Emotional Input for Character-based Interactive Storytelling

Marc Cavazza, David Pizzi, Fred Charles
University of Teesside, School of Computing
Borough Road, Middlesbrough
TS1 3BA, United Kingdom
{m.o.cavazza, d.pizzi, f.charles}@tees.ac.uk

Thurid Vogt, Elisabeth André
Multimedia Concepts and Applications
Faculty of Applied Informatics, University of Augsburg
Eichleitnerstr. 30, 86159 Augsburg, Germany
{thurid.vogt, elisabeth.andre}
@informatik.uni-augsburg.de

ABSTRACT

In most Interactive Storytelling systems, user interaction is based on natural language communication with virtual agents, either through isolated utterances or through dialogue. Natural language communication is also an essential element of interactive narratives in which the user is supposed to impersonate one of the story's characters. Whilst techniques for narrative generation and agent behaviour have made significant progress in recent years, natural language processing remains a bottleneck hampering the scalability of Interactive Storytelling systems. In this paper, we introduce a novel interaction technique based solely on emotional speech recognition. It allows the user to take part in dialogue with virtual actors without any constraints on style or expressivity, by mapping the recognised emotional categories to narrative situations and virtual characters feelings. Our Interactive Storytelling system uses an emotional planner to drive characters' behaviours. The main feature of this approach is that characters' feelings are part of the planning domain and are at the heart of narrative representations. The emotional speech recogniser analyses the speech signal to produce a variety of features which can be used to define ad-hoc categories on which to train the system. The content of our interactive narrative is an adaptation of one chapter of the XIXth century classic novel, *Madame Bovary*, which is well suited to a formalisation in terms of characters' feelings. At various stages of the narrative, the user can address the main character or respond to her, impersonating her lover. The emotional category extracted from the user utterance can be analysed in terms of the current narrative context, which includes characters' beliefs, feelings and expectations, to produce a specific influence on the target character, which will become visible through a change in its behaviour, achieving a high level of realism for the interaction. A limited number of emotional categories is sufficient to drive the narrative across multiple courses of actions, since it comprises over thirty narrative functions. We report results from a fully implemented prototype, both in terms of proof of concept and of usability through a preliminary user study.

Cite as: Emotional Input for Character-based Interactive Storytelling, Marc Cavazza, David Pizzi, Fred Charles, Thurid Vogt, Elisabeth André, *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009),* Decker, Sichman, Sierra and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. 313–320
Copyright © 2009, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org), All rights reserved.

Categories and Subject Descriptors

H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems.

General Terms

Algorithms, Human Factors.

Keywords

Interactive Narrative, Embodied Conversational Agents, Affective Interfaces.

1. INTRODUCTION

Interactive Storytelling (IS) is one of the main application areas for virtual actors integrating all aspects of deliberation, communication and expressivity. Whilst much progress has been achieved in narrative generation techniques, user interaction is increasingly appearing as a major bottleneck in the development of aspects of IS technologies. Interaction takes place primarily with virtual actors and is largely determined by the IS paradigm, which dictates the level of user involvement as well as the preferred modality of interaction with artificial actors and their environment. IS approaches can be classified according to the mode and intensity of involvement of the user, from occasional interventions in a spectators' perspective [3] [26] [33] [34] to total immersion into the narrative tending towards the Holodeck™ paradigm [20] [27]. The dominant modality of interaction is language, whether input as written text or spoken, as influence [3], or dialogue [18]. Although the narrative context can help focus the level of linguistic processing required, the understanding of user input remains beyond the state-of-the-art and in particular systems are difficultly scalable. The most challenging case is that of total immersion [27], as the need for anytime language processing coincides with aesthetic constraints, namely that user utterances should comply with the style of language expected from the narrative genre considered for the interactive narrative. This may be a condition for a realistic experience but soon becomes insurmountable in terms of Natural Language Processing (NLP) techniques. Another dimension of linguistic input is the use of spoken or written modality. The written modality undoubtedly brings additional robustness and supports various degrees of processing from shallow NLP [18] to the use of chatterbots [31]. It is better use in a dialogue mode, to preserve some form of realism and some appeal to the interaction. However, it inevitably faces limitations in terms of IS paradigms supported, as it is not compatible with immersion in the story, or a feeling of acting.

In this paper, we propose an approach to user interaction with virtual characters entirely based on emotional speech (Figure 1). This approach supports the unrestricted use of language by users and has been readily integrated within an IS system itself based on a form of Emotional Planning [8] [23]. The background narrative for our IS system is an adaptation of three chapters of the XIXth century classic *Madame Bovary* by Gustave Flaubert [6] (more specifically chapters 9-12 of Part II). In the next section, we discuss related work, before introducing our character-based storytelling approach based on emotional planning. We then describe the principles behind emotional speech recognition and how this can be naturally integrated within our IS system. We conclude by presenting early results from user testing of the fully-implemented system.

Figure 1. User Interaction in the EmoEmma Demonstrator.

