Under the Influence: using Natural Language in Interactive Storytelling

Marc Cavazza, Fred Charles and Steven J. Mead
School of Computing and Mathematics, University of Teesside, TS1 3BA Middlesbrough, UK
{m.o.cavazza, f,charles, steven.j.mead}@tees.ac.uk

Abstract: Interacting in natural language with virtual actors is an important aspect of the development of future Interactive Storytelling systems. We describe a paradigm for speech interfaces in interactive storytelling based on the notion of influence. In this paradigm, the user is mainly a spectator who is however able to interfere with the course of action by issuing advice to the characters. This is achieved by recognising corresponding speech acts and mapping them to the plans which implement characters' behaviours in the story. We discuss some examples based on a preliminary, yet fully implemented, prototype.

Key words: Interactive Storytelling, Virtual Actors, Speech Understanding.

1.   INTRODUCTION

In recent years, there has been considerable interest in developing Artificial Intelligence techniques for Interactive Storytelling [Young, 2000] [Swartout et al., 2001] [Mateas, 1999]. Substantial work has been dedicated to modelling plot and character behaviour to support the propagation of causal changes due to user intervention. Several paradigms of interactive storytelling are emerging, which can be classified according to two major dimensions: i) the nature of user involvement in the story and ii) the story generation mechanism (plot-centered or character-centered) [Sgouros et al., 1996] [Mateas, 1999]. An important objective of interactive storytelling is to enable natural modalities for user intervention: in this regard, interacting in natural language is certainly one of the most important endeavours in interactive storytelling research.

1

The paradigm we want to experiment with would implement the real-life experience of an audience shouting advice at their favourite characters when watching a film. Only that, in the case of interactive storytelling, the actors would actually be influenced by what the audience says. As this is obviously a long-term research goal our first prototype implements a simplified version of this concept. It consists in an interactive storytelling system, accepting speech input for a single user using a traditional microphone. The system presents itself as a real-time 3D animation in which the camera follows the feature character, implemented using the Unreal™ game engine. The animation behaves basically as a "silent movie" with character dialogue and important commentaries appearing as overlay text. The user can address the characters using a speech recognition system to give them advice, warning or information, which would alter their behaviour, hence the unfolding of the plot (Figure 1). This influence paradigm can thus be characterised as the ability for the user to interfere with the ongoing action, from an external position to the story, by intervening at any time, without being prompted for intervention by the system. Another important aspect is that user intervention is based on his current understanding of the story.

In this paper, we describe an ongoing research in the use of speech recognition to influence an interactive story. We first introduce the basic techniques underlying our system. We then describe an approach to natural language intervention derived from speech acts, and show how it relates to the mechanisms that support character behaviour and story generation. Throughout the paper, we will illustrate the discussion with examples actually implemented in our prototype.
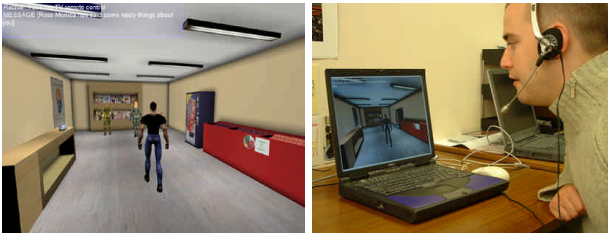

Figure 1. Giving spoken advice to characters.

2.   AI TECHNIQUES FOR CHARACTERS' BEHAVIOURS

Our system is a character-based system where each character's role is described as a plan. The plan actually contains several alternative actions for each sub-goal, making it a unified representation for all possible roles of a given character, which also facilitates the authoring of the baseline story. The planning technique we use is HTN planning [Erol et al., 1995] [Nau et al., 1998], which has been shown to be well suited to knowledge-intensive domains [Kambhampati and Hendler, 1992]. Our implementation supports interleaving planning and execution, as well as re-planning, which are essential requirements in the context of interactive storytelling [Young, 1999].

