Digital Signal unit (DSP), in charge of generating the final output audio file.

The second ETH main submodule, the Face Management, is connected to an external subsystem, named Responsive Face [Perlin, 1997], which defines a three-dimensional polygonal mesh. The face is animated by the application of relax and contract commands over the mesh edges (face muscles). The ETH module improves the Responsive Face [Perlin, 1997] features by adding the concept of visemes. Viseme is the name given to a mouth configuration for a specific phoneme. When initializing the system, the Face Management submodule builds a table of 16 visemes and 7 facial expressions (natural, frightened, angry, happy, annoyed, disappointed, and surprised). Each table entry stores the values for contracting/relaxing the face corresponding muscles commanding the Responsive Face. The Face Management submodule also builds a table defining the phoneme-viseme mapping.

The third and last ETH submodule is the Synchronization one, which is responsible for the fine synchronization between speech and facial muscle movements. Parallel to the audio file reproduction, the synchronizer polls the audio controller to check the effective playing instant. Using the information in the animation data structure, the Synchronizer discovers the current phoneme and the current character emotion. Then it asks the Face Manager to animate the face in order to achieve the corresponding viseme and facial emotion. The Synchronization submodule also includes components to control the movements of the head and of the eyes, so as to produce a more natural output.

4. The Virtual Narrator Environment

The integration of the ETH module with the LOGTELL module adds two extra dimensions to plot dramatization: the narrator perspective and the assistant one. The narrator perspective renders the animated plot into a genuine storytelling experience. The other perspective is the use of the narrator as a virtual assistant of the author during the specification and revision of the story genre, the detailed composition of the story plot, and the modification of an already-written plot. The assistant perspective is not fully implemented in the present work. Figure 4 (at the end of this paper) illustrates the communication between the LOGTELL and ETH modules in the virtual narrator environment.

In any perspective, the ETH module provides an additional medium to communicate information, by means of a live audio synchronized with a 3D emotive virtual narrator. During plot generation, the narrator can be used to complement what is presented, perhaps too concisely, in dialog text boxes. During dramatization, the virtual narrator can be used not only to read aloud the subtitles narrating the current action, but also to explain what is happening and reveal what lies "behind the scene". This is possible because all metadata (*i.e.* the internal definition of the genre, especially the pre-conditions and post conditions of operations and the goal-inference rules) stay available at runtime. In particular, we should point out that the ability of expressing emotions during dramatization is essential to increase the dramatic potential of the story. The result is a more engaging experience with a better comprehension of the story by the player-spectator.

The complementary explanation provided by the ETH module can be either produced in real-time, or pre-synthesized and later inserted in the appropriate context. The real-time strategy favors the necessary flexibility to offer assistance during plot generation. The user might want, for instance, to query IPG about details of a specific character at a certain point in the story. In this case, the answer to be given by the narrator should be generated at runtime. On the other hand, when Dramatization is activated, since (part of) the story to be told does not change, parallelism can be used to provide pre-synthesized speech for the next events while a previous one is being shown. In this way, CPU processing time can be saved and more attention can be paid to information contents, communicative efficacy and stylistic quality.

4.1 Graphical and Narrator Output

The graphical engine supports real-time rendering of the 3D elements, under the control of the Drama Manager. Characters in a generated plot are, as remarked before, treated as actors for the dramatization. The Drama Manager acts then as a 'director', without having to perform any intelligent processing with respect to plot generation. It essentially follows the ordered sequence of events generated at preceding stages of simulation and interaction. Each actor is implemented as a materialized reactive 3D agent, with minimal planning capabilities necessary to play the assigned role within an event. The Drama Manager controls, from a third-person perspective, the scene and the current actors' aspect and movements. Steering behaviors [Reynolds, 1999] are used by the actors for real-time interactions with the scene and, occasionally, with other actors.