2. RELATIONSHIP TO PREVIOUS WORK

Several Interactive Storytelling systems to date have incorporated natural language interaction, whether the underlying paradigm was one of complete user involvement [27] or user influence [3]. There is no previous report of large vocabulary speech-based system within IS, outside a strong context, generally task-oriented, such as the *Mission Rehearsal Exercise* (MRE) [3] [28] or *Justine* [10]. This is easily explained by the difficulties of speech understanding. The case of written interaction is slightly different and has been part of *FearNot!* [14] or *Façade* [18]. In *Fearnot!* written input was interpreted in terms of speech acts used to influence or comfort the virtual agent. *Façade* adopted a theatre-like environment heavily based on dialogue, which maintained user interest high, through a strong integration of narrative representations to dialogue acts. Emotional Planning has been originally described by Gratch [8] as based on emotional activation assessing threats to goal satisfaction during planning, which was later refined through the introduction of appraisal and coping [17]. More relevant to our context, they have described the use of emotions to alter beliefs [16]. The role of emotions in interactive narrative has also been discussed by Rank and Petta [24]. This has been primarily used in Interactive Storytelling applications, although not strictly speaking on those for which narrative was the main focus. Applications such as *MRE* or *FearNot!* were training or simulation systems for which emotional planning primarily conferred believability to the virtual actors but was not intended to address narrative or aesthetic issues. Both systems made use of emotional planning but did not attempt to integrate emotional language processing and emotional planning within a unified framework.

In order to increase a player's level of immersion and engagement (see [7]), a number of affective games encourage a gamer to express his or her emotive states and dynamically adapt actions and events to them. In *SenToy*, the user interacts with a tangible doll to communicate one of six emotions through gestures (see [22]). Another example of an affective interface is the computer game *FinFin* where the user may influence the emotional state of a half bird, half dolphin creature via talking and waving. [5] conducted an experiment in order to investigate how to induce an optimal state of arousal in the *Tetris* game. Surprisingly, hardly any attempts have been made to adapt a story to a user's emotional state. In the commercial game *"Through the eyes of the girls"* produced by Girland[1], young girls may use an emotion wheel to input their emotive states which then influence the events in a dating scenario. The manual input of emotional states leads, however, to an interruption of the story and is likely to negatively influence the gamers' experience. [19] developed a system for children with learning disabilities that monitors the children's emotional state using skin conductivity sensors and adapts the behaviour of pedagogical agents and the difficulty of tasks to it. Even though the system makes use of narrative, the story is not directly influenced by the children's emotional state. Rather, emotional monitoring is used to increase the children's learning performance as opposed to driving a story. A system which makes use of vocal emotion recognition in games has been presented by [9]. Here, the gamers navigate through obstacles by high or low arousal in their voice. If the gamers' voice portrays a negative emotion, the character shrinks. Thus it is able to get through small gaps. In the case a positive emotion is expressed, the character grows and gets stronger so that it is, for example, able to destroy a wall. Even though the gamer is able to influence events in the game, the system does not make explicit use of narrative like we report in this paper.

3. EMOTIONAL PLANNING FOR INTERACTIVE STORYTELLING

Expressing emotions may be an important aspect of the believability of any virtual actor, but the actual relationship between emotions and narrative is of much greater sophistication. This is best understood by considering that, within novels themselves, the psychology of characters is not usually described at a cognitive level, in terms of basic emotions, but tends to be intertwined with the literary presentation of the text. In other words, the psychology of its characters is represented within the narrative through their feelings. These feelings tend to form part of a finer-grained ontology than commonly described emotions, and depart from traditional emotional models, e.g. of Ekmanian emotions. Using such feelings for narrative representations would bring a new perspective for character-based IS: i) character's feelings would determine their actions, ii) interactions between characters (determined by one character acting upon an another

