The Knowledge Engineering Review, Vol. 23:4, 339–367. © 2008, Cambridge University Press
doi:10.1017/S0269888908000039 Printed in the United Kingdom

On the implications of the virtual storyteller's point of view

JESÚS IBÁÑEZ[1], RUTH AYLETT[2], CARLOS DELGADO-MATA[3] and BEATRIZ MOLINUEVO[4]

[1]*Departamento de Tecnología, Universidad Pompeu Fabra, Barcelona, Spain;*
e-mail: jesus.ibanez@upf.edu;
[2]*School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, UK;*
e-mail: ruth@macs.hw.ac.uk;
[3]*CINAVI, Universidad Bonaterra, Aguascalientes, Mexico;*
e-mail: cdelgado@bonaterra.edu.mx;
[4]*Departamento de Psiquiatria, Universitat Autònoma de Barcelona, Barcelona, Spain;*
e-mail: beatriz.molinuevo@uab.es

Abstract

This paper deals with the implications of showing the storyteller's perspective when telling stories in virtual environments. The paper proposes an open and reusable architecture for the construction of virtual guides who tell stories about the worlds they inhabit from their own perspective. The system consists of two main components: the structure of a knowledge base to code the narrative knowledge about the worlds and a novel hybrid algorithm for the generation of narratives showing the storyteller's viewpoint. The complete architecture has been developed. Furthermore, we carried out a study with users in order to both validate the implemented storytelling system and test whether the users consider it significantly better or worse (in terms of several aspects related to believability) adding the virtual guide perspective (as both text and animations showing emotions) to the guide discourse. The interesting results are discussed.

1 Introduction

It is now possible to create and display *virtual environments* (VEs) for a fraction of the price needed 10 years ago, and as a result their application is increasing. Drawing on the graphics hardware and, sometimes, software, developed for computer games, many recently developed VEs recreate real spaces with an impressive degree of visual realism. Such VEs include large-scale digital cities, some available on the Internet, including interactive 3D representations of their real-world counterparts.

However, we argue that visual realism is by no means enough for a usable VE. For example, real-world heritage sites have long moved beyond the purely visual in order to support visitors with contextual knowledge through signage, hand-held recorded commentaries and human guides. Virtual heritage, though visually compelling, has rarely moved in this direction. As Hoorn *et al.* (2003) have argued, it is now necessary to augment *relevance*, instead of realism, in order to improve the user experience in VEs. The use of *artificial intelligence* (AI) techniques is beginning to be seen by VE designers—again, following a trend in computer games—as a tool to qualitatively improve their models. On the other hand, many groups of AI researchers see VEs as a powerful test bed for their technologies. In this context, a new field called *intelligent virtual environments* (Aylett & Luck, 2000; Aylett & Cavazza, 2001) has begun to be considered. This new area is formed by the increasing overlap of the technologies involved in 3D real-time interactive graphics environments and AI/artificial life technologies. This paper presents research carried out towards the improvement of current VEs from an intelligent agent approach.

In the real world, people relate the environments that surround them to the stories they know about places and objects. When we visit a new city, we learn about it through stories, sometimes told by friends, sometimes by tour guides, sometimes from reading; or through our experience of these places, stored as stories in our own memory. Thus we argue that in order to obtain more relevant VEs, we need to embed a narrative layer within them that can support stories related to the places and objects in these worlds.

However, stories necessarily suppose a particular point of view. As Tozzi (2000) argues in relation to history—a story domain as its name suggests—a striking feature of historical investigation is the coexistence of multiple interpretations of the same event or process. The same historical events can be told as different stories depending on the storyteller's point of view. The story of the same battle between two cities, for example, will be different depending on whether the storyteller identifies with the loser or the winner. Thus we see the need for a virtual guide who tells the stories of a specific VE from her own explicit perspective. Such a guide would also be very useful for educational purposes: storytellers with different profiles could be treated as different sources and used as the basis for developing historical empathy and the ability to understand many points of view in school students. This kind of virtual storyteller would be particularly useful for virtual museums or other virtual heritage worlds. It would also be useful for computer games, where different characters populating the environment could employ the same narrative information to construct different narratives depending on their point of view.

How can one tell whether such a guide is successful? We suggest that it is the *believability* of the guide as a character that is at stake here. Believability was a term introduced by Bates (1994) that is not easy to define. It is however usually taken to refer to characters perceived by the user as compelling and engaging, for which they are willing to suspend their disbelief and which they treat as if they had an independent inner life. Note that this is not at all the same thing as naturalism or graphical realism: cartoon characters may be very believable without being either. It is our hypothesis that by giving a guide the ability to make its own stance on a story clear, it becomes more believable.

This paper describes a novel model for storytelling, which allows a virtual guide to tell stories *from her own perspective*. In our model the guide begins at a particular location and starts to navigate the world telling the user stories related to the places she visits. Our guide tries to emulate a real guide's behaviour in such a situation, behaving spontaneously as a character who knows stories about the places in the virtual world but has not prepared an exhaustive tour or a pre-scripted storyline. Our guide tells stories by improvizing at each step where to go and what to tell next. The proposed model has been implemented and then evaluated through user experiments.

The structure of the paper begins with a review of related work. Following this, we present our proposal by describing the overall architecture, the representation model, the narrative construction and the navigation algorithm. We also outline the implementation in relation to two specific domains. We then present the experimental study. Results are analysed providing the basis for the conclusions.

2 Related work

In this section, we first describe work on computer-based storytelling, discussing general trends in this area. Next, we review research on computer-generated presentations. Finally, we introduce research that particularly inspired our proposal.

2.1 Computer-based storytelling

Computer-based storytelling is a flourishing area. Many research groups are dealing with the problems that this subject provides. Different groups are working on different problems and are approaching them from different perspectives. However, some general approaches can be distinguished. Some groups are trying to obtain methods and mechanisms to automatically generate narratives (this is the plot-based approach), while others are more concentrated on trying to obtain believable characters (Mateas, 1999) (character-based approach). A few works try to reconciliate both approaches (Mateas & Stern, 2000).