The test scenario we have been using is inspired from the popular "Friends" sitcom [Cavazza et al., 2001a]. The rationale for using a sitcom is that the story ending and intermediate situations are equally relevant, which constitutes a good test case for story generation and interaction. Several roles are defined as plans for each feature character. The decomposition of a plan into sub-goals reflects different stages of the action, while the lower layers correspond to various ways to achieve these goals. For instance, in order to invite Rachel out for dinner, Ross needs to acquire information about her, to gain her friendship, find a way to talk to her in private, etc. He his faced with several possibilities at each stage, e.g. to gain information he can steal her diary, ask one of her friends, phone her mother, etc., each of these sub-goals being further refined until they can be described in terms of terminal actions to be played in the virtual environment. Figure 2 represents the corresponding HTN for the main character, 'Ross'. Heuristic values are associated to the OR nodes, which can be used to bias search towards certain actions [Weyhrauch, 1997]. These heuristics constitute another narrative representation and can be altered dynamically to reflect e.g. change in emotional state of the character.

The same planning techniques support story generation, which arises from the dynamic interaction of the various characters' plans, and user intervention, which alters the plan executions. However, top-down planning is not the sole source of story variation: as the characters share a common environment, their independent activities generate emerging situations (competition for action resources, random encounters between characters) not explicitly foreseen by the plan descriptions, which contribute further variation to the plot. In addition characters have reactive mechanisms to cope with certain situations [Cavazza et al., 2001b]. For instance, Rachel
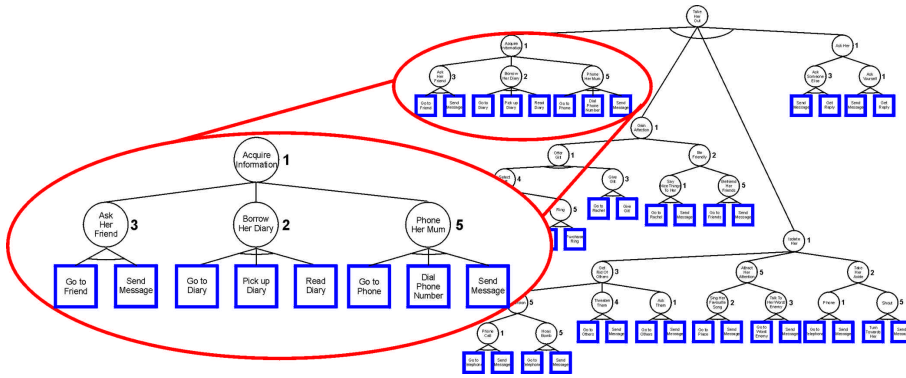
Figure 2. An HTN for the main character.

might get jealous whenever she sees Ross interacting with another female character.

3.   INTERACTING IN NATURAL LANGUAGE

The storytelling dimension of speech influence mostly consists in either contrasting or favouring the perceived actions of the virtual characters. The user could naturally issue direct commands to the virtual characters, for instance ordering Ross to go and talk to Phoebe. However, we wish to demonstrate that this would contradict the whole paradigm of user intervention as influence. A recent analysis of computer games by Bolter and Grusin [1998] has suggested that in most of these games, the player adopts simultaneously the role of a spectator and of the director (this can be grasped intuitively if one considers third-person adventure games). On the other hand, the idea of interactive storytelling is to depart from this director role to privilege the "active spectator" perspective. In that sense the spoken input should take the form of realistic advice rather than commands and be embedded within the story. For instance, rather than saying "go talk to Phoebe" the user will say something like "Phoebe has the information you need". An additional effect is that third parties watching the interactive story, but not interacting themselves, would get a better understanding of the interactive plot as spectators and hence a better experience.