[1] http://www.girland.com/

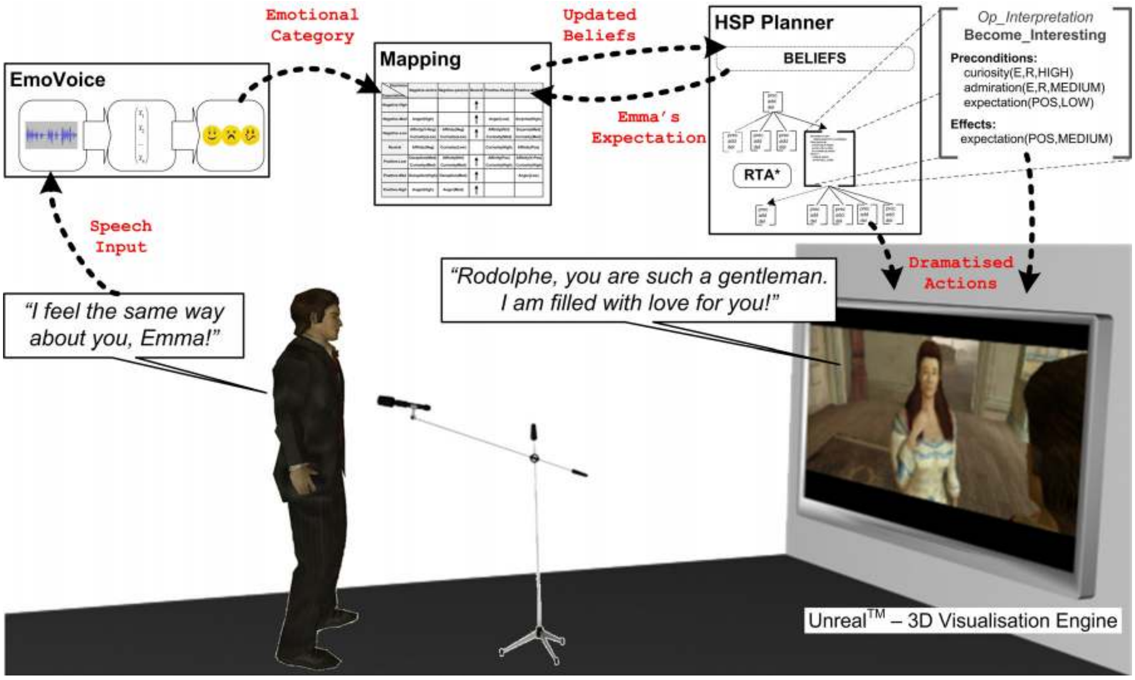

Figure 2. System Overview: the User Interacts with the Narrative by Impersonating *Rodolphe*.

according to its own feelings) will modify and update their respective feelings, and iii) user intervention is similar in nature to an interaction between virtual characters, the user utterance being interpreted as the same type of communicative actions between characters which modify their respective feelings.

The general principle underlying our storytelling engine is the following: the basic representation is the current set of narrative feelings for any given character, hence the planning domain itself is derived from the set of possible feelings (Figure 2). Planning operators determine which actions a character may undertake based on its feelings; for instance, Emma may declare her love for Rodolphe. In turn, evolution of feelings is driven by the interactions between characters (or with the user). In that sense, Emotional Planning is implemented by having the planning domain itself based on an ontology of narrative feelings, rather than associating emotions to plan progression [8]. Each feature character in the interactive narrative would be under the control of its own Planner: in the case of *Madame Bovary*, for those chapters we have adapted, this is only required for Emma, as the user plays the role of Rodolphe, and Charles' role is so simple as to be compatible with direct scripting. A Planner can be defined through its representational formalism, its algorithm and the set of facts over which it operates, or planning domain. Defining the set of feelings for the characters would appear even a greater challenge than the one normally faced with formalizing a traditional narrative to adapt it to IS. Fortunately, it so happens that recent research in literary studies has uncovered, amongst Flaubert's preparatory work for the novel and as part of his drafts, a detailed description of the characters' psychology, in particular for Emma Bovary, down to a specification of characteristic feelings [12], sometimes extremely specific, such as *emboldened-by-love*, *irritated-by-vice*, or *jealousy-curiosity* [12].

Our Planning algorithm is based on a standard Heuristic Search Planning (HSP, [1] [2]) approach with standard heuristic calculation through the VI algorithm [13]. As traditionally implemented for IS systems, the planner has to cope with dynamic environments, because of interactions with other characters or user intervention. To support real-time planning within an HSP approach it is sufficient to use a real-time heuristic search algorithm as the underlying search algorithm, and in our case we have implemented a simple variant of RTA* [11] [23]. One variant is our system is that the definition of planning goals is not fixed for the whole duration of the narrative. The goal against which HSP searches for the best next operator is defined as a conjunction of world states and character's feelings, and can be updated by the addition or the deletion of predicates. This derived from reflections on the representational nature of goals in interactive narratives, and the finding that it was too strong an assumption to equate a goal state to the narrative "ending", especially when various characters have different goals. Anecdotically, in the case of *Madame Bovary*, her final suicide cannot be said to respond to her long-term goal of escaping boredom, as it was prompted by despair of her reputation and her financial situation. We have thus substituted the weakest notion of drivers to goals to account for that finding, although this does not affect the algorithmic process underlying plan progression. We have further specified planning operators as belonging to three categories depending on the type of modification they operate on the planning domain and the actions they trigger on the virtual stage. Physical operators determine character motion and navigation, for instance moving to a room where another character is waiting so as to render interaction possible, or leaving the room to manifest discontent. Communication operators produce communicative actions which aim at influencing another character's feelings. Finally, Interpretation operators analyse the emotional contents of dialogue act performed by another