Table 1 General approaches

More predefined	More emergent
Plot	Character
Story pieces	Events
Large granules	Small granules
Knowledge based	Behaviour based
Structuralism	Transformationalism

From an information point of view, another distinction can be pointed out. Some researchers work with story fragments, while other researchers work with simple events. People working with story fragments usually approach the problem by one of two ways: connecting these pieces in a hypertext-like network (and then navigating through this network to get a story) or using some narratology theories that establish how story structures should be, in order to combine the story pieces. These works rely on a long tradition of work that suggests that traditional narrative, both oral and written, has well-defined high-level structures. In this sense, the theories of Polti (1922), Propp (1968), Thompson (1955) and Branigan (1992) are especially interesting. The application of Propp's morphological approach to interactive storytelling was suggested by Murray (1998). For a recent example of the application of Propp's theory, see the work of Braun (2002) and Spierling *et al*. (2002). For an example of the application of Branigan's ideas, see the work of Brooks (1999).

The granularity of the information used to construct stories is a key factor. In general, the larger the granules used to build with, the easier it is to build. The smaller the granule, the more precise and smooth the building. Systems with large granules are easier to maintain and their related story generation process is simpler, but the generated stories are more 'predefined'. Systems with small granules are more difficult to maintain and their related story generation process is more complex (at least if the story is expected to show narrative devices as climax, tension, etc.), but the generated stories are more surprising.

As pointed out by Brooks (1999), from an AI point of view, two different approaches have been considered to try to solve the narrative generation problem. The knowledge-based approach makes an *a priori* attempt to capture the rules for successfully solving or navigating a domain, the behaviour-based approach instead relies on a set of lower level competences that are each 'experts' at solving one small part of the larger problem domain. In this sense, the groups trying to apply Propp's ideas to generate a narrative are using a knowledge-based approach, while works giving every character autonomy in order to get an emergent narrative (Aylett, 1999; Louchart & Aylett, 2002) from their interactions are using a behaviour-based approach.

From a more linguistic point of view, as pointed out by Liu & Singh (2002), previous work on story generation has generally taken one of two approaches: structuralist and transformationalist. Structuralists use real-world story structures such as canned story sequences and story grammars to generate stories, while transformationalists believe that storytelling expertise can be encoded by rules, or narrative goals that are applied to story elements such as settings and characters.

In general, as we show in Table 1, more predefined stories are obtained by using a plot-based approach, story pieces, large granules, a knowledge-based approach or structuralism. More surprising and emergent stories are obtained by using a character-based approach, events, small granules, a behaviour-based approach or transformationalism.

2.2 Computer-generated presentations

Apart from the work trying to obtain narratives, there are many approaches oriented towards generating presentations. Although both ways try to create texts in order to communicate information, the expected characteristics of the constructed texts are very different. The works trying to generate presentations construct pieces of textual information describing objects or facts, but do not deal with rhetorical devices (climax, tension, etc.).

In this sense, some works explore how to create presentations adapted to a particular context. ILEX (Intelligent Labelling Explorer) (Oberlander *et al.*, 1998) is a dynamic hypertext system that allows exploring catalogues of objects previously labelled and organized in a structure called the content potential, which is basically a graph containing entity-nodes (which represent objects), fact nodes (which represent facts about objects) and relation-nodes (which represent relations between facts). The system takes into account types of visit (according to the amount of time that is intended to be spent in the museum), the interests of visitors (e.g. in a gallery of 20th century jewellery, some visitors may be more interested in styles, and their evolution, while others may be more interested in the individual designers), and their evolving knowledge during a visit (elimination of redundancy and comparisons of the current object to previously seen objects or type of objects). ILEX uses a system (O'Donnell *et al.*, 2000) to generate text from relational databases that need to be enhanced with additional information defining the semantics of the database, as a database itself says nothing about the nature of their contents. For a survey and comparison of existing dynamic hypertext systems (like ILEX), see the work of Knott *et al.* (1996).

Presentations (as well as stories in general) are multi-modal. That is, they can be composed not only by textual sentences but also by images, sound, videos, 3D scenes, etc. Some work is intended not to generate textual presentations but to propose additional components to preexistent textual presentations. For example, ARIA (Annotation and Retrieval Integration Agent) (Lieberman *et al.*, 2001) is a software agent that acts as an assistant to a user writing e-mail or Web pages. As the user types a story, it does continuous retrieval and ranking on a photo database. It can use descriptions in the story to semi-automatically annotate pictures. To improve the associations beyond simple keyword matching, the system uses natural language parsing techniques to extract important roles played by text such as 'who, what, where, when'. A detailed description of the techniques employed can be found in the work of Lieberman & Liu (2002) and Liu & Lieberman (2002).

2.3 Inspiring our work

Some systems are not intended to generate either complex stories or simple presentations, but little stories in some point between both extremes. In this sense, Terminal Time (Domike *et al.*, 2002) is a system that combines historical events, ideological rhetoric, familiar forms of TV documentary, consumer polls and AI algorithms to create hybrid cinematic experiences for mass audiences that are different every time. Through an audience response measuring device connected to a computer, viewing audiences respond to periodic questions reminiscent of marketing polls. Their answers to these questions allow the computer program to create historical narratives that attempt to mirror and often exaggerate their biases and desires. The engine uses the past 1000 years of world history as 'fuel' for creating these custom-made historical documentaries.

Our proposal, as Terminal Time, tries to generate narratives from a particular perspective or ideology. However, while Terminal Time takes into account an ideology biased by the users (the audience), we take into account the virtual guide profile. And while Terminal Time selects the story elements to tell according to its ability to show a plan predefined to support its ideology, we select the story elements according to several factors (as described in Section 5).