When accompanying dramatization, the virtual narrator is responsible for synchronously narrating the ongoing actions being performed by the actors. In our test scenario, we use a small subclass of the popular Swords-and-Dragons genre. The participants are a Princess, called Marian (the potential victim), Draco the dragon as a villain, and two heroes, the knights Brian and Hoel. Currently, we make use of simple templates (Prolog lists intercalating variables and fixed character strings) to translate the formal terms denoting the events into the natural language sentences that are used for narration. Examples of the subtitles automatically generated by the system are: 1) The protection of the Princess's castle is reduced; 2).Draco kidnaps Marian; 3) Brian kills Draco; 4) Brian frees Marian; and 5) Brian and Marian get married.

Since the rendering duration of most of the actions can be previously ascertained, usually taking no less than 10 seconds, the virtual narrator has enough time to describe the events being dramatized and to add relevant contextual information. This extra material can be readily extracted by IPG, which has access to a number of data, such as: the properties of the characters, the places (at each state) reached by the plot simulation, and the logical specification of the genre. The logical chaining of events, determined by specified causes, effects and goals, is an essential part of the narration. As an example, consider the abduction of the victim (Princess Marian) by the villain (Draco). A pre-condition for this event is the fragility of the victim and, as post condition, the kidnapped princess is confined to the villain's castle. What ultimately motivates the event is the villain's goal of kidnapping an unprotected victim. Since one common heroic goal is to free any damsel in distress, the kidnapped victim's situation arouses in the hero the desire (goal) of rescuing her. The simulated execution of a plan to achieve this goal leads, in turn, to a new state wherein other goals are inferred, thus causing the story to move forward. We have already implemented a text generation module that generates this kind of explanation, and are still working on text stylistic improvements to better incorporate the text generation feature into the environment.

The ETH module is responsible for dialog synthesis, in real time, and also for handing over the speech audio and the phoneme sequences to be spoken in a synchronized way. For each phoneme, there is an associated viseme, and to visualize the viseme the narrator facial muscles are moved, as mentioned in Section 3. Each event-producing operation in the story is hitched with an emotion. This emotional information is used by the virtual narrator in the exact instant when it tells the story. On the other hand, it knows that, for each word, sentence or paragraph, there is a facial expression. Internally, operations must somehow be mapped into emotions, as indicated in Table 1.

Table 1: Operation-and-Emotion Mapping

Operation(*)	Emotion
`go(CH,PL)`	natural
`reduce_protection(VIC,PL)`	annoyed
`kidnap(VIL,VIC)`	frightened
`attack(CH,PL)`	surprised
`fight(CH1,CH2)`	angry
`kill(CH1, CH2)`	angry
`free(HERO, VIC)`	happy
`marry(CH1, CH2)`	happy

(*) CH is a character, PL is a place, VIC is a victim, VIL is a villain, and HERO is a hero.

Another important enhancement for telling the story is the description of effects of the actions in a more dramatic way, conveying the appropriate emotion. The event "Brian frees Marian" has a side-effect that is essential for understanding the story: the level of affection of the princess for the hero is raised to 100. With a simple conditional template, in Prolog notation, such as:

```
template(affection(CH1,CH2, Level),
  [CH1, 'feels now ', Aff, ' for ', CH2]):-
    (Level = 0, !, Aff = 'absolutely nothing';
    Level =< 50, !, Aff = 'a moderate liking';
    Level =< 99, !, Aff = 'some tenderness';
    Level = 100, !, Aff = 'a perfect love').
```

a sentence, with appropriate emotion tags, would be sent to the virtual narrator, inducing it to comment, with a happy smile: "Marian feels now a perfect love for Brian".

In addition to speech, it is also possible to incorporate a background music line. The music can play throughout different narration phases, reflecting the varying emotions associated with the events. Figure 5 (at the end of this paper) portrays the visual aspect of the environment.

4.2 Implementation Issues

When a user requests plot dramatization, each event is processed following the connected sequence drawn by the user in the Plot Manager interface. The dramatization process involves the delivery of all specific data associated with the current event to both the Drama Manager and the narrator. In order to do that, for each individual event, the Drama Manager initially consults the IPG module to obtain the information required for describing the event, including subtitles and dialogs. The Drama Manager determines when an event dramatization has been finalized, and, in this case, requests a new one from the Plot Manager. All modules are implemented in Java, except the Drama Manager, which is implemented in C++/OpenGL.