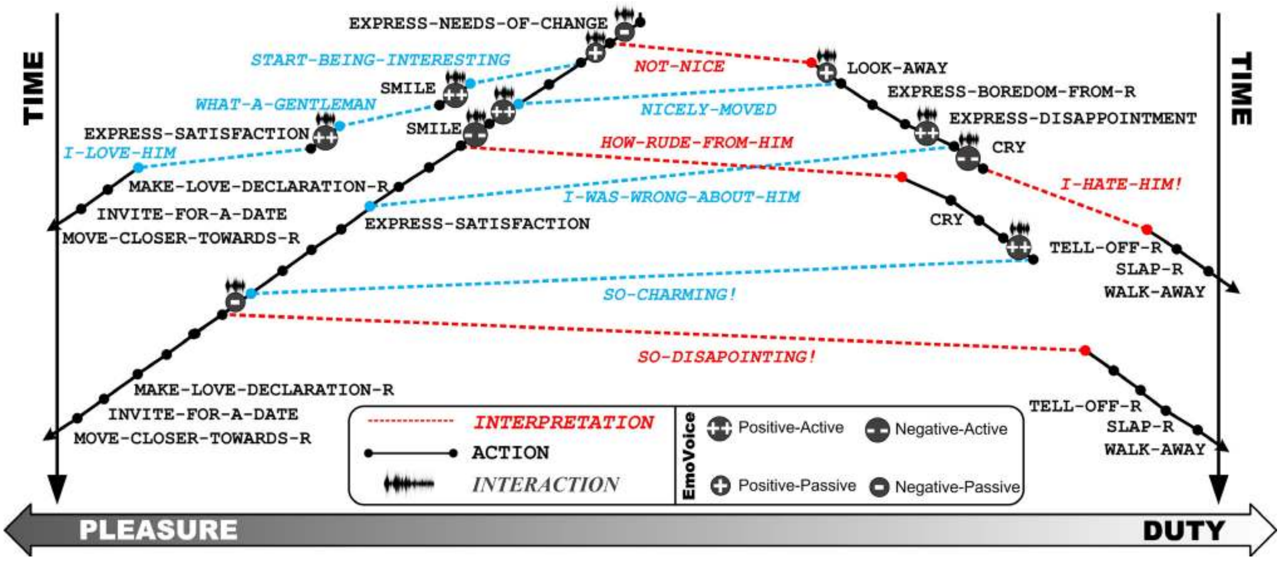

Figure 3. Story Variability and Impact of Emotional Speech: the diagram represents different evolutions depending on the emotional categories recognised. Multiple opportunities for interaction leverage the impact of Emotional Speech, and account for significant variability despite the limited number of emotional categories.

character (or the user) to determine, taking into account the current emotional context, its actual impact. The same search algorithm at the heart of our HSP implementation can deal with all types of operators in a uniform fashion since their pre-conditions and post-conditions are based on the same elements of the planning domain.

Figure 3 represents a story progression diagram adapted to our approach, in which the narrative progression towards one particular type of ending is accounted for by the list of operators activated within a particular plan generated for Emma Bovary, being the feature character. To characterize how the interactive narrative may depart from the baseline story, we have designed a dimensional model opposing *duty* to *pleasure*, these characterising the driving forces behind Emma's behaviour. This ending only characterizes the three chapters under consideration. What is at stake here is whether Emma will be unfaithful or not, but the interactive narrative still admits multiple variants: she can remain faithful to her estranged husband, reconcile herself with him, take Rodolphe as a lover, or even escape from Yonville with him…

4. EMOTIONAL SPEECH RECOGNITION

EmoVoice identifies affect conveyed by the voice. No semantic information is extracted - the recognition relies on the acoustic signal only. For the integration into the showcase, this has to be done in real-time while the user is interacting, which so far has scarcely been attempted.

The major steps in speech emotion recognition are audio segmentation, which means finding appropriate acoustic segments as emotion classification units, feature extraction to find those characteristics of the acoustic signal that best describe emotions and to represent each segmented acoustic unit as a (series of) feature vector(s), and lastly the actual classification of the feature vectors into emotional states.

EmoVoice, our toolkit for vocal emotion recognition, consists of two modules, one for the offline creation and analysis of an emotional speech corpus, and one for the online tracking of affect in voice while someone is talking (see Figure 4). The first module is a set of tools for audio segmentation, feature extraction, feature selection and classification of an emotional speech corpus, and a graphical user interface to easily record speech files and create a classifier. This classifier can then be used for the second module, the online emotion recognition. Here, classification results are obtained continuously during talking, there is no "push-to-talk".

Both the offline creation and analysis of the emotional speech corpus and the real-time recognition of vocal emotions are a three-step process. First, the acoustic input signal coming continuously from the microphone is segmented into chunks by Voice Activity Detection (VAD), which segments the signal into speech frames with no pauses within longer than about 0.5 seconds. Next, from this speech frame, a number of features relevant to affect are extracted. The features are based on pitch, energy, Mel Frequency Cepstral Coefficients (MFCC), the frequency spectrum, the harmonics-to-noise ratio, duration and pauses. The actual feature vector is then obtained by calculating statistics (mean, maximum, minimum, *etc.*) over the speech frame ending up with around 1300 features. A full account of the feature extraction strategy can be found in [30].

In the last step, the feature vector is classified into an affective state. Currently, two classification algorithms are integrated in EmoVoice: a naïve Bayes (NB) classifier and a support vector machine (SVM) classifier (from the LibSVM library [4]). The NB classifier is very fast, even for high-dimensional feature vectors, and therefore especially suitable for real-time processing. However, it has slightly lower classification rates than the SVM classifier which is a very common algorithm used in offline emotion recognition. In combination with feature selection and thereby a reduction of the number of features to less than 100, SVM is also feasible in real-time.

