
This full text version, available on TeesRep, is the post-print (final version prior to publication) of:

Cavazza, M. O. et al. (2007) 'Madame Bovary on the holodeck: immersive interactive storytelling', Proceedings of the 15th international conference on Multimedia, 2007, Augsburg, Germany, September 24 - 29. New York: ACM, pp.651-660.

"© ACM, 2007. This is the author's version of the work. It is [posted] included here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in Proceedings of the 15th international conference on Multimedia, 2007

<http://dx.doi.org/10.1145/1291233.1291387>

When citing this source, please use the final published version as above.

This document was downloaded from <http://tees.openrepository.com/tees/handle/10149/58296>
Please do not use this version for citation purposes.

All items in TeesRep are protected by copyright, with all rights reserved, unless otherwise indicated.

Madame Bovary on the Holodeck: Immersive Interactive Storytelling

Marc Cavazza

School of Computing
University of Teesside
Middlesbrough,
TS1 3BA, UK

m.o.cavazza@tees.ac.uk

Jean-Luc Lugrin

School of Computing
University of Teesside
Middlesbrough,
TS1 3BA, UK

j-l.lugrin@tees.ac.uk

David Pizzi,

Fred Charles
School of Computing
University of Teesside
Middlesbrough, TS1 3BA, UK

d.pizzi@tees.ac.uk

ABSTRACT

In this paper, we describe a small-scale, yet complete, integration of a real-time immersive interactive storytelling system. While significant progress has been achieved in recent years on the individual component technologies of interactive storytelling, the main objective of this work is to investigate the concept of interactive storytelling in a fully immersive context. We describe each individual component of immersive interactive storytelling from a technical perspective. We have used a commercial game engine as a development environment, supporting real-time visualisation as well as the inclusion of Artificial Intelligence components controlling virtual actors. This visualisation engine has been ported to an immersive setting using dedicated software and hardware supporting real-time stereoscopic visualisation. The hardware platform is built around a 4-sided CAVE-like immersive display operated by a PC-cluster. The interactive storytelling engine is constituted by a planning system based on characters motivations and emotional states. The user can interact with the virtual world using multimodal interaction. We illustrate the system's behaviour on the implementation of excerpts from *Madame Bovary*, a classic XIXth century novel, and demonstrate the ability for the user to play the role of one of the characters and influence the unfolding of the story by his actions.

Categories and Subject Descriptors

H.5.1 [Multimedia Information Systems] Artificial, Augmented and Virtual Reality - *Virtual Reality for Art and Entertainment*

General Terms

Theory, Algorithms, Design and Experimentation

Keywords

Interactive Storytelling, Virtual Reality, Virtual Actors, Multimodal Interfaces.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MM'07, September 23–28, 2007, Augsburg, Bavaria, Germany.

Copyright 2007 ACM 978-1-59593-701-8/07/0009...\$5.00.

1. INTRODUCTION AND OBJECTIVES

The emergence of interactive storytelling technologies has been one of the most exciting developments in the field of new media. Interactive storytelling is a long-term endeavour, which involves not only advances in interactive technologies, but also new modes of media content production. A frequent metaphor for interactive storytelling is that of the Holodeck™ [13] [29] [34], the science-fiction ultimate entertainment system, where narratives take the form of virtual reality world in which the user is totally immersed, interacting with other characters and the environment in a way which drives the evolution of the narrative. In recent years, over 30 research papers in the field of virtual reality and new media have made reference to this popular metaphor. From a technological perspective, it is now 10 years since Altman and Nakatsu [1] hypothesised the convergence of virtual reality and traditional media such as cinema, to create new immersive interactive experiences of that kind.

It is thus a worthy enterprise to attempt to create a first prototype in pursuit of the Holodeck™, which would implement an immersive interactive storytelling system synthesising the current state of the art in interactive storytelling technologies, within a fully immersive virtual reality setting.

In this paper, we report the integration of such a prototype of immersive interactive storytelling system, which capitalises on previous research in interactive storytelling.

There are several objectives to this research. In the first instance, this technology demonstrator aims at measuring the actual complexity of prototype development. In addition, we need to gain understanding of some more systemic aspects, such as the constraints on user interaction which arise in an immersive setting and how they impact on the type of interactive narrative in terms of user involvement. This is best discovered through experimenting with an early prototype. Still, as an output of these first experiments, we aim at demonstrating at least the feasibility of immersive interactive storytelling and its potential in terms of user experience. Due to the scale and complexity of the overall implementation, we are essentially aiming at a proof-of-concept demonstrator, rather than a prototype that could undergo user evaluation. Our aim is thus to demonstrate the feasibility of immersive interactive storytelling by assembling a working prototype which would satisfy the combined requirements of immersion and interactive storytelling, i.e. real-time display and user real-time response.

The choice of the background narrative is poised to have an impact on user experience. Different narrative genres bring different constraints to their interactive counterparts. To emphasise that this research is focused on developing new media rather than simulation, we have decided to use a classic XIXth century French novel, *Madame Bovary* [12] by Gustave Flaubert, as the background for our immersive interactive storytelling experiments. To some extent this experiment will be a *remediation* [2] of the novel into an interactive format, which is why we need to represent baseline narrative actions from the original novel to implement the interactive version. This is rather ambitious and, in view of the complexity of the narrative, our prototype focuses on one single episode of the novel, which consists of the love affair between Emma and Rodolphe (chapters 9-12 of Part II of the original novel). The scene which will illustrate the prototype (in the forthcoming sections) is an encounter between Emma Bovary and Rodolphe (whose role is played by the user), taking place after they have already started their affair. This scene offers good opportunities for interaction, as Rodolphe's response will determine the outcome of their affair and influence the subsequent unfolding of the narrative.

2. IMMERSIVE STORYTELLING: ARCHITECTURE AND OVERVIEW

The system consists of a CAVE-like system [9] allowing the user to be immersed into a virtual world. The narrative unfolds as a real-time stereoscopic 3D animation, featuring virtual actors, and controlled by the interactive storytelling engine. Characters express themselves using speech synthesis as well as body animations (and elementary facial animations with lip synchronisation) and the user can interact with them naturally, using speech and attitudes, as if acting on stage (Figure 1).

Our baseline interactive storytelling technology has been developed as a desktop application using game engine technology (Unreal Tournament™, henceforth UT) to support visualisation

and interaction [17], while we have developed a family of interactive storytelling engines based on AI Planning technology (see below). To extend this approach to the context of immersive interactive storytelling, we have resorted to a specific software package, CaveUT 2.0.