These more natural forms of expression, based on implicit background information, characterise the influence paradigm of speech interaction as another implementation of speech acts. The speech act nature of spoken advice can be illustrated by considering the meaning of the same sentence in different contexts. An utterance such "Phoebe is in Rachel's room" will convey different information depending on the context in which it is uttered. If Ross is trying to reach Phoebe in order to obtain information about

Rachel, it will give him Phoebe's location (*information provision*). However, if Ross is trying to acquire the same information by stealing Rachel's diary in her room, it can also signal to Ross that he won't be able to do so, because Phoebe will object to that (*warning*). To compute the consequences of a speech act, the information it contains is to be matched to the various pre-conditions of task operators or to the executability conditions of terminal actions subsumed by the current active task.

4.   SPEECH PROCESSING

The system comprises a speech recognition step and a natural language processing step that carries out the actual speech act recognition and computes its influence on the character's plan. Speech processing is based on a commercial speech recognition system, the EAR^{TM} SDK from Babel Technologies. The recognised sub-language has to be defined through a context-free grammar, though the existence of optional symbols extends its descriptive power (and supports greater flexibility, such as the use of multi keyword spotting). Naturally in order to preserve the believability of the experience, the user cannot be imposed to learn the syntax of the sub-language recognised by the system. The solution to this problem is generally called habitability: it consists in encoding a sufficient number of surface variants of the main syntactic constructs. In this way the user can express himself relatively freely, without realising the constrained of the underlying language. This approach however is only practical is specific areas, but the kind of spoken advice used in interactive storytelling falls into that category.

Overall, processing can be described as a two-level template system that takes advantage of the recognition language definition used in the EAR^{TM} SDK. The first level identifies surface forms for the advice. The second level associates semantic features, which make possible to fill slots with selectional restrictions and instantiate templates for candidate speech acts. The example below shows an actual template for informing Ross about some gossip.

"?Character [Attrib_Male] I have ?ToHear ?Character [Attrib_Female] ?ToSay…"

In the next section, we describe the various forms of influence in narrative terms and the implementation of their effects on the characters' plans.

5.   RESULTS

The user advice can directly provide the information sought by one of the characters as part of its plan. For instance, the user can suggest Ross that "Rachel would prefer chocolates", which solves Ross' goal of discovering Rachel's preferences (Figure 3). Here the primary template should identify information about a character's preferences through verb categories and case structure. The secondary template associates various preferences to a given character in terms of semantic categories. From this secondary template it is possible to compare information with currently active tasks, by matching information categories in the HTN to preference categories in the template. Any match will result in labelling the task as solved and instantiating its post-conditions to the value of the information provided by the user.
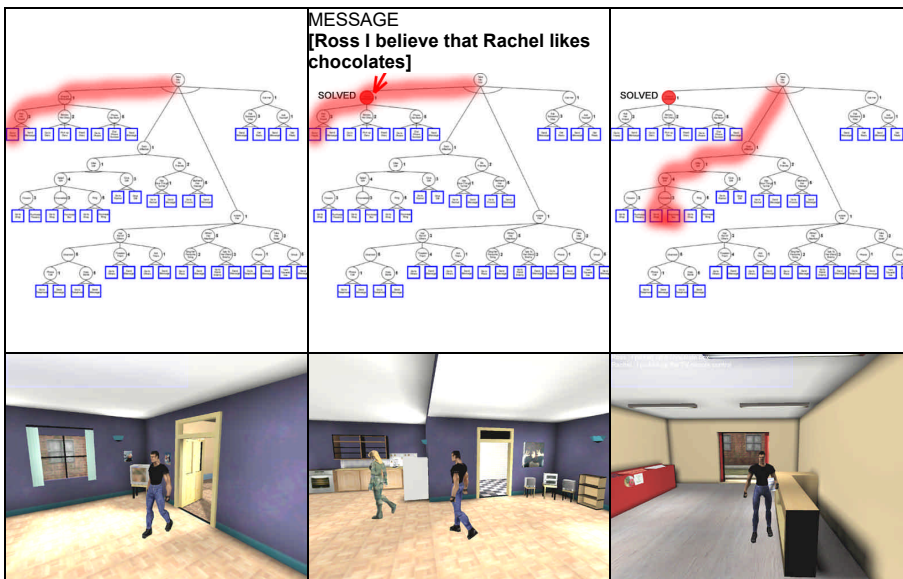

Figure 3. Caption??.