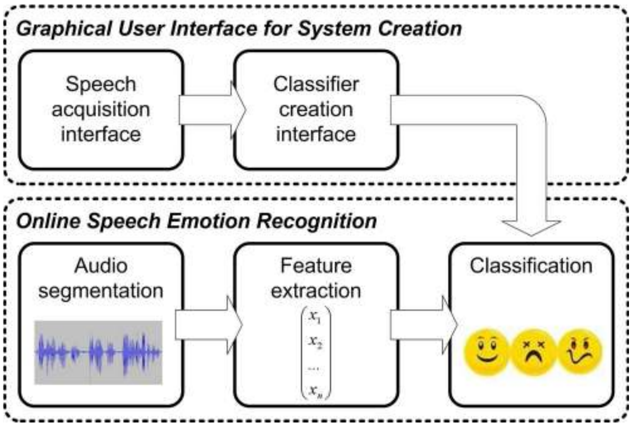

Figure 4. The EmoVoice Emotional Speech Recogniser comprises an offline system to to generate classifiers and an online system to be integrated in IS applications.

EmoVoice integrates an easy-to-use interface for recording and training an emotional speech corpus, which is meant to increase accuracy in application-dependent contexts. The method used for emotion elicitation was inspired the Velten mood induction technique [29] as used in [32] where subjects have to read out loud a set of emotional sentences that should set them into the desired emotional state. The system comes with a list of such sentences, which we have completed with actual excerpts from *Madame Bovary*'s dialogues. For our IS application we have concentrated on a small set of five categories (each corresponding to combinations of valence and arousal): *NegativeActive*, *NegativePassive*, *Neutral*, *PositiveActive* and *PositivePassive*. The rationale for such a reduced set of emotional inputs has been that these categories will be further interpreted, taking into account the context in which they are recognised. However, developers making use of the system are encouraged to change sentences according to their own emotional experiences: For a good speaker dependent system, about 40 sentences per emotion usually suffice [30]. We have initially trained EmoVoice with three subjects using various test sentences, some of which extracted from the actual dialogues of *Madame Bovary*, with an average 40 sentences per category. Overall we have achieved a recognition score of 66% for those five categories, obtained with speakers outside those having contributed to system training. This score is consistent (and probably on the upper end) with those previously reported for the EmoVoice system [30]. The implications of this recognition score for overall system performance will be discussed in section 6.

5. SYSTEM ARCHITECTURE

The prototype comprises the traditional elements of an IS system, namely a narrative engine, a visualisation module, and an interaction module (see Figure 2). The narrative engine consists of the real-time HSP planner described above, which only controls the main character, Emma Bovary. Story visualisation is based on the Unreal Tournament 2003™ game engine, which supports staging and character animation, including some form of facial animation allowing more realistic close-ups of the characters during dialogue scenes, which constitute the vast majority of the narrative action in that genre (Emma's dialogue being conveyed using a Text-To-Speech system). Communicative actions produced by the planner activate Unreal scripts controlling Emma's animations as well as sound generation. The interaction module consists of the emotional speech recogniser EmoVoice described above [30]. The system's philosophy, which integrates emotional planning for IS with emotional speech recognition, is implemented through a mapping between the latter modules output and input. Such a mapping relies on the contextual interpretation of a reduced number of emotional categories extracted from the user's speech (as described in the above section): *NegativeActive*, *NegativePassive*, *Neutral*, *PositiveActive* and *PositivePassive*. These can be translated into modifications of the character's emotional states via the notion of expectation. Communicative actions triggered by Emma are dictated by the current goal but can also be categorised according to the type of response expected in context. The notion of expectation, attached to a communicative action, is thus used to relate narrative context to the interpretation of user input.

EmoVoice / Expectation	Negative-*Active*	Negative-*Passive*	Positive-*Passive*	Positive-*Active*
Negative-*High*				
Negative-*Med*	Anger(High)		Anger(Low)	Surprise(High)
Negative-*Low*	Affinity(V-Neg) Curiosity(Low)	Affinity(Neg) Curiosity(Low)	Affinity(Ntrl) Curiosity(Med)	Surprise(Med) Curiosity(Med)
Neutral	Affinity(Neg)	Curiosity(Low)	Curiosity(High)	Affinity(Pos)
Positive-*Low*	Disappointment (Med) Curiosity(Med)	Affinity(Ntrl) Curiosity(Med)	Affinity(Pos) Curiosity(High)	Affinity(V-Pos) Curiosity(High)
Positive-*Med*	Disappointment (High)	Disappointment (Med)		Anger(Low)
Positive-*High*	Anger(High)	Anger(Med)		

Figure 5. Matrix for the Contextual Interpretation of Emotional Input. The character's expectations determine the actual interpretation of a given input.

User utterances are interpreted contextually as a function of the relation between the EmoVoice category and the character's expectation. The affective response to the user's reply is amplified by Emma's current emotional status: for instance, a lukewarm attitude from Rodolphe would upset Emma all the more that her expectations run higher at any given stage. For instance, in case of high expectations from Emma, *NegativePassive* and *NegativeActive* utterances will be interpreted as feelings of disappointment, of levels of intensity determined by the *Active*/*Passive* component. This is also a mechanism to incorporate the dynamics of the relationship, as expectations would vary according to the status and progression of the characters' relationship throughout the narrative. It makes so that a similar affective response would have dramatically different effects at various stages of the unfolding narrative. Contextual interpretation is determined by the matrix depicted on Figure 5. It should be noted that this matrix only contains generic associations (such as the one between surprise and low expectations, or between disappointment and high expectations), and that the actual selection of operators may involve an additional contextual elements, which is the specific nature of the relation. For instance, in the genre considered, disappointment about a romantic relationship may lead to estranging, etc.