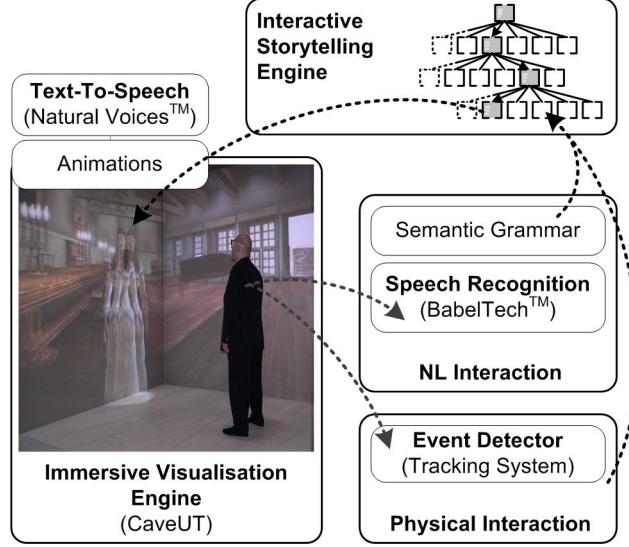


Figure 1: Software Architecture for Immersive Interactive Storytelling.

CaveUT 2.0 [16] is an open source package, developed largely using UT's scripting language UnrealScript, which supports the use of any interactive application developed around the UT game engine within a CAVE-like virtual reality system, including real-time head and hand tracking, as well as stereoscopic display at up to 60 frames per second (Figure 2). The four-screen display based on CaveUT relies on a server computer connected by a standard LAN to 8 client computers (two for each screen, to support stereoscopic display). On each client, the CaveUT code rotates

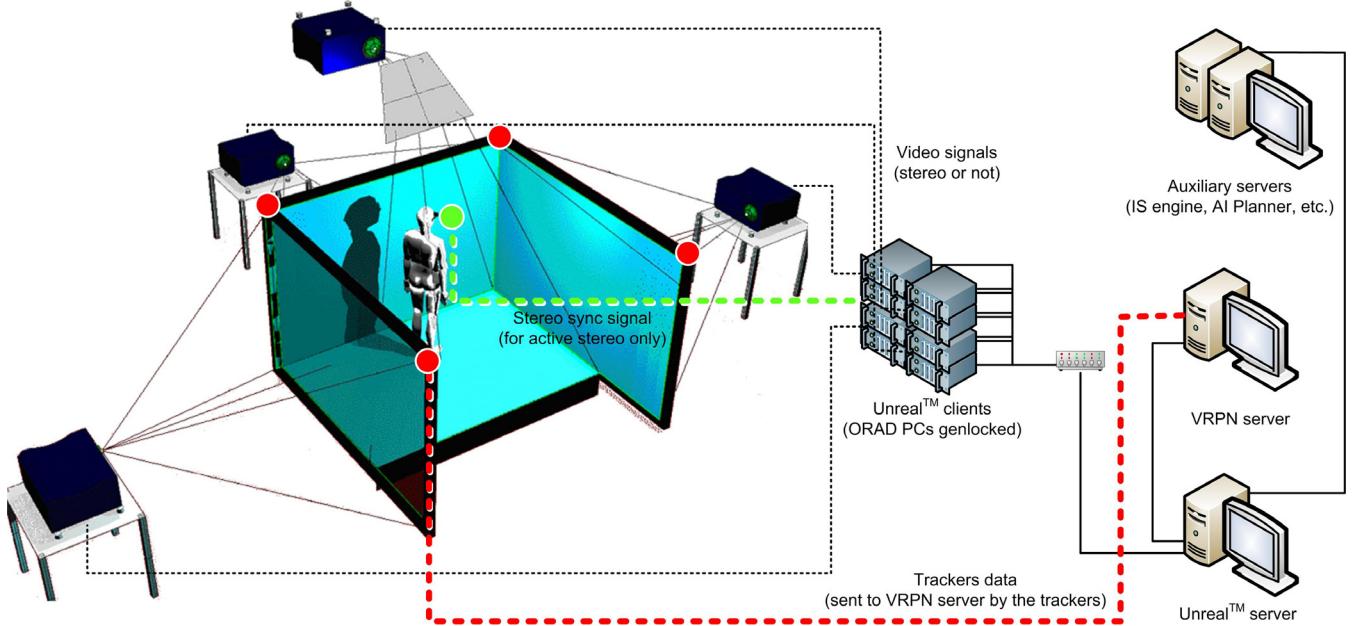


Figure 2: Hardware Architecture for the PC-based CAVE-like Immersive Display.

the view according to parameters defined in a configuration file. In addition, the perspective correction on each screen must be adjusted so that the ideal viewing point for each screen is located at the same point in physical space. This is achieved using a modified OpenGL wrapper library, called VRGL™¹, which operates between OpenGL and the UT engine.



Figure 3: The CaveUT Software supports immersive visualisation using the UT game engine.

The CaveUT 2.0 installation supports stereographic display by using two ORAD computers for each projection screen, one for each eye view, and uses specific video hardware (DVG cards) in their ORAD cluster to mix the two and send it to a single stereographic projector which will alternate between the left and right eye views at 120 frames per second. The user wears shutter glasses to view the stereoscopic image, as is traditional in CAVEs (Figure 1). All of the screens in the composite display must also switch view in a synchronised fashion, which is ensured via a process known as “genlocking”, also handled by the DVG hardware (Figure 3).

CaveUT makes it possible to port almost seamlessly an interactive storytelling system developed for a desktop configuration to an immersive setting. However, there exist specific constraints on visualisation, navigation and “physical” interaction in an immersive setting. Firstly, the user visualises the scene in a first person mode and all physical interactions are mediated by tracking his position (he is equipped with a head and hand tracking system). The immersive context, where all action revolves around the user can thus dispense with the use of sophisticated (narrative-based) camera systems used with desktop interactive storytelling systems, which track events of importance and choose the appropriate viewpoints and camera shots (as the viewpoint is almost always the user’s).

Secondly, the immersive context requires an adaptation of collision detection so as to preserve a minimal level of physical realism, e.g. avoiding that the user traverses characters, doors or furniture. This is obtained by centring a bounding box on the user position in the virtual world, obtained in real-time from head tracking data. This will grant him physical existence including in terms of collisions with virtual objects or virtual actors. CaveUT also manages tracker data, which in this experiment correspond to head and hand tracking: head tracking is an essential part of the user-centred visualisation component, as well as interaction. As physical motion is limited by the floor surface (as in most

CAVEs), the user is equipped with a handheld controller to navigate into the virtual environment, the same device enabling hand tracking.

Finally, the production of synchronised stereoscopic animations running on the eight clients of the PC cluster required specific extensions to some native mechanisms from the UT game engine to be developed. The native UT replication mechanisms for animations in a networked environment use traditional, physics-based approximations. Consequently, the native replication system cannot deal with character animations relying on procedural animation (i.e. skeletal and vertex, which are used for character expressions). Our extension uses replication statements and facilities present within the UT scripting language (UnrealScript) to allow any client to constantly listen which animation is being played on the server, and at which rate. As a result, our prototype easily supports up to three simultaneous animated virtual characters in the CAVE, with no perceptible gap between client animations.