Our proposal deals, in particular, with the problem of guiding in a VE. This problem, which has its own special characteristics, has been explored in Kyoto Tour Guide (Isbister & Doyle, 1999), a project aimed to develop an agent that is integrated in an on-line 3D website tour, a digital version of Kyoto. Kyoto Tour Guide explores ways to make an agent's narrative effectively adaptive to different groups who take the tour, like an expert human tour guide. The authors focused on developing storytelling strategies that produce an engaging experience for tour takers. They derived a list of target abilities for the agent by researching the behaviour of actual tour guides in Kyoto. They found that tour guides made use of illustrative stories frequently, supplementing the rich visual environment of the city with explanations of how Japanese people, both past and present, made use of these settings. Stories included things such as descriptions of how a given site was constructed and its history of destruction and reconstruction, descriptions of peak historic

events and customary activities that occur at the site. While visitors looked at the buildings, and took things in visually, the guide would create a narrative context for the site with these stories, providing visitors with stories that they could share with their fellow tour members, as well as with people back home.

In this sense, our proposal deals with the same problem as Kyoto Tour Guide does. However, while Kyoto Tour Guide uses a database of stories explicitly related to the sites, our system constructs stories, step-by-step, from story elements (encoding historical facts) by means of processes that take into account several factors.

A virtual guide is expected to tell short stories about the places she visits. Therefore, our system is expected to generate this kind of short stories. MAKEBELIEVE (Liu & Singh, 2002) is a story generation agent that can generate short fictional text of 5–20 lines when the user supplies the first line of the story. MAKEBELIEVE uses a subset of sentences extracted from Open Mind Commonsense (OMCS) (Singh, 2002) that describes simple commonsense causations. By inferring over these causations, the system is able to generate a text like this: 'John became very lazy at work. John lost his job. John decided to get drunk. He started to commit crimes. John went to prison. He experienced bruises. John cried. He looked at himself differently'. The authors reported that although the system does not incorporate plot devices such as motifs, climax, tension, etc., many users involved in the evaluation nonetheless felt that these devices were present in the generated stories.

As shown in Section 5, we use this approach in our system. However, while MAKEBELIEVE generates complete short stories just by inferring on commonsense rules, our system infers on commonsense rules just to enhance story elements previously selected and translated to show the guide virtual perspective. Note that MAKEBELIEVE can generate the complete story by inferring on commonsense rules because the stories it generates are fictional, while we cannot do it because we construct stories based on historical facts.

3 Global architecture

As shown in Figure 1, the VE runs on a VE server and the guide is connected to this server through the Internet via sockets. The knowledge the system uses includes a common level of meta-information, global application-dependent meta-information specially suitable for narrative construction, and guide-dependent information that defines features specifying the guide's characteristics, and therefore allow generating narratives from the guide's perspective.

At a conceptual level, the main processes are the *broker* (controls the operation of the other modules), the *communications manager* (deals with the communication with the VE server), the *KB manager* (controls the knowledge base operations), the *narrative constructor* (constructs the narrative and performs the play) and the *user interface* (allows configuring the guide).

The storytelling process involves several mechanisms, which can be summarized as follows. With every step, the virtual guide decides (improvises) where to go and what to tell there, through a process that selects a location and a story element (this improvization tries to emulate what occurs in the mind of a real guide by taking into account the spatial structure of the world, her memories, and her own interests). Then the story element is extended and translated from the virtual guide perspective and enhanced with information generated through simple commonsense rules. At the same time, a related guide attitude is induced. Next, the new story elements are used to generate text sentences to tell and special effects to show, while the guide attitude is used to select suitable guide actions. Finally, the storytelling process is carried out, that is, the guide is moved to the selected location, the environment is updated with the required special effects and the textual sentences are told while the guide shows her attitude through the selected actions.

4 Representation model

Ibáñez (2004) proposed a representation model. This model provides a layer of semantic information that constitutes a common lower level for application-dependent models. In this sense, the