6. EXPERIMENTAL RESULTS AND USER EVALUATION

In order to evaluate the system, we have conducted tests with 14 subjects. The setting consisted in the interactive narrative being displayed on a 30'' screen, with a high-quality microphone positioned in front of it. Subjects were first asked to read aloud several excerpts from the original novel's dialogues, in order to test optimal acoustic signal strength for emotional speech recognition. They were given instructions describing the narrative, the part they were supposed to play impersonating Rodolphe, and the fact that Emma would react to the emotional content of their responses (however, they were not given any detail on the actual techniques underlying the system, such as the fact that it did not recognise word meaning). The IS system was then started, generating real-time 3D animations, with a voice over giving the background for the early stages of the narrative at which no interaction was allowed. The user had no control over navigation of his character, and was presented the stage in third-person mode, with Rodolphe as his avatar. An automatic camera system (part of the visualisation engine) was centred on Emma and would follow her on stage: an example of camera positioning from a dialogue phase is visible on Figure 5.

As the IS system starts, generating the first encounters being Emma and Rodolphe, the user can either address Emma spontaneously or respond to one of her questions or declarations which are enacted through a corresponding animation with Text-To-Speech voice synthesis. After each utterance from Emma, the user has the choice of responding her with various level of enthusiasm, empathy or disapprobation, or not to respond, which in some cases will also give raise to an interpretation, based on the level of Emma's expectation. The user can experience the subsequent unfolding of the interactive narrative, whether he continues to interact or not: his replies may show immediate or deferred effects, or no effects at all. Multiple interactions are allowed throughout the scene, as Emma repeatedly addresses Rodolphe as part of her role/plan. At no stage does the user receive any indication of the emotional category perceived from his utterance, and his only feedback is via the interactive narrative itself.

All 14 subjects successfully completed the experiment, which resulted in each case in a complete session, generating an interactive narrative until its normal ending. The average duration of the interactive narrative was 2.9 minutes (with extremes varying from 2 to 6 minutes) and ended up with either Emma leaving the stage in despair ("negative" ending), or engaging with Rodolphe ("positive" ending, which actually occurs in the original novel). Subjects were not instructed to favour a particular outcome, nor were they described any given outcome as normal: as a result, their interventions were balanced in nature, leading to an almost equal split between each possible ending (57% positive versus 43% negative). The actual sequence of narrative events was of much greater variability and its constituency depended on the nature and number of user interventions. Longer stories emerged as the user gave successive contradicting messages, which lead Emma through opposite feelings, provided none is so extreme as to accelerate the ending. In a similar fashion, high intensity emotional categories (active), regardless of their valence tended to lead more quickly to the story ending. An alternative explanation would correspond to users trying to correct the impact of EmoVoice recognition errors, which they perceive through inappropriate responses from the Emma character, by repeating a similar type of utterance to the one they see as having been unsuccessful. However, because of the relative robustness of the system, and the rather unconstrained nature of the experiments, these contradicting messages could correspond to exploratory behaviour by the subjects.

"I hope it is, Emma, and I hope it is with you."
"No, your life is definitely not what it was meant to be."
"I feel the same way about you, Emma."
"Of course, I'd declared my love for you many times."
"We must leave together now and make a new life."

Figure 6. Example User Utterances.

One explanation for the overall system robustness despite a 66% emotion recognition score can be found in the actual type of recognition errors. A study of system logs during user experiments showed that the most severe errors, in which opposite valence categories such as *NegativeActive* are recognised instead of the reference *PositiveActive*, only occur in about 5% of utterances. Most of the errors do not affect valence and, notwithstanding the value of expectations at the point at which they occur, tend to produce similar results in terms of narrative impact, or more often have no impact at all, and the dynamics of the story offers 'second chances' for correcting this seamlessly.

These evaluations do not aim at measuring the intrinsic aesthetic quality of the novel, but tend to validate the overall concept, and assess user engagement with the system. The average number of interactions (user utterances) during a session was 7.4 ± 5, and there was no clear correlation between interactive narrative duration and the number of interventions (which can be accounted for by the redundancy of some interventions). The average length of user utterances was 7.5 words with a significant proportion of utterances exceeding 10 words, again suggesting that the users were comfortable interacting with the system (see Figure 6).

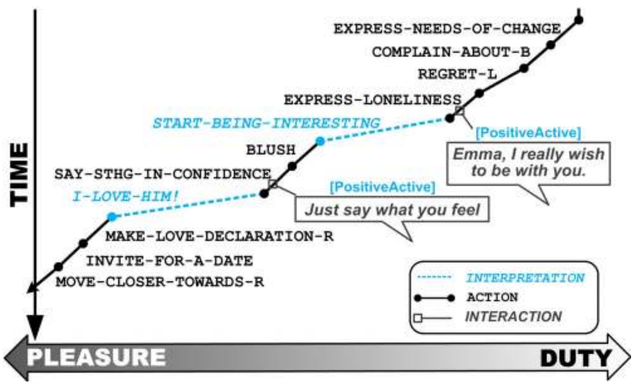

Figure 7. In this session, the user is interacting freely with Emma, although not always in a consistent fashion. This results in some instability in the evolution of the narrative as well as an unexpected ending from the early evolution of the story. This particular example demonstrates the overall stability of the system in the face of unrestricted interactions.