3. THE INTERACTIVE STORYTELLING ENGINE

This section describes the basic mechanisms supporting interactive storytelling, i.e. the creation of an immersive narrative featuring artificial actors, which can also respond consistently to user interaction. The processing of multimodal user interaction in an immersive context, as well as examples illustrating their impact on story evolution, will be described in the next section.

Interactive storytelling systems rely on an “interactive storytelling engine”, which is in charge of generating a sequence of narrative actions as well as maintaining narrative consistency despite variability in user interventions. There is now a consensus among researchers in interactive storytelling on the use of planning technologies to support narrative actions generation [6] [7] [15] [24] [32]. Planning is naturally geared towards action generation, which now makes it the AI technique of choice. In addition, it can propagate the consequences of user interaction on the unfolding of the narrative, either through real-time planning or through the use of re-planning.

Madame Bovary, is quintessentially a psychological novel, in which relationships between characters, and characters’ feelings play a much greater role than in other genres (such as epics or tales) that may instead emphasise the outcome of actions. It is thus necessary to reflect the central role of characters’ feelings in defining the various stages of the narrative and in determining their actions.

We consider a narrative state to be described by the feelings of the main characters relative to their situation, with a specific emphasis on those of the feature character Emma Bovary. From the inventory of feelings described by Flaubert himself in his preparatory works to the novel [20] we have described a list of 21 feelings each taking a value in the {low, medium, high} interval. Examples of such feelings include anger-towards (?x, ?y); accepts-adultery-risk (?x, ?y), which represents Emma’s awareness of social constraints and willing to transgress them; power-over (?x, ?y), which represents the influence of one character over another.

¹ by Willem de Jonge and PublicVR.

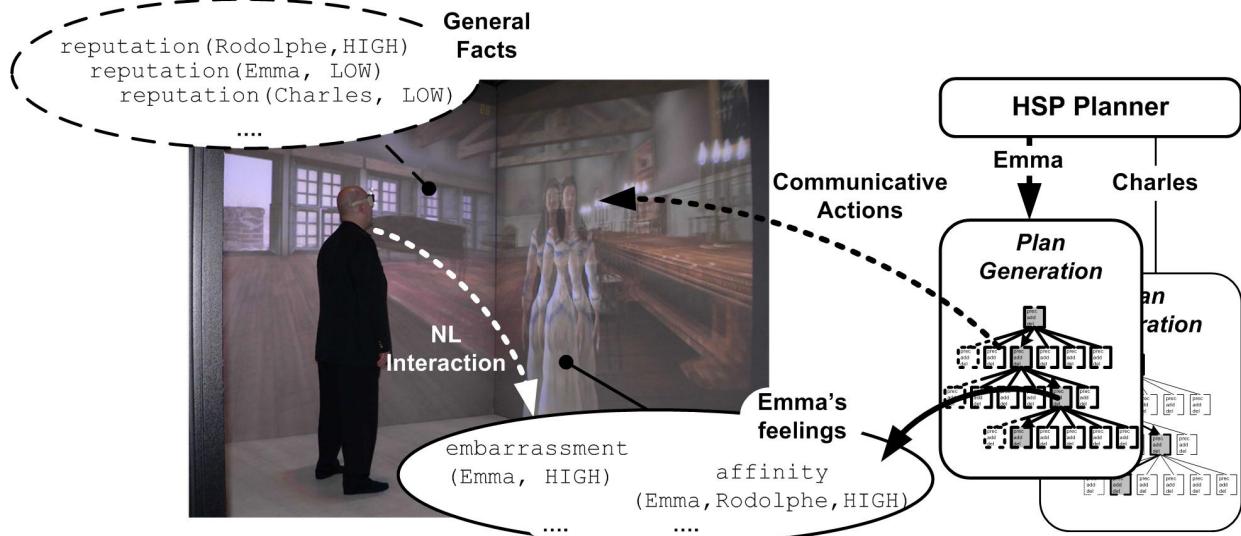


Figure 4: Emotional Planning for Interactive Storytelling: narrative states are based on characters' feelings.

To reflect the fact that characters' behaviours are driven mostly by their psychological states, we use conjunctions of elementary feelings as goals for the planner (Figure 4). As new situations will emerge (as a result of other characters' actions, including the user) that may alter character's desires, our system supports dynamic goals, i.e. allows for the addition or deletion of certain feelings in the current goal.

We use an implementation of Heuristic Search Planning (HSP) [3] [4] for the planner which is at the heart of our interactive storytelling engine, in which each character is operated by his own HSP planner. We have used a simple STRIPS-like [11] description for operators. One such operator is depicted on Figure 5: it illustrates how a given action (Emma asking Rodolphe to take her away from her boring life) can be determined by a characters' emotional state, defined as a conjunction of elementary feelings of the type described above. HSP can be used in a real-time context simply by incorporating a real-time search algorithm: to that effect, we use RTA* [19] to search the space of applicable operators so as to be able to plan in the dynamic environment modified by user interaction. In our HSP implementation, the heuristic value, which is based on the contribution of that operator to reaching the goal state, is computed using the Value Iteration method (VI) of Liu et al [23]. The resulting planner performance is fully compatible with an interactive system as it determines a character's next action or updates characters' feelings in less than 200 ms on average.

We have chosen a character-based approach rather than a plot-based approach for the interactive storytelling engine. This means that each character will be driven independently by its own feelings. At the same time, because certain feelings will drive them to interact with other characters, we incorporated mechanisms synchronising characters interaction (whether this interaction is between two virtual actors or between the user and a virtual actor). One such mechanism is the definition of specific operators managing an interaction between two characters. Requests or invitations, which require a response from the other character, naturally fall under that category. For instance, Emma may request Rodolphe to take her away from Yonville (the small

town where she lives her boring life) to the capital. After producing such a request, she will enter into a waiting state where planning new actions is suspended until she receives a response from the other character (or the user) or the waiting duration exceeds a maximum time limit. Receiving an answer would cause her to resume her planning activity, which will naturally incorporate any new emotional states brought up by the user's reply. This is achieved using two types of operators: actions and interpretations operators (the former communicate with other actors while the latter update emotional states following interaction - see also Figures 9 and 10).

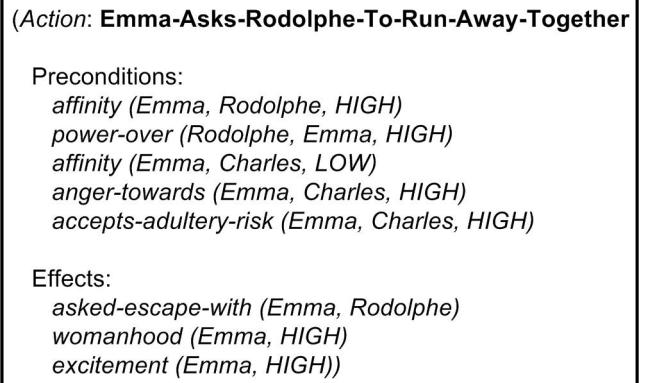


Figure 5: An example Operator used for by the Interactive Storytelling engine.