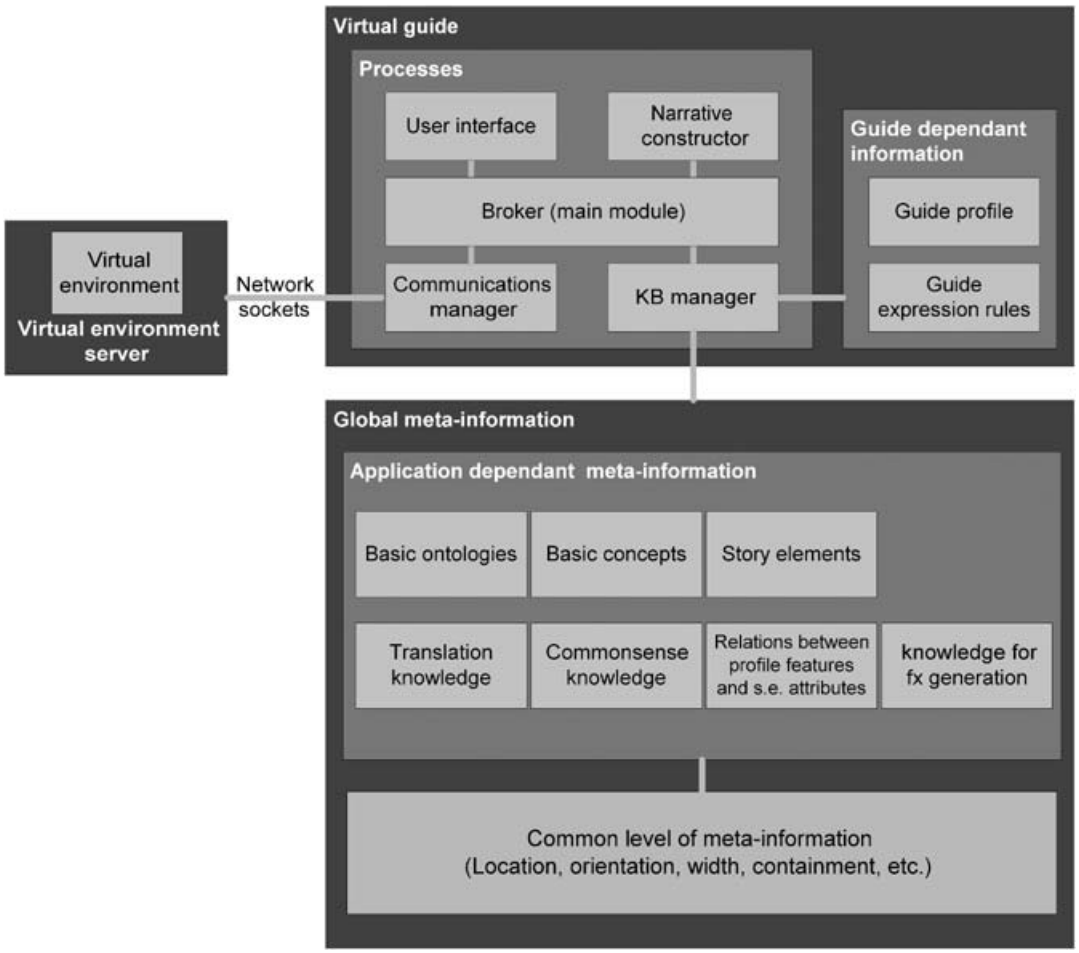

Figure 1 Global architecture of the storyteller system

storytelling system extends the proposed model by adding new kinds of elements that are useful for this concrete application. In this section, we describe these elements, which include basic ontologies, basic concepts, story elements, guide-dependent information and knowledge encoded as rule sets.

4.1 Basic ontologies

The system defines two basic ontologies: a *guide profile ontology*, which allows describing guide profiles (in particular our guide profile ontology defines roles and interests), and an *ontology of story element attributes*, which describes the attributes that can be used to annotate the story elements. Different versions of these ontologies can be defined for environments. For example, next we show a part of an attribute ontology built for the case of a virtual recreation of the Mayan city of Palenque:

```
society
   alimentation
   clothes
   ...
science
   astronomy
   mathematics
   ...
```

which in fact is not appropriate to annotate story elements for telling stories in a virtual *dungeon & dragons* environment, although it could be reused to annotate story elements for a virtual antique Roman city.

4.2 Basic concepts

Some *basic concepts* (objects, characters or abstract concepts) are defined in terms of a textual description. These definitions are used to introduce the related concepts the first time they appear in a particular storytelling process. These concepts also serve as clues to calculate the similarity between every pair of story elements, and therefore to calculate how much a story element is triggered off by the already-told story elements, in a remembrance-like process (see Section 5.1).

4.3 Story elements

Story elements are pieces of narrative knowledge expressed in a special format. They are the basic pieces the virtual guide uses to construct stories. These elements are, in principle, ideologically agnostic, that is, they do not reflect any particular ideology. Story elements are annotated with the features shown in Table 2.

4.4 Guide-dependent information

4.4.1 Guide profile

The guide profile is composed of roles and interests. Figure 2 shows an example of a guide profile.

Table 2 Features of the story elements

Feature	Description
Name	Every story element in a narrative space is uniquely identified by its name.
Type of event	Each story element belongs to a particular kind of event.
Location	Each story element is related to one or more locations. These locations indicate the places where the story element can be told. We consider the following kinds of location: *object* (the story element location is specified by an object name indicating that the story element is related to the location of that object), *space* (the story element location is specified by a space name indicating that the story element is related to the location of that space), *node* (the story element location is specified by a navigational node name indicating that the story element is related to the location of that navigational node), *vector* (the story element location is specified by a 3D vector) and *any* (the story element can be told anywhere).
Attributes	Each story element is tagged according to several attributes. These attributes describe the nature of the story element (for example artistic, religious, political, social).
Basic concepts	Basic elements are specially tagged in the story elements so that the system knows that there is a definition for this concept in the knowledge base. The first time that the narrative constructor uses a particular basic concept, it will retrieve and use its related definition (description of the concept) in order to inform the user. The basic concepts are also used to simulate the guide's memories.
Date	Each story element is related to a particular historical moment or period.
FXs	Special environment conditions such as weather, light, etc. are annotated, indicating that these conditions must be generated when telling this story element.
Granularity	The *size* of the story element. It allows the narrative constructor to construct storyboards with a suitable size.
Effects	A list of events (other story elements) that are supposed to be caused by this story element.
Subjects	A list of subjects of the fact encoded by the story element.
Objects	A list of objects of the fact encoded by the story element.
Text	A string of text expressing the fact encoded by the story element.