The user may select whether he would like to see the story with 3D scenes and a narration, or only with 3D scenes. The former is the default option, with both visual and speech narrations. If the purely visual option is chosen, the narrator is simply not created.

5. Conclusion

In this paper, we present the main concepts and strategies of a 3D interactive environment for story generation and dramatization using an expressive avatar to augment user immersion and emotional experience. These results create a new form of participatory game, in which the player is co-author of the plot and provide a glimpse of what might be the interactive TV/Cinema of the future.

The flexibility of the system for incorporating different kinds of modules increases its ability to cope with an ample variety of applications. In fact, besides its usage in participatory games, the proposed system can be adapted and applied in many other areas such as authoring systems, business training, news presentation, distance learning, and e-commerce [Ciarlini and Furtado].

The environment presented in this paper has two main components: a plot-based storytelling module, called LOGTELL, and an expressive talking head module, called ETH. In the proposed environment, the talking head is used as a story narrator, integrated with a 3D rendering module, with voice output generated on the fly. Moreover, the avatar exhibits emotional facial expressions in order to enhance the user perception during storytelling. As far as the authors are awared, there is no other work in the literature with both aspects of a logic-based plot generation model and an emotional and expressive virtual narrator.

A planned extension to the proposed system is the modification of the text generation module to include stylistic improvements combined with automatic generation of emotion tags. Another issue currently under investigation by the present authors is on how the narrator's capabilities can be fully used to cooperate with the user during plot generation. Another future work is to investigate how to assign roles to the narrator as an emotional and reactive actor in the story. Also a model of emotion is currently being investigated by the authors to support many of the above-mentioned future features. An intriguing possibility for further research is to have avatars (working as narrators or actors) interacting vocally with the user, helping him/her to conduct the story. In this kind of interaction, a face recognition system (with cameras) is being investigated in order to link the virtual narrator to the audience response. Experiments on 3D stereo environments is also under investigation. Finally, better models for eyes and head movements are being investigated by the present authors.

Acknowledgements

The authors would like to thank CNPq for their individual grants and FINEP for the research contract 01.04.0945 Ref. 3110/04.

References

BLACK, A. AND TAYLOR, P., 2004. Festival Speech Synthesis, Software Package, version 2.0.

BORLAND, J., 2005. *Blurring the line between games and life* [online] CNET News.com. Available from: http://news.com.com/Blurring+the+line+between+games+and+life/2100-1024_3-5590956.html [Accessed 27 August 2006].

BUI, T., HEYLEN, D., NIJHOLT, A., 2004. Combination of facial movements on a 3d talking head. In Computer Graphics International, 284-291.

CASSELL, J., BICKMORE, T. W., BILLINGHURST, M., CAMPBELL, L., VILHJALMSSIN, H. H., YAN, H. 1999. Embodiment in conversational interfaces: Rea, In Proceedings of CHI, 520-527.

CAVAZZA, M., CHARLES, F., MEAD, S., 2002. Characterbased interactive storing, IEEE Intelligent Systems, Special Issue on AI in Interactive Entertainment 17(4), 17-24.

CIARLINI, A. AND FURTADO, A. L., 2002. Understanding and simulating narratives in the context of information systems. In Proceedings of the 21st. International Conference on Conceptual Modeling – ER'2002, 291-306.

CIARLINI, A., POZZER, C. T., FURTADO, A. L., FEIJÓ, B., 2005. A logic-based tool for interactive generation and dramatization of stories, ACM SIGCHI International Conference on Advances in Computer Entertainment Technology – ACE 2005, Valencia, Spain, 133-140.

CORRADINI, A., MEHTA, M., BERNSEN, N., CHARFUELAN, M., 2005. Animating an interactive conversational character for an educational game system. In Proceedings of the 10th Int. Conf. on Intelligent User Interfaces, 183-190.

DUTOIT, T., PAGEL, V., 1996. The MBROLA Project. Available from: http://www.tcts.fpms.ac.be/synthesis [Accessed 28 August 2006].

GLASSNER, A, 2004. Interactive Storytelling: Techniques for 21st Century Fiction. AK Peters, Ltd.

LAUREL, B, 1996. Computers as Theatre. Addison-Wesley Professional, Reprinted edition.