The immersive narrative, as perceived by the user, is composed of a succession of real-time animations showing the characters moving around on stage, performing actions and expressing themselves through utterances, body attitudes and gestures. All these animations are generated by elementary actions associated to planning operators which stage narrative actions (e.g., Emma walking towards the user if wanting to address him, Emma making a love declaration, Emma complaining about her boring life, etc.) and dramatise them using expressive animations.

The user takes part in the interactive narrative by playing a role as one of the novel's characters (only Rodolphe or Charles in our current implementation), in a first person mode, being embodied by an avatar. This means that he needs to act and perform according to the role, although he is free to make decisions within that role. In the first instance, a significant part of the action is driven by Emma's desires who will take the initiative of interacting with Rodolphe (the user character). This is why the user's role will be facilitated in that it will consist in responding to the main character's action. By doing so, he will influence the emotional states of other actors, which in turn will affect future events in the narrative.

The majority of narrative actions played by virtual characters consist of communicative actions, which to be properly staged require an appropriate soundtrack to be generated. Communicative actions are thus associated dialogues that are transmitted to a Text-To-Speech (TTS) engine. The limited quality of off-the-shelf TTS engines constrains the length of utterances, whose length has been restricted to less than 15 words in most cases. In this *Bovary* scenario, we used as utterances actual excerpts from the novel's dialogues attached to planning operators such as those of Figure 8 (using the English translation of the novel). A given operator can be associated similar utterances in context (several invitations from Emma to Rodolphe to take her away from Yonville, for instance) from which one is selected randomly to ensure additional variability.

4. MULTIMODAL INTERACTION AND USER INVOLVEMENT

The only input modality compatible with an immersive interactive storytelling experience is multimodal interaction, based on natural modalities such as speech and various forms of physical interaction (such as user position, body attitude and gestures). It is however clearly beyond the state-of-the-art of speech recognition to support free dialogue with artificial actors, even more so if these dialogues should achieve the level of sophistication typical of narrative dialogues. It is thus necessary to compensate for the limited performance of speech recognition systems, trying to make use of the strong context provided by the underlying narrative. One approach, known as habitable language [35] consists in defining a recognition syntax using a speech recognition grammar (such as the one provided by the off-the-shelf speech recognition system we are using, from BabelTech™) which will take into account a sufficient number of syntactic variants not to require the user to learn a command syntax (Figures 6 and 7). Further, in the context of an interactive narrative, users may be replying to dialogues initiated by the virtual actors (in our case, Emma), which potentially restricts user utterances in terms of topics and complexity [18]. We still expect users to require a short briefing session before being able to use the prototype, being shown examples of spoken interaction before using the system, which in particular will insist on the limited length of utterances. In other words, in order to be able to take part in the interactive narrative, the user must gain some familiarity with the baseline plot and with typical dialogues. This is somehow consistent with the original concept of the Holodeck™, where the user was aware of the baseline plot and his expected role. Another aspect potentially facilitating speech recognition is that the user is expected to act, which should

(hopefully) limit disfluencies, but we were unable to test that hypothesis.

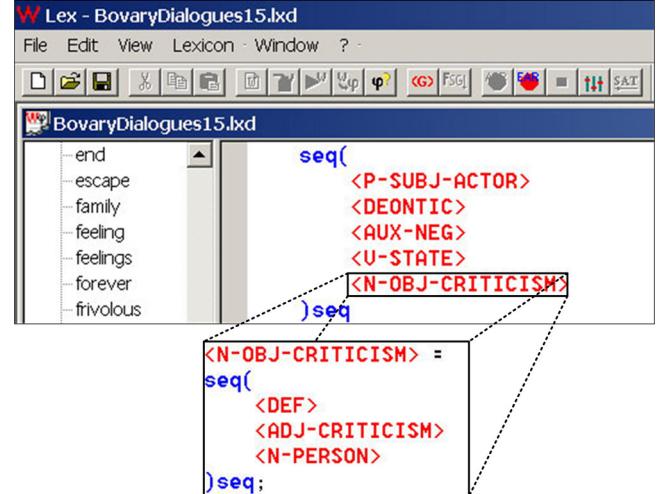


Figure 6: Semantic grammar definition for Speech Recognition (here, a criticism speech act).

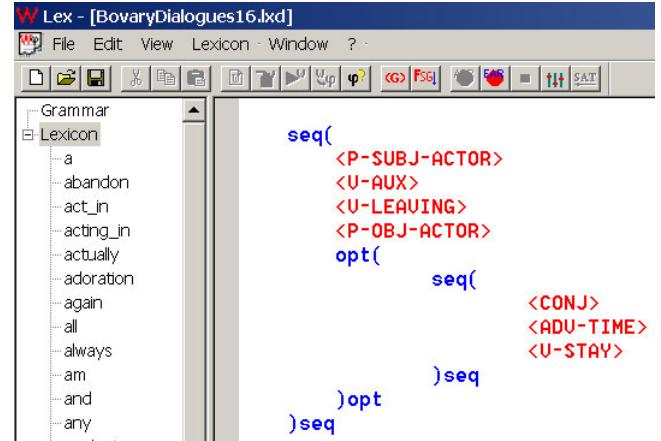


Figure 7: Developing the grammar for the Speech recognition system. The instantiation of a template can also be used for semantic interpretation.

In the prototype we describe here, the speech recognition grammar is defined as a semantic grammar to integrate recognition and a first step of interpretation. Habitability is supported by i) the multiple syntactic constructs associated to a semantic category and ii) the lexical variants for terminal vocabulary. Figure 6 illustrates the template for a set of speech acts advising Emma of her improper behaviour. The template makes explicit reference to deontic categories (should, can) and supports variants for the types of criticism involved (selfishness, frivolity, etc.).

Essentially, patterns in the semantic grammar have been defined for a set of communicative actions most likely to take place in a given state of the narrative. Recognition is thus expectation-driven: in order to have an influence on the story, the user's interactions take the form of a limited number of speech acts responding to the narrative situations defined by the baseline plot. For instance, there are narrative constraints on the types of responses Rodolphe can produce to Emma's love declarations,

such as reciprocal feelings, disdain or indifference, concern, etc. Each of these will correspond to a small number of speech acts, such as Rodolphe announcing his intention to leave, reminding Emma of social constraints, reminding Emma of her family, or criticising her conduct.

The recognition of a spoken utterance instantiates at the same time a semantic template, whose contents can then be matched to elements of the planning domain using a set of dedicated rules.