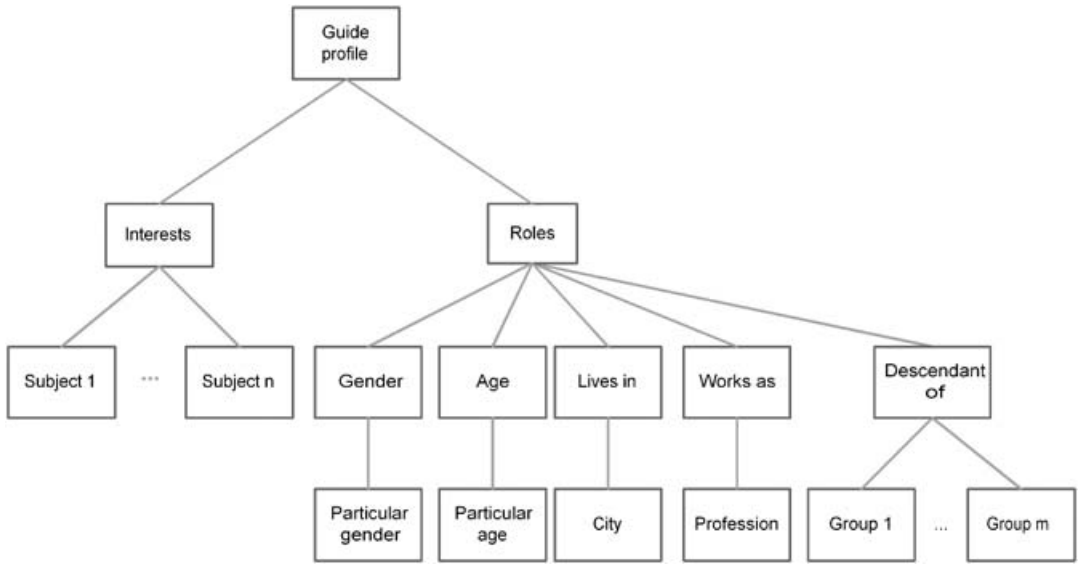

Figure 2 A sample guide profile

4.4.2 Guide expression rules

These are rules that allow translating guide attitudes into guide actions (particularly 3D animations), which allow the guide to express herself. By decoupling the attitudes from the actions, the system is able to express the guide attitudes (generated by a process common to all the guides) through particular actions that exactly reflect the particular guide personality.

4.5 Knowledge encoded as rule sets

4.5.1 Translation knowledge

This knowledge is represented by a set of rules that allow translating story elements to show a particular guide perspective. For example, a story element describing a battle between Spaniards and Mexicans where Spaniards lose will be translated to a new story element showing the guide profile implication (if the guide is Spanish the new story element will reflect a loser perspective, while if the guide is Mexican the new story element will reflect a winner perspective).

4.5.2 Commonsense knowledge

These are rules representing simple commonsense causations. These rules are used to enhance the story elements previously selected to be told once these story elements have been translated from the guide perspective. Therefore, this knowledge allows the guide to extend the basic information represented as story elements by using her 'supposed' commonsense.

4.5.3 Relations between guide profile features and story element attributes

These are relations that are established between interests and roles on the one hand, and attributes on the other hand. These relations describe the importance that a particular story element attribute could present for a guide with a particular interest or role.

4.5.4 Knowledge for special effects generation

These are rules that allow generating the special effects corresponding to a particular story element. They describe how special environment conditions should be performed in the VE.

5 Narrative construction

In our model the guide begins at a particular location and starts to navigate the world telling the user stories related to the places she visits. Our guide tries to emulate a real guide's behaviour in

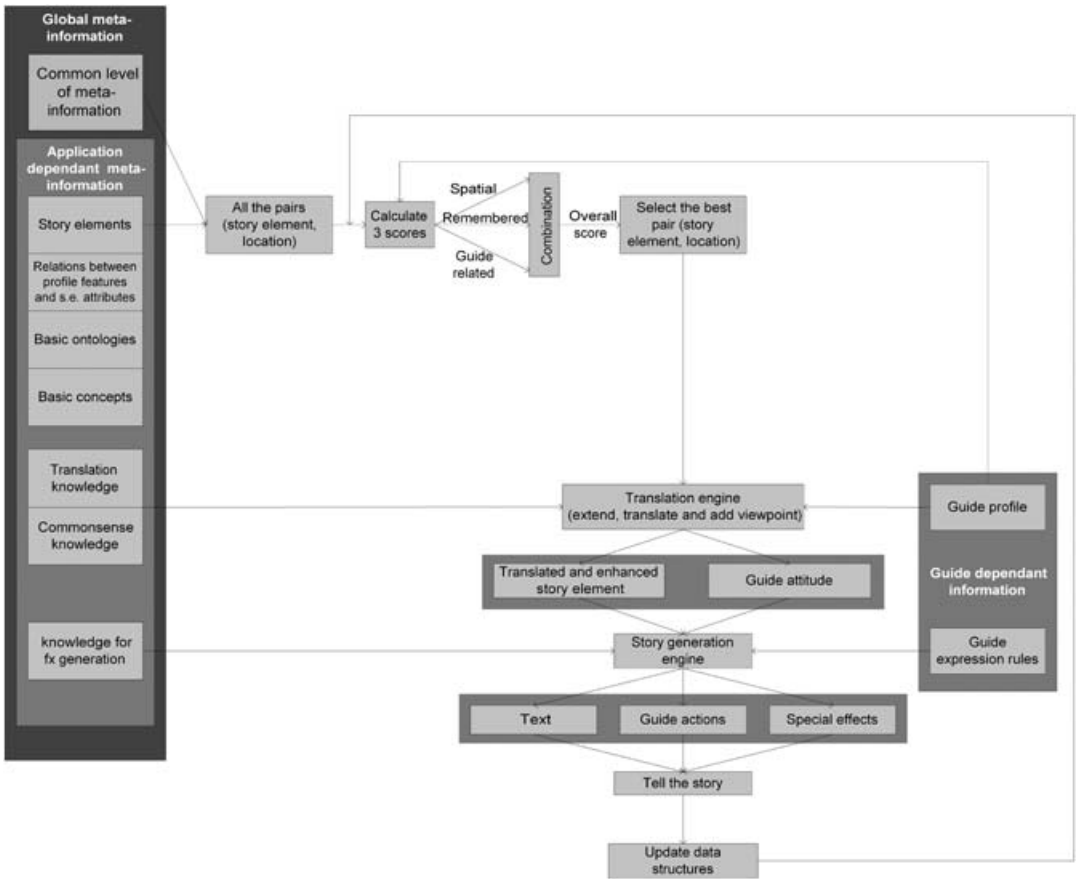

Figure 3 Algorithm diagram

such a situation. In particular, she behaves as a spontaneous real guide who knows stories about the places in the virtual world but has prepared neither an exhaustive tour nor a storyline. Therefore, our guide tells stories by improvizing, at each step, where to go and what to tell next, and this improvization tries to emulate what occurs in the mind of a real guide.