For each experiment, we record the step-by-step evolution of the generated narrative, by logging the planning operators activated, together with the emotional categories characterizing each user intervention (in addition to the utterance itself on a separate channel). These data can be represented as the narrative evolution diagram of Figure 7.

Q1:	*I had the feeling that Emma understood what I was saying.*
Q2:	*I had the feeling that Emma was responding emotionally to what I was saying.*
Q3:	*I had the feeling that Emma was expressing emotions.*
Q4:	*Emma's speech reflected the changes in the story.*

Figure 8. Questionnaire Assessing the IS experience.

Finally, each subject is handed a questionnaire about his experience with the interactive narrative, whose questions are reproduced on Figure 8. In terms of the overall experience, subjects responded very positively to the installation (Figure 9), confirming their perception of Emma Bovary as a believable character (Q3: 3.7 ± 0.7 ; Q4: 3.9 ± 0.6), responding appropriately to their interaction (Q1: 3.6 ± 0.7 ; Q2: 3.9 ± 0.7).

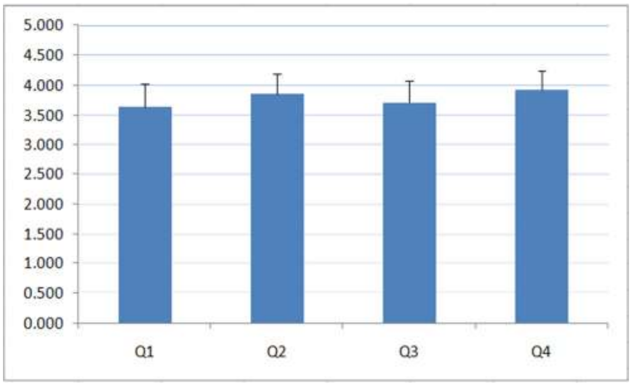

Figure 9. Results from the User Experience Questionnaire.

7. CONCLUSIONS

We have introduced a new approach to IS, in which affective interaction allows unconstrained linguistic expression, as part of a dialogue with the feature character of the interactive narrative (although strictly speaking dialogue is restricted to pairs of utterances, without any extended dialogue phenomena). Current limitations of this work arise from its definition of emotional categories, and the fact that their impact depend on genre considerations: the novel supporting our experiments contained many elements (such as the character's expectations and interpretation of situations) which contributed to story generativity, despite a reduced number of emotional categories for input. The prototype we described implements a single IS paradigm, which is one of constant involvement of the user, where the user actually plays the role of one character. Whilst the conditions for its successful use would require a more thorough investigation, such an approach already opens new perspectives, in particular for the adoption of IS techniques as part of digital entertainment systems, as it could operate without the need for large-scale NLP. This could favour the long-awaited adoption of IS technologies for digital entertainment or edutainment, but would also benefit research in the field, by providing a stable framework for interaction whilst investigating other open challenges such as multiple plots or the relations between narrative generation and presentation.

8. ACKNOWLEDGMENTS

This work has been funded in part by the European Commission under grant agreements CALLAS (FP-ICT-034800) and IRIS (FP7-ICT-231824). We would like to thank Jean-Luc Lugrin for the development of the game engine modification.

9. REFERENCES

[1] Bonet, B., Geffner, H. 1999. Planning as Heuristic Search: New Results, Proceedings of the European Conference on Planning (ECP'99), pp. 360-372.

[2] Bonet, B., Geffner, H. 2001. Planning as Heuristic Search. Artificial Intelligence Special Issue on Heuristic Search, 129, n.1, 2001, pp. 5-33.

[3] Cavazza, M., Charles, F. and Mead, S.J., 2002. Character-based Interactive Storytelling. IEEE Intelligent Systems, 17, 4, 2002, pp. 17-24.

[4] Chang, C.-C. and Lin, C.-J. 2001. LIB-SVM: a library for support vector machines. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.

[5] Diener, H. and Oertel, K. 2006. Experimental Approach to Affective Interaction in Games. Edutainment 2006: pp. 507-518.

[6] Flaubert, G. 1856. Madame Bovary. La revue de Paris (Ed.), France. (in French).

[7] Freeman, D. 2003. Creating Emotion in Games: The Craft and Art of Emotioneering. New Riders. 2003.

[8] Gratch, J. Why you should buy an emotional planner, Proceedings of the Autonomous Agents 1999 Workshop on Emotion-based Agent Architectures (EBAA'99).

[9] Jones, C. and Sutherland, J. 2007. Acoustic emotion recognition for affective computer gaming. In Peter, C. and Beale, R., editors, Affect and Emotion in Human-Computer Interaction, volume 4868 of LNCS. Springer, Heidelberg, Germany.

[10] Kenny, G., P., Parsons, T., D., Gratch, J., Rizzo, A., 2008. Evaluation of Justina: A Virtual Patient with PTSD. In Proceedings of the conference on IVA 2008, pp. 394-408.