One typical example is constituted by narrative actions such as abandoning Emma, for which there is a direct match to a recognition template covering a set of alternative utterances. Figure 8 illustrates such an interaction by also showing how the planning system modifies Emma's emotional state and influences her subsequent behaviour (only the steps immediately following the intervention are represented).

This exchange takes place at a time where their affair has been going on for some time and Emma is comfortable with its implications (`accepts-adultery-risk(E, C, HIGH)`). At the same time, Rodolphe is exerting a real influence on her (`power-over(R, E, HIGH)`). When Emma reminds Rodolphe of her feelings (`Says-something-in-confidence`) for him, Rodolphe (the user) replies by announcing his intention to leave her ("I will leave you and never see you again"). This action from Rodolphe has a direct impact on Emma's feelings by generating anger towards him (`anger-on`

towards(E, R, High)). This emotional state will actually trigger a chain of events, which is actually the key mechanism for story generation. In particular, Emma will re-assess her situation, which will lead her to gradually recover from the influence Rodolphe had on her (`Regrets-falling-for-Rodolphe`). This is where her behaviour becomes driven by her longer-term goal again, which is (to put it simply) to pursue happiness. In the absence of any romantic solicitations, she will find happiness in her family life (`Emphasises-Motherhood`).

Other types of speech acts have a more implicit content. For instance (Figure 9), Rodolphe responds to Emma's begging him to take her away from her boring life by reminding her of social constraints and/or criticising her attitude ("you should not be one of these frivolous women"). Implicit speech acts don't have a direct mapping to narrative actions; rather, they can be interpreted in terms of emotional states in the planning domain. This is achieved by mapping certain semantic categories instantiated in the speech recognition template to emotional states in the planning domain using ad hoc rules. In the example of Figure 9, the semantic interpretation of this utterance consists of generating a feeling of embarrassment (`embarrassment(E, HIGH)`). Speech recognition templates corresponding to those interventions are of a more generic nature: for instance all those templates containing deontic categories (expressing permission or obligation) will be the starting point for that type of interpretation.

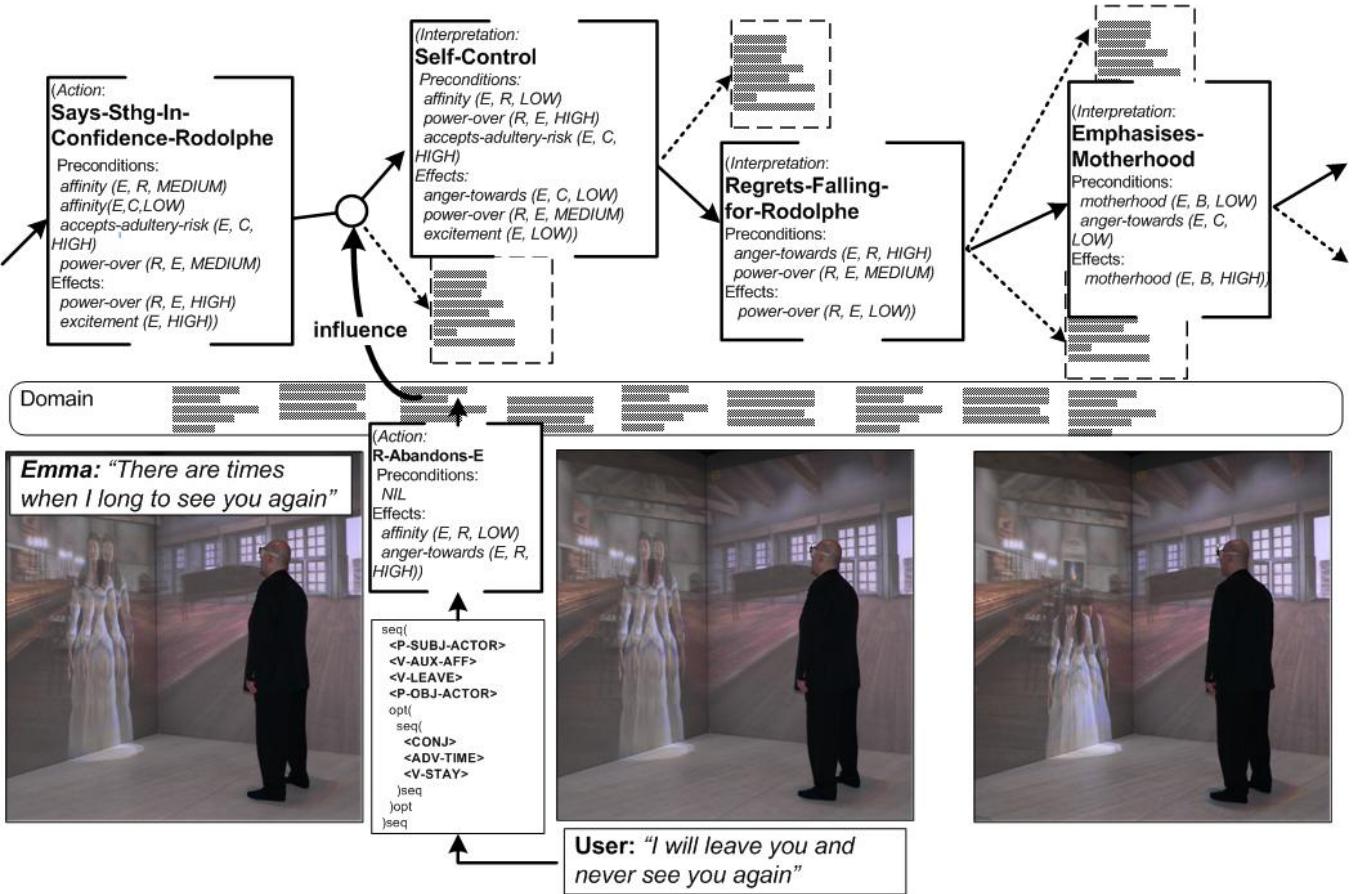


Figure 8: An example interaction during the immersive narrative: Rodolphe (the user) abandoning Emma.

In this second example (Figure 9), the relation between E and R is even more advanced ($\text{affinity}(E, R, \text{High})$), Emma is ready to take any risk ($\text{accepts-adultery-risk}(E, C, \text{High})$), as she has estranged her husband Charles ($\text{angers}(E, C, \text{High})$). She is willing to run away with Rodolphe and expresses this to him (the user). This stage corresponds to an actual situation from the novel, re-generated from baseline data. Here the user criticises her attitude (“*you should not be one of these frivolous women*”), and this criticism is recognised by the speech interpretation system, which updates Emma’s feeling to one of embarrassment ($\text{embarrassment}(E, \text{High})$). The subsequent behaviour of Emma, influenced by that feeling will result in her being more aware of the risks ($\text{accepts-adultery-risk}(E, C, \text{Low})$) and eventually feeling closer to her husband.