Furthermore, our guide tells stories from her own perspective, that is, she narrates historical facts and events taking into account her own interests and roles. In fact, she extends the stories she tells with comments that show her own point of view. This mixture of neutral information and personal comments is what we can expect from a real guide who, on the one hand, has to tell the information he has learnt (contained in the script provided by his company), but on the other hand, cannot hide his feelings, opinions, etc. about the information he is telling.

We have designed a hybrid algorithm that models virtual guide behaviour taking into account all the aspects described above. Figure 3 shows a diagram of the overall functioning. The processes and mechanisms involved in the algorithm can be separated in three global parts. These three global processes are carried out with every step. First, the guide should decide where to go and what to tell there (actually, she only decides something to tell but not all the information that she will tell, that is she decides that she wishes to tell something). Then, during the next process, she extends this information with additional story elements, translates these story elements from her own perspective, and again extends the selected information depending on her perspective, that is, with information showing her viewpoint. Finally, the system generates a story. The next three subsections detail these general phases.

5.1 Finding a spot in the guide's memory

Given a particular step in the navigation-storytelling process (i.e. the virtual guide is at a particular location and she has previously narrated a series of story pieces), the guide should decide where to

go and what to tell there. To emulate a real guide's behaviour, the virtual guide evaluates every candidate pair (*story element, location*) taking into account three different factors: the distance from the current location to *location*, the already-told story elements at the current moment and the affinity between *story element* and the guide's profile.

A real guide will usually prefer nearer locations, as distant locations involve long displacements, which lead to unnatural and boring delays among the narrated story elements. In this sense, our guide prefers nearer locations too, and therefore shorter displacements.

When a real guide is telling stories in an improvizational way, the already-narrated story elements make him recall by association related story elements, that is, some concepts, objects or characters used in previously narrated story elements trigger off recollections in his mind. In a spontaneous way, a real guide tends to tell these recently remembered stories. In this sense, our guide prefers story elements related (metaphorically remembered) to the ones previously narrated.

Finally, a real guide tends to tell stories related to his own interests (hobbies, preferences, etc.) or roles (gender, job, religion, etc.). In fact, this could be a special case of the above-commented remembrance by association, as the guide's interests and roles could be seen as persistent triggers of recollections in the guide's mind. In this sense, our guide prefers story elements related to her own profile.

The system evaluates every candidate pair (*story element*, *location*) such that there is an entry in the knowledge base that relates *story element* to *location* (note that this means that *story element* can be narrated in *location*) and such that *story element* has not been narrated yet. In particular, three scores corresponding to the previously commented factors are calculated. These three scores are then combined to calculate an overall score for every candidate pair. The user can tune the weight of each score in the combination process, as shown in Figure 7, which displays the user interface of the current implementation. Finally, the system chooses the pair with the highest overall score value. Next, we show the steps followed to find the spot:

//preliminary calculations are carried out
For each $pair = (se, loc) \in PAIRS$
  $distance\ (pair) = navigationalDistance\ (loc, currentLocation)$
  For each concept $c \in concepts\ (se)$
    $rememberedScore(pair) = rememberedScore(pair) + recentMemory(c)$
For each attribute $a \in attributes\ (se)$
    $guideRelatedScore(pair) = guideRelatedScore(pair)$
    $+(relevance(a, guideProfile) \cdot attributeValue(se, a))$

$$maxDistance = max_i\{distance(pair_i) \mid pair_i \in PAIRS\}$$

//the scores are normalized

For each $pair = (se, loc) \in PAIRS$

$$spatialScore(pair) = \frac{maxDistance - distance(pair)}{maxDistance}$$

$$rememberedScore(pair) = \frac{rememberedScore(pair)}{max_i\{rememberedScore(pair_i) \mid pair_i \in PAIRS\}}$$

$$guideRelatedScore(pair) = \frac{guideRelatedScore(pair)}{max_i\{guideRelatedScore(pair_i) \mid pair_i \in PAIRS\}}$$

//the scores are combined

$$overallScore(pair) = spatialScoreWeight \cdot spatialScore(pair)$$
$$+rememberedScoreWeight \cdot rememberedScore(pair)$$
$$+guideRelatedScoreWeight \cdot guideRelatedScore(pair)$$

//the spot is selected

$$spot = pair \in PAIRS \mid overallScore(pair) \geq overallScore(pair') \forall pair' \in PAIRS$$

where *concepts* (*se*) gives the set of concepts annotated in the story element *se*. *recentMemory* (*c*) gives the value stored in recentMemory for the concept *c*. *relevance*(*a*, *prof*) gives the relevance of the attribute *a* for the profile *prof*. *Attributes*(*se*) gives the set of attributes associated with the story element *se*. *attributesValue* (*se*, *a*) gives the value associated with the attribute *a* in the story element *se*. *currentLocation*. is the location where the guide is located at the moment. *distance*(*pair*) gives the distance from the location of the pair *pair* to *currentLocation*. *NavigationalDistance*(*loc*,*loc*′) gives the distance from location *loc* to location *loc*′ in terms of navigation. That is, it is not the Euclidian distance but the distance that should be navigated to reach *loc*′ from *loc*. *PAIRS* is the set of pairs (*story element*, *location*) such that there is an entry in the knowledge base that relates *story element* to *location* and such that *story element* has not been narrated yet. Furthermore, pairs (*story element*, *location*) where *location* is of type 'any' are excluded. *guideProfile* is the profile of the virtual guide. *spatialScoreWeight*, *rememberedScore-Weight* and *guideRelatedScoreWeight* are the weights of the spatial score, remembered score and guide-related score, respectively. These weights parameterize the combination of the three scores.

5.2 Extending and contextualizing the information

Figure 4a represents a part of the general memory the guide uses. This memory contains story elements that are interconnected with one another in terms of different relations. In particular, in our case, cause–effect and subject–object relations interconnect the story elements. Two story elements X and Y are connected by a cause–effect link (see Figure 5) if one of the following cases is fulfilled: X is cause of Y; X is effect of Y; X and Y are causes of the same third story element Z; X and Y are effects of the same third story element Z. A subject–object link between the story elements X and Y is established if any of the following cases is fulfilled: the subjects of X and Y are the same; the objects of X and Y are the same; the subject of X is the object of Y; the object of X is the subject of Y.