[11] Korf, R.E. Real-time heuristic search. Artificial Intelligence, 42:2-3, 1990, pp. 189-211.

[12] Leclerc, Y (Ed.). Plans et Scenarios de Madame Bovary, CNRS Editions, France, 1995 (in French).

[13] 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, pp. 484-491, 2002.

[14] Louchart, S., Romano, D., M., Aylett, R., and Pickering, J. 2004. Speaking and acting - interacting language and action for an expressive character. In Proceedings for the AISB workshop, University of Leeds, UK.

[15] Lyons, G., Sharma, P., Baker, M., O'Malley, S., and Shanahan, A. 2003. A computer gamebased EMG biofeedback system for muscle rehabilitation. In: Engineering in Medicine and Biology Society, 2003.

Proceedings of the 25th Annual International Conference of the IEEE. volume 2. S. 1625–1628. Sept. 2003.

[16] Marsella, S. and Gratch, J. 2002. A Step Towards Irrationality: Using Emotion to Change Belief, in Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems, Bologna, Italy, July 2002.

[17] Marsella, S. and Gratch, J. 2003. Modeling coping behavior in virtual humans: don't worry, be happy. In Proceedings of the Second international Joint Conference on Autonomous Agents and Multiagent Systems (Melbourne, Australia, July 14 - 18, 2003). AAMAS '03. ACM, New York, NY, pp. 313-320.

[18] Mateas, M., Stern, A., 2004. Natural Language Understanding in Façade: Surface-Text Processing. TIDSE 2004: 3-13.

[19] Mohamad, Y., Velasco, C., A, Berlage, T. 2003. An Approach toward building of adaptive training and therapeutic systems considering the users affective state. HCI International Proceedings 2003, Crete, Greece.

[20] Murray, J. 1997. Hamlet on the Holodeck: The Future of Narrative in Cyberspace. MIT Press, Cambridge, 1997.

[21] Nazir, A., Lim, M., Y., Kriegel, M., Aylett, R., Cawsey, A., Enz, S., Rizzo, P. and Hall, L. 2008. ORIENT: An Inter-Cultural Role-Play Game. In Narrative in Interactive Learning Environments NILE 2008 Conference.

[22] Paiva, A., Chaves, R., Piedade, M., Bullock, A., Andersson, G., and Höök, K. 2003. Sentoy: A tangible interface to control the emotions of a synthetic character. In: Proc. of AAMAS 2003. pp. 1088–1089. 2003.

[23] Pizzi, D., Charles, F., Lugrin J-L., and Cavazza, M., 2007. Interactive Storytelling with Literary Feelings. In: Paiva, A.C.R., Prada, R., Picard, R.W. (eds.) ACII 2007. LNCS, vol. 4738, pp. 630-641.

[24] Rank, S., Petta, P. and Trappl R. 2006. Features of Emotional Planning in Software Agents, in Dellariccia G., et al.(eds.), Decision Theory and Multi-Agent Planning, Springer Wien/New York, 2006.

[25] Rehm, M., Vogt, T., Wissner, M. and Bee, N. 2008. Dancing the Night Away — Controlling a Virtual Karaoke Dancer by Multimodal Expressive Cues. In Proceedings of AAMAS, 2008.

[26] Riedl, M.O., Young, R.M. 2004. 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.

[27] Swartout, W., Gratch, J., Hill, R. Hovy, E., Marsella, S. Rickel, J. and Traum, D., 2006. Toward Virtual Humans in AI Magazine, v.27(1).

[28] Traum, D., Rickel, J., Gratch, J., and Marsella, S. 2003. Negotiation over tasks in hybrid human-agent teams for simulation-based training. In Proceedings of the Second international Joint Conference on Autonomous Agents and Multiagent Systems (Melbourne, Australia, July 14 - 18, 2003). AAMAS '03. ACM, New York, NY, pp. 441-448.

[29] Velten, E. 1968. A laboratory task for induction of mood states. Behavior Research & Therapy, (6):473-482.

[30] Vogt, T., André, E. and Bee, N. 2008. EmoVoice - A framework for online recognition of emotions from voice. In Proceedings of Workshop on Perception and Interactive Technologies for Speech-Based Systems, Springer, Kloster Irsee, Germany, (June 2008).

[31] Weiß, S., Müller, W., Spierling, U., Steimle, F., 2005. Scenejo – An Interactive Storytelling Platform. In: Virtual Storytelling, Using Virtual Reality Technologies for Storytelling, Third International Conference, Proceedings, Strasbourg, France, pp. 77—80.

[32] Wilting, J., Krahmer, E., and Swerts, M. 2006. Real vs. acted emotional speech. In Proceedings of Interspeech 2006 | ICSLP, Pittsburgh, PA, USA.

[33] Young, R., M. 2000. Creating Interactive Narrative Structures: The Potential for AI Approaches. AAAI Spring Symposium in Artificial Intelligence and Entertainment. AAAI Press, 2000.

[34] Young, R., M. 2001. An Overview of the Mimesis Architecture: Integrating Intelligent Narrative Control into an Existing Gaming Environment, The Working Notes of the AAAI Spring Symposium on Artificial Intelligence and Interactive Entertainment, Stanford, CA, March 2001.