Speech acts templates differ from those used in previous interactive storytelling systems [8] [27] in several ways. We have imposed stylistic constraints on the user’s expressions, which should be consistent with the language used by *Madame Bovary*’s characters. Secondly, the mapping between speech acts and narrative actions is less straightforward than in other systems. While some speech acts, such as *criticism*, are similar to those described in [27], in other cases it is the semantic content of the utterance which will be matched to emotional representations [31]. As a result, while the speech interface is based on well-defined principles, it should be portable to other domains and

genres (by redefining style and vocabulary) rather than directly re-usable.

Finally, the immersive interface naturally entails some kind of “physical” interaction through body attitude, proxemics, as well as elementary gestures. These may take place in conjunction with a user utterance, or can be interpreted in isolation as non-verbal communicative actions. In our first prototype, only the detection of the user’s body orientation and distance relative to another character has been implemented. For instance, in response to an invitation by Emma (the same discussed in the example above), the user (again playing the part of Rodolphe) can turn his back on Emma and that in itself is recognised as a negative answer to the specific interaction initiated by Emma (Figure 10). More specifically, Emma’s request creates a conversation state, which upon receiving a response from the user will trigger a new interpretation of her emotional state based on the contents of that response. The user’s attitude (non-verbal behaviour, turning his back on Emma) is actually interpreted as a response, namely as refusing the discussion (creating the state *interrupted-conversation*(E, R)). The interpretation mechanism is based on a contextual rule which can be triggered when the current operator is a request from Emma. This rule tracks the position of the user and detects specific attitudes which form part of a pre-defined list of non-verbal “expressions”. Communicative acts can thus be associated to non-verbal behaviour provided the context in which they occur can be specified.

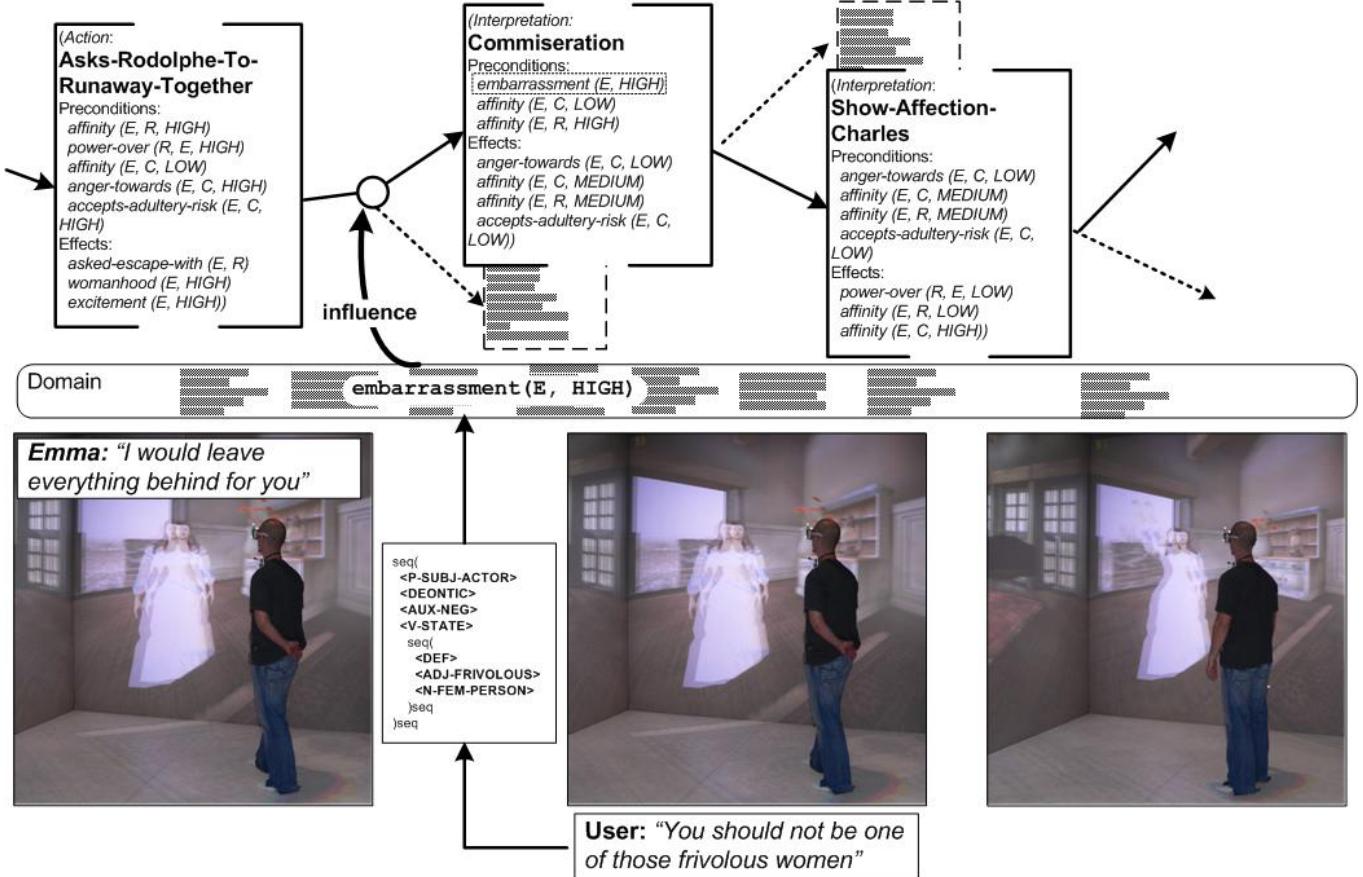


Figure 9: An Example interaction in the immersive narrative: Rodolphe (the user) criticising Emma.