Figure 4b shows the same part of the memory where a story element has been selected by obtaining the best overall score described in the previous section.

If the granularity provided by the selected story element is not considered to be large enough to generate a little story, then more story elements are selected. The additional story elements are chosen according to a particular criterion or a combination of several criteria (cause–effect and subject–object in our case). This process can be considered as navigating the memory from the original story element. Figure 4c shows the same part of the memory where three additional story elements have been selected by navigating from the original story element. The system allows the user to tune the relative weight of each criterion in this memory navigation process as well as the desired degree of granularity (see Section 7.1 and Figure 7).

Once the granularity provided by the selected story elements is considered to be large enough, the selected story elements are translated, if possible, from the virtual guide perspective (see Figure 4d). For this task, the system takes into account the guide profile and the rules stored in the knowledge base that are intended to situate the guide perspective. The translation process also generates guide attitudes that reflect the emotional impact that these story elements cause her. We demonstrate this by a simple example. Let us assume the following information extracted from a selected story element:

```
fact(colonization, spanish, aztec)
```

meaning that the Spanish people colonized the Aztec. And let us assume the following meta-rules included in the knowledge base, aimed to situate the guide perspective

```
fact(colonization, Colonizer, Colonized) and profile(Colonizer) =>
   fact(colonizerColonization, Colonizer, Colonized) and
   guideattitude(contradictory)

fact(colonization, Colonizer, Colonized) and profile(Colonized) =>
   fact(colonizedColonization, Colonizer, Colonized) and
   guideattitude(anger)
```

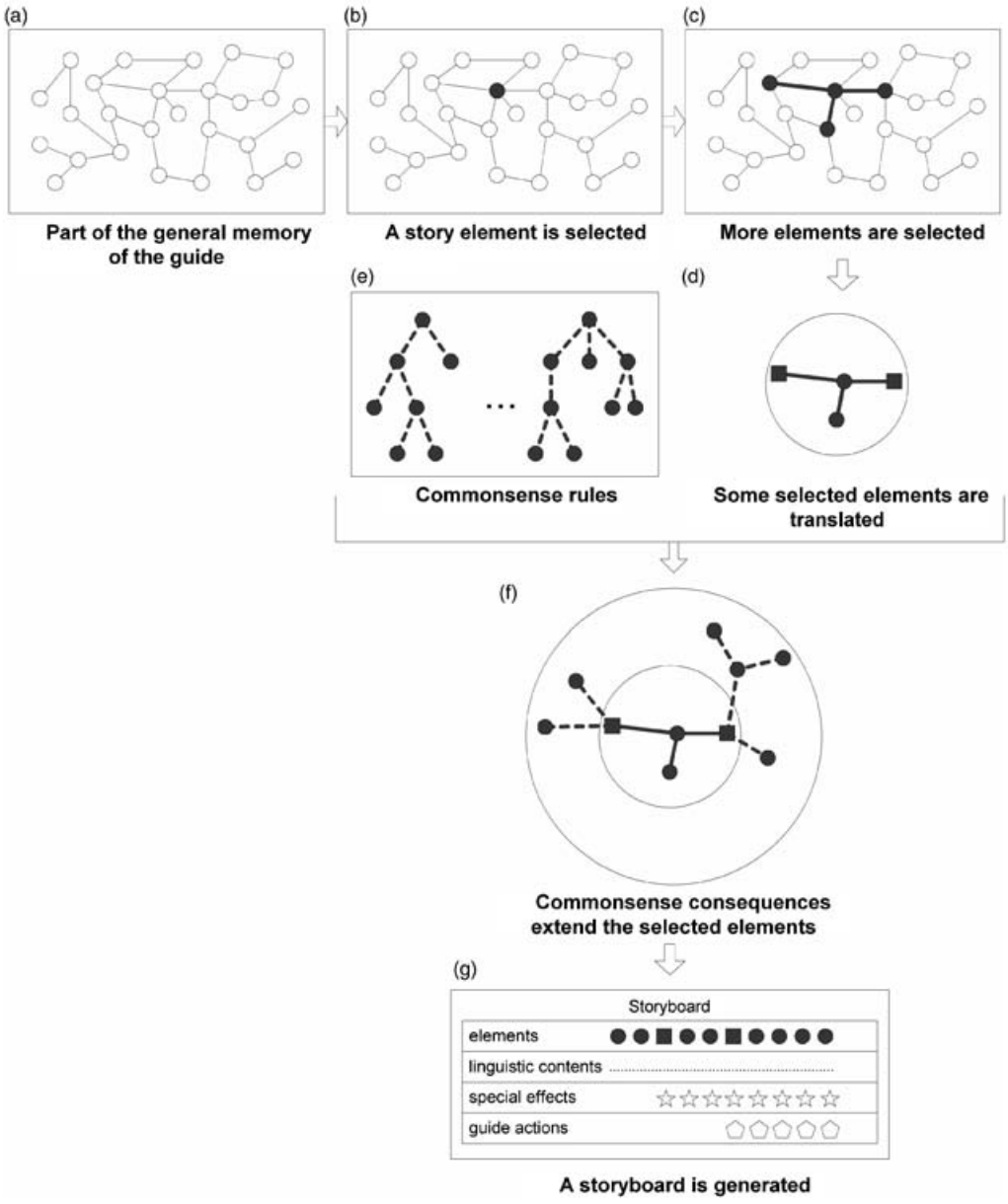

Figure 4 Storyboard construction

meaning that a colonizerColonization fact and a contradictory guide's attitude should be inferred if a colonization fact is included in the story element and the guide profile matches the second argument of this fact, that is, the guide is the Colonizer. In the same way, a colonizedColonization fact and anger as the guide's attitude should be inferred if a colonization fact is included in the story element and the guide profile matches the third argument of this fact, that is, the guide is the Colonized. Whatever rule is activated, the new inferred fact represents the original one but from the guide's perspective.