MASSARO, D. W., 2003. A computer-animated tutor for spoken and written language learning. In Proceedings of the 5th. Int. Conf. on Multimodal Interfaces, ICMI 2003, 173-175.

MATEAS, M., 1997. An Oz-Centric Review of Interactive Drama and Beliavable Agents. Technical Report, School of Computer Science, Carnegie Mellon University, Pittisburgh.

MICROSOFT, 2003. Microsoft Agent. Available from: http://www.microsoft.com/msagent/default.asp [Accessed 28 August 2006]

MURRAY, J.H., 1998. Hamlet on the Holodeck: the Future of Narrative in Cyberspace. The MIT Press.

PANDZIC, I. AND FORCHEIMER, R., 2002. MPEG-4 Facial Animation: The Standard, Implementation and Applications. ISBN: 0-470-84465-5.

PARKE, F.L. AND WATERS, K., 1996. Computer facial animation. Wellesley: AK Peters.

PERLIN, K., 1997. Responsive Face. Technical Report, Media Research Lab, New York University, USA. Available from: http://mrl.nyu.edu/˜perlin/demox/Face.html [Accessed 28 August 2006]

REYNOLDS, C.W., 1999. Steering behaviors for autonomous characters. In Proceedings of Game Developers Conference, San Jose, 763-782.

RODRIGUES, R.F., LUCENA-RODRIGUES, P.S., FEIJÓ, B., VELHO, L., SOARES, L.F.G. 2002. Cross-media and elastic time adaptive presentations: the integration of a talking head tool into a hypermedia formatter. In Adaptive Hypermedia and Adaptive Wev-Based Systems, Eindhoven, 215-234.

SGOUROS, N., 1999. Dynamic generation, managing and resolution of interactive plots. Artificial Intelligence 107, 29-62.

SILVA, A., VALA, M., PAIVA, A., 2001 Papous: the virtual storyteller. In Proceedings of the Intelligent Virtual Agents, 3rd. Int. Workshop – IVA 2001, Madrid, Spain, Lectures Notes in Computer Science, vol. 2190, 171-180.

SPIERLING, U., BRAUN, N., IURGEL, I., GRASBON, D., 2002. Setting the scene: playing digital director in interactive storytelling and creation. Computer and Graphics 26, 31-44.

SZULBORSKI, D., 2005. This is not a Game: a Guide to Alternate Reality Gaming. 2nd digital ed., Lulu Press.

THALMANN, N. M. AND THALMANN, D., 1995. Digital actors for interactive television. In Proceedings of the IEEE Special Issue on Digital Television, Part 2, vol. 83, 1022-1031.

THEUNE, M., FAAS, S., NIJHOLT, A., HEYLEN, D., 2003. The virtual storyteller: Story creation by intelligent agents. In Technologies for Interactive Digital Storytelling and Entertainment (TIDSE) Conference, 204-215.

YOUNG, R., 2000. Creating interactive narrative structures: The potential for AI approaches. In the Working Notes of the AAAI Spring Symposium in Artificial Intelligence and Interactive Entertainment, Stanford, CA. Available from: http://liquidnarrative.csc.ncsu.edu:16080/pubs/potential.pdf [Accessed 27 August 2006]

ZHANG, Q., LIU, Z., GUO, B., SHUM, H., 2003. Geometry Driven photorealistic facial expression synthesis. In 2003 ACM/SIGGRAPH/Eurographics Symposium on Computer Animation, 177-189.