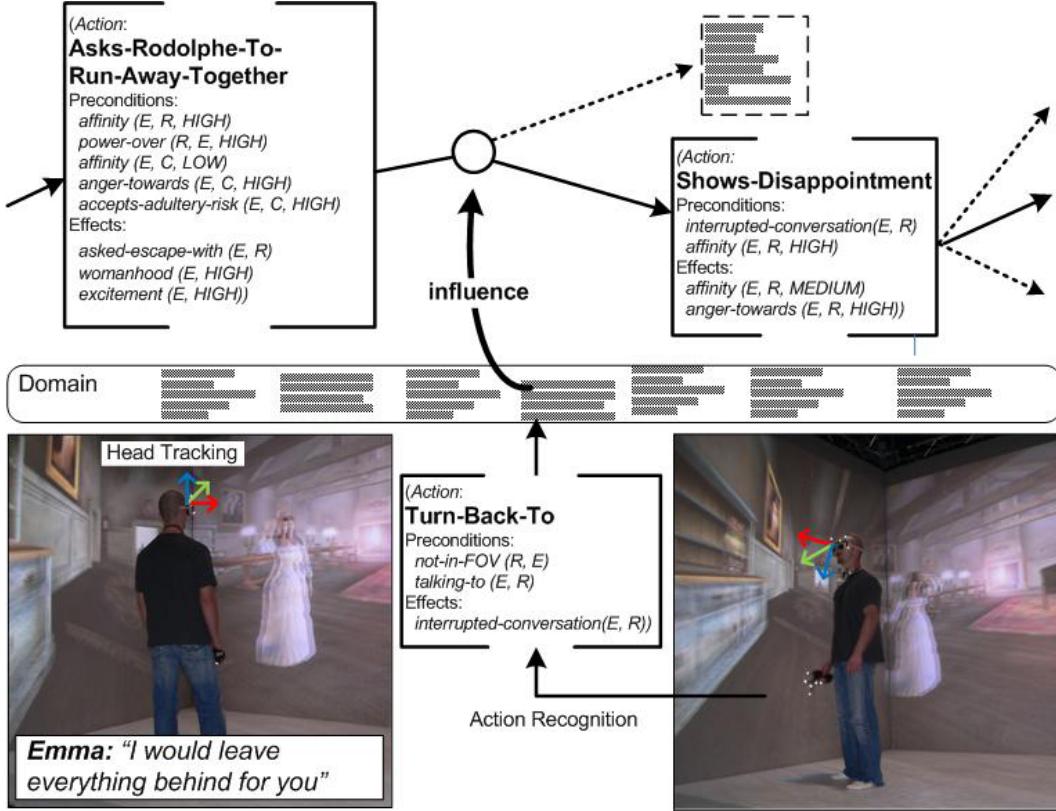


Figure 10: Physical interaction with the immersive narrative: the user turns his back on Emma, which can be recognised in terms of non-verbal behaviour.

5. RELATION TO PREVIOUS WORK

While various types of interactive storytelling systems have been demonstrated over the past 15 years, the field has only developed significantly in recent years, with the emergence of AI-based storytelling systems, which rely on explicit plot representations. This has made the development of IS system a highly multidisciplinary enterprise involving AI, HCI, Computer Graphics, and narrative theory. In that sense, there have been very few large-scale attempts at integrating every single component of interactive storytelling (including multimodal interaction). One well known example has been the Mission Rehearsal Exercise at ICT, which also made explicit reference to the Holodeck™ concept [34]. However, probably because of multi-user/spectator constraints it did not make use of stereoscopic fully immersive environments. Another major difference with the work described here is the purpose and genre of the narrative: the MRE mostly enriches simulation with narrative and is not preoccupied with literary or aesthetic aspects of the narrative itself, only its believability and the relevance of the emotions manifested [26] which are mostly Ekmanian emotions rather than literary feelings.

Façade [27], on the other hand, was the first large-scale integration of natural language input in an interactive narrative. Its particularity is that the narrative is almost entirely dialogue based and the user plays a predefined role as the “third character” in the narrative. *Façade*’s high quality comes at the cost of significant authoring work [25], dedicated to narrative actions and how natural language input should match them. The augmented

reality version, *AR-Façade* [10], is getting more similar to immersive storytelling, including in its use of speech recognition, but is not addressing the creation of an entirely artificial environment (for the narrative) in the sense of the Holodeck™. Another difference is that within our baseline plot there is more freedom to play different characters as the user is not assigned a fixed role.

The main difference is of course that our system is totally immersive, which has an impact on various aspects of visualisation and user interaction, as described in section 4.

6. CONCLUSIONS

We have presented a first implementation of an immersive interactive storytelling application, which integrates state-of-the-art interactive storytelling technologies into an immersive installation, inspired from the long-term endeavour of the Holodeck™. We shall not discuss here the relevance or value of immersive storytelling in itself, which has already been debated by various authors [1] [5] [8] [29] [33]. Instead, we shall discuss the overall feasibility of the whole enterprise, also addressing development effort and authoring issues.

Although evaluation is an important concept in most interactive systems, one cannot apply traditional user evaluation methods to proof-of-concept demonstrators, where the challenge rests with actually implementing concepts to a level of technical performance (real-time response, accuracy), which will make possible evaluation at a later stage. This is why in the following

sections we only include some qualitative findings: we can nevertheless conclude to the stability of the system, which in multiple trials over a period of one week has demonstrated regular and reproducible performance in terms of stereoscopic visualisation, frame rate, speech recognition performance (which on average was 90% accurate, with a response time always < 2 sec.), and real-time behaviour of the interactive storytelling engine. However, the next step in system evaluation should consist in measuring the impact of user interventions using story progression diagrams as described in [31]. These diagrams plot the evolution of the narrative towards certain types of ending states, thereby visualising its overall evolution. They constitute a good analytical tool to evaluate the actual impact of user interventions, including when these have mostly long-term effects on the characters.

From a visual content development perspective, most of the environment and character animations developed for a desktop interactive system are readily transposable to an immersive context. We have illustrated this point by first developing the system described here in a desktop configuration.

When implementing an immersive approach, one should take into consideration how it affects interactive storytelling paradigms, defined in terms of user involvement and participation. The traditional Holodeck™ concept implies constant user involvement (as one of the story characters), which is probably beyond the capabilities of our current interactive storytelling systems. Constant involvement also puts additional strain in terms of narrative consistency: traditional interactive storytelling systems can accommodate user intervention and recalculate the baseline plot according to its consequences but are faced with complexity limitations when trying to maintain narrative consistency in the case of constant involvement.

There is agreement in the interactive storytelling community to measure the scale achieved by an interactive narrative using story duration divided by the number of beats, as introduced by Mateas [28], who proposed as a first milestone for interactive storytelling to achieve a 10-minute duration with one beat per minute. While we have been able to achieve story lengths of up to 6 minutes with the same baseline plot and the same engine on a desktop version [30], porting the system to an immersive context with permanent user involvement significantly affected scale, as additional developments were required to incorporate immersive multimodal interaction (in addition to the limitations in story complexity outlined above). As a consequence, the longest interactive stories we were able to obtain were only 2 minutes in duration.

This tends to measure how far we still are from immersive interactive storytelling but is also encouraging considering the unparalleled level of immersion, both in the narrative and in the virtual world potentially achieved by such systems.

7. ACKNOWLEDGEMENTS

Marc Le Renard is thanked for his assistance in operating the SAS-Cube™. Sean Crooks managed the visual design of the virtual environment and the virtual actors. This work has been funded in part by the DTI, under the Technology Programme, through the BARDS Project. CAVE™, Holodeck™ and Unreal Tournament™ are registered trademarks.