The generic class of *warnings* consists in preventing certain situations to happen in the story. We consider here that warnings aim at avoiding adverse situations only, hence being rather helpful to the character they are directed to. We can further specify warnings into anticipating action failure or avoiding negative outcomes. One example of the former case consists in warning Ross, who intended to steal Rachel's diary that Rachel is currently writing in her diary. This form of advice is a kind of time saver for Ross and avoids the negative impact of action failure (Figure 4). Another possibility consists in creating rules for behaviour that will be checked against dynamic situations. Examples of such advice would be "don't leave Rachel alone" or

"don't let Rachel see you with Phoebe". These will generate small-size, situated plans, triggered by the emergence of the corresponding situation (more precisely detecting specific pre-conditions). For instance, to avoid being seen by Rachel with another girl can be seen as a situated reasoning procedure whose pre-condition consists in any task comprising interaction with a female character. The situated reasoning will check for Rachel presence, aborting the current task if necessary.
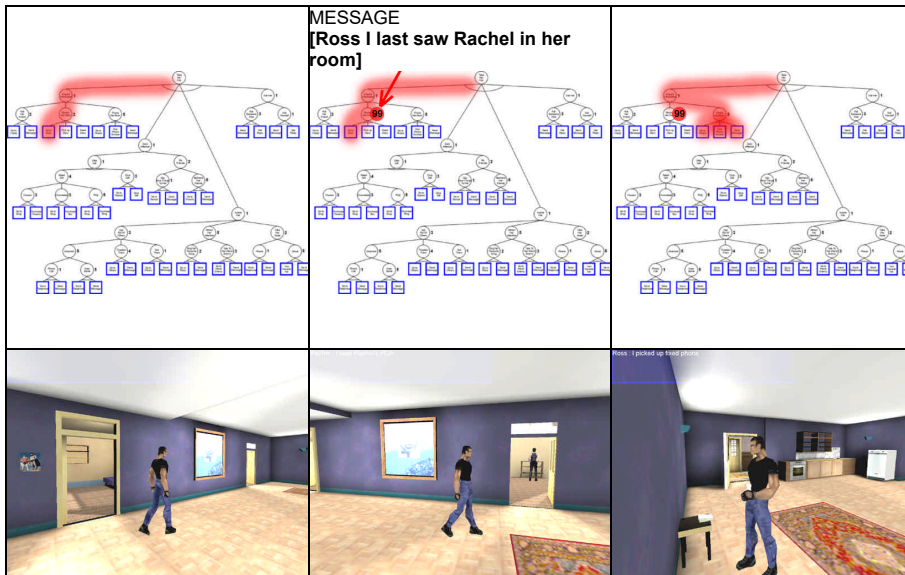

Figure 4. Caption??.

On the other hand, causing situations to occur can be either helpful or negative. For instance, the user can influence Phoebe so that she would talk to Monica rather than Rachel, leaving Ross free to interact with her. Conversely, he could mislead Ross into talking to Phoebe somewhere where Rachel can see them, making her jealous. There's a difference here in influencing feature characters or secondary characters. The latter don't have complex goals and their plans mostly manage their activities. Influencing them essentially shifts their activity focus. For principal characters, the mechanisms can be at least in part shared with other forms of influence (e.g., provision of information), except that the action is carried out for its side effects.