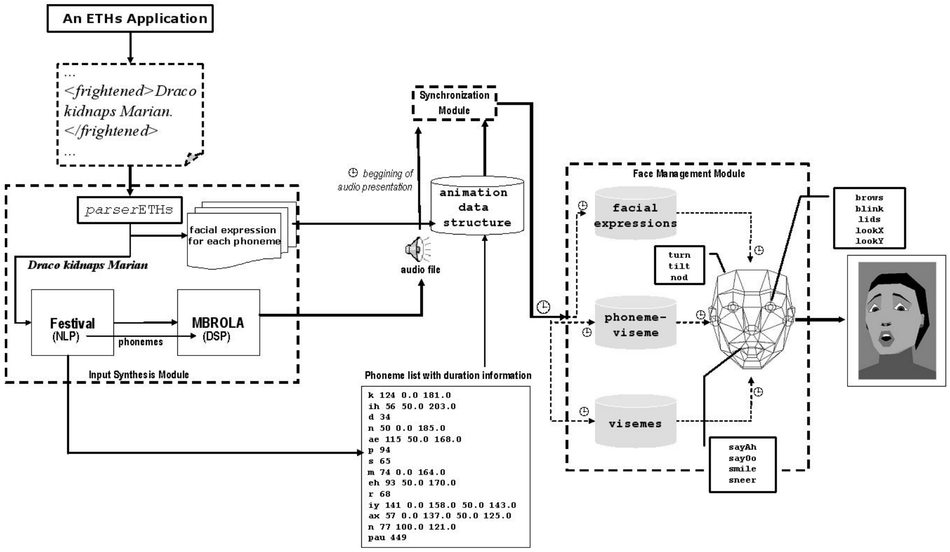

Figure 3: An overview of the ETH architecture and its submodules.

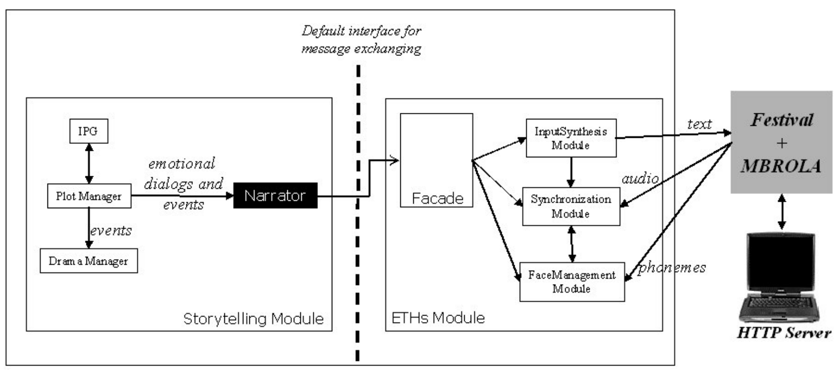

Figure 4: The virtual narrator system.

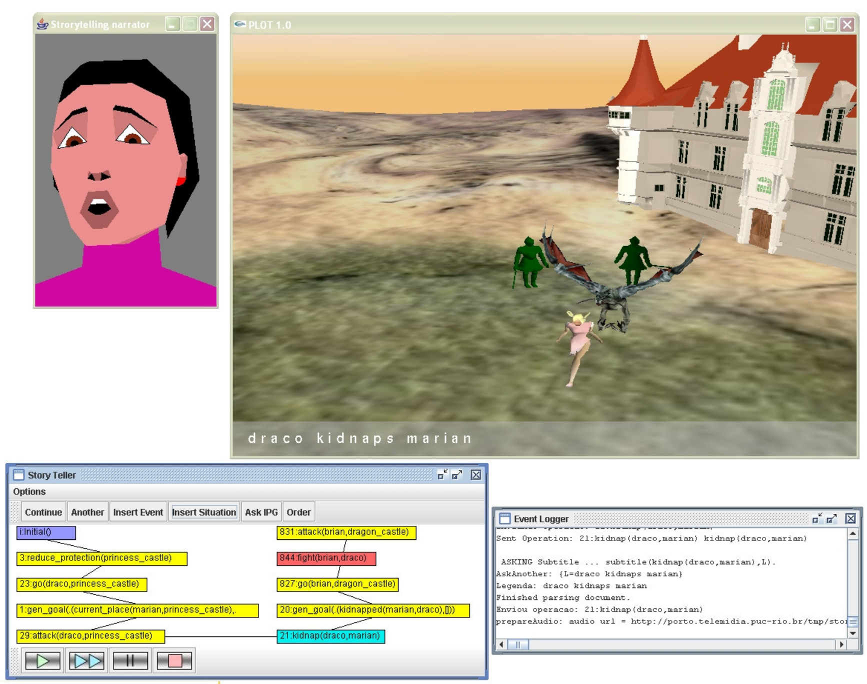

Figure 5: A snapshot of the virtual narrator environment.