8. REFERENCES

- [1] Altman E.J., and R. Nakatsu. Interactive Movies: Techniques, Technology and Content. Course Notes n.16, ACM SIGGRAPH'97, 1997.
- [2] Bolter, J.D. and Grusin, X. Remediation: Understanding New Media. Cambridge: MIT Press, 1999.
- [3] Bonet, B. and Geffner. H. Planning as Heuristic Search. Artificial Intelligence Special Issue on Heuristic Search, 129, n.1, 2001, pp. 5-33.
- [4] Bonet, B and Geffner, H., Planning as Heuristic Search: New Results, Proceedings of the European Conference on Planning (ECP'99), 1999, pp. 360-372.
- [5] Brooks, K., There is Nothing Virtual About Immersion: Narrative Immersion for VR and Other Interfaces. 2003 (<http://alumni.media.mit.edu/~brooks/storybiz/immersiveNotVirtual.pdf>).
- [6] Cavazza, M., Charles, F. and Mead, S.J. Interacting with Virtual Characters in Interactive Storytelling. First ACM Joint Conference on Autonomous Agents and Multi-Agent Systems, Bologna, Italy, pp. 318-325.
- [7] Cavazza, M., Charles, F. and Mead, S.J. Character-based Interactive Storytelling. IEEE Intelligent Systems, 17, 4, 2002, pp. 17-24.
- [8] Charles F., Martin O., Cavazza M., Mead S.J., Nandi A. and Marichal X. Compelling Experiences in Mixed Reality Interactive Storytelling. International Conference on Advances in Computer Entertainment Technology, ACE 2004, ACM Press, Singapore, 2004, pp. 32-41.
- [9] Cruz-Neira, C., Sandin, D.J. and DeFanti, T.A. Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE. Proceedings of the ACM SIGGRAPH 1993 Conference, 1993, pp. 135-142.
- [10] Dow, S., Mehta, M., Lausier, A., MacIntyre, B., and Mateas, M. Initial Lessons from AR-Façade, An Interactive Augmented Reality Drama. In ACM SIGCHI Conference on Advances in Computer Entertainment (ACE'06), Los Angeles, CA, 2006.
- [11] Filkes, R. and Nilsson, N. STRIPS: A New Approach to the Application of Theorem Proving to Problem Solving. Proceedings of the Second International Joint Conference on Artificial Intelligence, 1971, pp. 608-620.
- [12] Flaubert, G. Madame Bovary. La revue de Paris (Publisher), France, 1856 (in French).
- [13] Foote, J. Where's My Holodeck? The New Frontier of Media Display. IEEE MultiMedia 13, n. 3 (Jul. 2006), 104.
- [14] Gratch, J. Emile: Marshalling Passions in Training and Education, in Proceedings of the 4th International Conference on Autonomous Agents, Barcelona, Spain, June 2000.
- [15] Gratch, J. Why you should buy an emotional planner, Proceedings of the Autonomous Agents 1999 Workshop on Emotion-based Agent Architectures (EBAA'99).

- [16] Jacobson, J. Using CaveUT to Build Immersive Displays With the Unreal Tournament Engine and a PC Cluster, ACM SIGGRAPH 2003 Symposium on Interactive 3D Graphics, Monterrey, California, 2003.
- [17] Jacobson, J., and Hwang, Z. Unreal Tournament for Immersive Interactive Theater, Communications of the ACM, 45, 2002, pp. 39-42.
- [18] Julia, L and Cheyer, A., Is Talking to Virtual More Realistic? Sixth European Conference on Speech Communication and Technology, EUROSPEECH'99, Budapest, Hungary, September 1999.
- [19] Korf, R.E. Real-time heuristic search. Artificial Intelligence, 42:2-3, 1990, pp. 189-211.
- [20] Leclerc, Y. Plans et Scenarios de Madame Bovary, CNRS Editions, France, 1995 (in French).
- [21] Lewis, M., and Jacobson, J. Game Engines In Scientific Research. Communications of the ACM, 45, 2002, pp. 27-31.
- [22] Liu, H. and Mihalcea, R., AAAI Workshop on "Computational Aesthetics: AI Approaches to Beauty & Happiness", Boston, MA, 2006.
- [23] Liu Y., Koenig S. and Furcy D., Speeding Up the Calculation of Heuristics for Heuristic Search-Based Planning, In Proceedings of the Eighteenth National Conference on Artificial Intelligence, 2002, pp. 484-491.
- [24] Magerko, B., Laird, J.E., Assarie, M., Kerfoot, A. and Stokes, D. AI Characters and Directors for Interactive Computer Games. Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI 2004), 2004, pp. 877-883.
- [25] Manish Mehta, Steven Dow, Michael Mateas, Blair MacIntyre, Evaluating a Conversation-Centered Interactive drama, Proceedings of the Sixth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2007), Honolulu, Hawaii, USA, May 2007.
- [26] Marsella, S. and Gratch, J. Modeling Coping Behaviour in Virtual Humans: Don't Worry, Be Happy. Proceedings of the Second International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2003), ACM Press, 2003, pp. 313-320.
- [27] Mateas, M., Stern, A. Natural Language Understanding in Façade: Surface-Text Processing. 2nd International Conference on Technologies for Interactive Digital Storytelling and Entertainment, TIDSE 2004, pp. 3-13.
- [28] Mateas, M., A Neo-Aristotelian Theory of Interactive Drama, Working Notes of the AAAI Spring Symposium on Artificial Intelligence and Interactive Entertainment, AAAI Press 2000.
- [29] Murray, J.H. Hamlet on the Holodeck: The Future of Narrative in Cyberspace, MIT Press, Cambridge, 1997.
- [30] Pizzi, D., Cavazza, M., and Lugrin J-L. Extending Character-based Storytelling with Awareness and Feelings. ACM Joint Conference on Autonomous Agents and Multi-Agent Systems, AAMAS 2007, Hawaii, May 2007.
- [31] Pizzi, D., Charles, F., Lugrin J-L., and Cavazza, M. Interactive Storytelling with Literary Feelings. The second International Conference on Affective Computing and Intelligent Interaction (ACII2007), Lisbon, Portugal, September 2007.
- [32] Riedl, M.O., Young, R.M. An Intent-Driven Planner for Multi-Agent Story Generation. Third ACM Joint Conference on Autonomous Agents and Multi-Agent Systems, New York, USA, 2004, pp. 186-193.
- [33] Ryan, M.L. Narrative as Virtual Reality: Immersion and Interactivity in Literature and Electronic Media. Baltimore: Johns Hopkins University Press, 2001.
- [34] Swartout, W., Hill, R., Gratch, J., Johnson, W.L., Kyriakakis, C., LaBore, C., Lindheim, R., Marsella, S., Miraglia, D., Moore, B., Morie, J., Rickel, J., Thiebaux, M., Tuch, L., Whitney, R., Douglas, J. Toward the Holodeck: Integrating graphics, sound, character and story. Proceedings 5th International Conference on Autonomous Agents (Agents-01), ACM Press, New York, 2001.
- [35] Zoltan-Ford, E., How to get people to say and type what computers can understand. The International Journal of Man-Machine Studies, 34, 1991, pp. 527-547.