Finally, another mode of influence consists in altering the social relations between characters (through gossip), which will affect any further interactions between them (e.g., "Phoebe, Monica said that your songs are daft" or "I heard Monica say horrible things about you"). This can have a

variable impact on the story: it can for instance cause Ross to avoid seeking information from Phoebe. Or it can result in avoidance behaviour between certain characters, which can play a substantial role in the situations generated during the plot. The result is cause premature termination of all sub-tasks that involve interacting with a disliked character, with associated re-planning. For instance, if Ross dislikes Phoebe he will rather find another source of information about Rachel, even if his current plan initially prescribes getting this information from Phoebe. The mechanism by which this is implemented again consists in activating specific situated reasoning

6.   CONCLUSIONS

We have presented a first prototype of interactive storytelling integrating speech recognition within an influence paradigm. Our approach differs from previous works having used natural language in interactive storytelling, which have been based on dialogue techniques, as in these systems the user is taking part in the story as another actor [Swartout et al., 2001] [Mateas and Stern, 2002]. It also extends previous work in language-based animation [Webber et al., 1995] by explicitly relating plan modifications to speech acts. Further work will pursue our initial hypotheses on the relations between narrative functions and speech acts [Cavazza, 2000].

Acknowledgements

References

Bolter, JD, Grusin, R. *Remediation: Understanding New Media*. Cambridge; Massachussets. MIT Press, 1998.

Cavazza M. Mapping Dialogue Acts to Narrative Functions for Interactive Storytelling. Proceedings of "Learning to Behave"; Twente Workshop in Natural Language Technology n. 17; University of Twente, 2000.

Cavazza M, Charles F, Mead SJ. Characters in Search of an Author: AI-based Virtual Storytelling. Proceedings of the First International Conference on Virtual Storytelling (ICVS 2001); Avignon; France: Lecture Notes in Computer Science; n. 2197; Springer Verlag; pp. 145-154, 2001a.

Cavazza M, Charles F, Mead SJ. AI-based Animation for Interactive Storytelling. Proceedings of IEEE Computer Animation; Seoul; Korea, 2001b.

Erol K, Hendler J, Nau DS, Tsuneto R. A Critical Look at Critics in HTN Planning. Proceedings of IJCAI-95, 1995.

Kambhampati S, Hendler JA. A Validation Structure Based Theory of Plan Modification and Reuse, Artificial Intelligence, 55(2-3):193-258, 1992.

Mateas M. An Oz-Centric Review of Interactive Drama and Believable Agents. In M Wooldridge and M Veloso, (Eds.); AI Today: Recent Trends and Developments, Lecture Notes in AI 1600, Berlin, New York: Springer, 1999.

Mateas M, Stern A. A Behaviour Language for Story-based Believable Agents. In: Papers from the AAAI Spring Symposium on Artificial Intelligence and Interactive Entertainment; AAAI Press, 2002.

Nau DS, Smith SJJ, Erol K. Control Strategies in HTN Planning: Theory versus Practice. Proceedings of AAAI/IAAI-98; pp. 1127-1133, 1998.

Sgouros NM, Papakonstantinou G, Tsanakas P. A Framework for Plot Control in Interactive Story Systems. Proceedings AAAI'96; Portland. AAAI Press, 1996.

Swartout W, Hill R, Gratch J, Johnson WL, 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 of the Autonomous Agents 2001 Conference; Montreal, Canada, 2001.

Webber B, Badler N, Di Eugenio B, Geib C, Levison L, Moore M. Instructions, Intentions and Expectations. Artificial Intelligence Journal; 73; pp. 253-269, 1995.

Weyhrauch P. Guiding Interactive Drama. PhD Thesis; School of Computer Science; Carnegie Mellon University. Technical Report; CMU-CS-97-109, Pittsburgh, PA, 1997.

Young RM. Notes on the Use of Plan Structures in the Creation of Interactive Plot. AAAI Fall Symposium on Narrative Intelligence. AAAI Press, 1999.

Young RM. Creating Interactive Narrative Structures: The Potential for AI Approaches. AAAI Spring Symposium in Artificial Intelligence and Entertainment. AAAI Press, 2000.