In addition, the new translated story elements are enhanced by means of new information items generated by inferring simple commonsense rules, allowing the guide to add some comments showing her perspective. The guide uses the new contextualized story elements (Figure 4d) as input for the rules that codify commonsense (Figure 4e). By inferring these rules, the guide obtains consequences that are added to the contextualized story elements, obtaining a new data structure

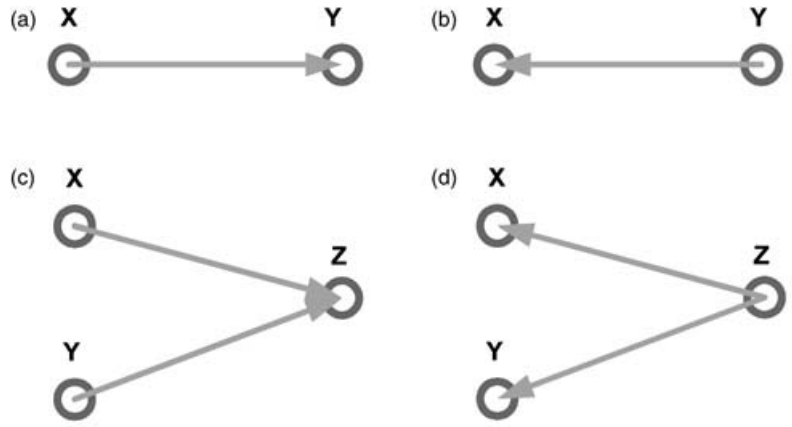

Figure 5 Possible cause–effect relations between two story elements X and Y: (a) X is cause of Y, (b) X is effect of Y, (c) X and Y are causes of the same third story element Z, and (d) X and Y are effects of the same third story element Z

(Figure 4f) that codifies the information that should be told. We continue with the previous example. Let us assume the following commonsense rules:

```
fact(colonizerColonization, Colonizer, Colonized)=>
   fact(culturalImprovement, Colonized) and
   fact(religiousSalvation, Colonized)

fact(colonizedColonization, Colonizer, Colonized)=>
   fact(culturalDestruction, Colonized) and
   fact(religionChange, Colonized)
```

meaning that the colonizer's view of a colonization process implies the cultural improvement and religious salvation of the colonized. The view of the colonized implies the destruction of their culture and the change of their religion. Therefore, if in our example the guide were Aztec, the story element to be told would be enhanced with the facts culturalDestruction and religionChange. Note that in turn these facts could fire other commonsense rules producing additional facts.

5.3 Generating the story

As a result of the previous processes, the guide obtains a set of inter-related information items to tell (Figure 4f). These elements are stored as a structure that reflects the relations among them, as well as the reasons why each one was selected. Some elements are also related to particular guide attitudes. Now, the system generates the text to tell (expressing these elements) as well as special effects and the guide actions to show while telling the story. The phases of this story generation process are as follows:

1. As the information should be told in a sequential way, the first step is then to order the data elements. Several criteria can be considered. In particular, we use the four following concrete criteria (its weights can be tuned by the user):
   - Selection order: The elements are arranged according to the selection order. That is, if an element *Y* was selected after another element *X* (whatever the criterion employed for this selection) then *Y* is located after *X*. In addition, commonsense consequences are immediately located after the elements which generated them. If several commonsense consequences were triggered by the same element then these consequences are also arranged according to its selection order.
   - Cause–effect: The elements are ordered in such a way that if an element Y was caused by another element X, then X should precede Y.

- Subject–object: The elements are ordered in such a way that the elements whose subject/object are similar are grouped together.
- Classic climax: The first selected story element (the one that obtained the best overall score) is supposed to be the climax of the narration, and therefore all the rest of the elements are arranged taking it into account.

2. The text corresponding to the ordered set of elements is generated. The complexity of this process depends on the particular generation mechanism and the degree of granularity employed. The granularity we are employing at the moment is one sentence per story element and the generation mechanism we have designed can be seen as a Natural Language Generation (NLG) system, which is simple in some aspects but complex in others. NLG consists of three tasks: *discourse planning*, *sentence planning* and *text realization*. Next, we comment how these tasks are carried out by our system. The *discourse planning* task is relatively complex as we select the story elements to tell taking into account several factors such as the spatial structure of the world, the guide profile and the guide memories. Furthermore, the selected elements can be ordered according to several criteria. Moreover, the system not only plans a textual discourse but also guides actions and special effects to perform together with the discourse. In addition, the system adds descriptions of basic concepts annotated in the story elements the first time they appear. The *sentence planning* task involves substituting the subjects of sentences for the appropriate pronouns if convenient. The *text realization* is based on templates. Particular elements (e.g. the subject) of the sentences are annotated with tags. These tags allow the system to carry out transformations of the sentences (e.g. they substitute the subject for the appropriate pronoun). In addition, we have started to employ the Rhetorical Structure Theory (RST). So far we have utilized the RST cause–effect relation and we plan to use more RST relations in the near future.
3. A process that relies on the guide expression rules (the set of rules that translate abstract guide attitudes in particular guide actions) generates a set of guide actions (each one related to a particular story element). By this process, for example, an Aztec guide will show an anger gesture when talking about the Spanish colonization of the Aztec territories, while a Spanish guide will show a posture of distress when talking about the Aztec religious sacrifices.
4. Special effects are prepared to show the particular environment conditions annotated in concrete story elements. This work is done by employing the set of rules that describe how special environment conditions should be performed in the VE.

Thus, finally, a storyboard like the one shown in Figure 4g is obtained. Then the guide navigates towards the new target location (the location of the *spot*). Once the guide reaches it, she performs the story. In addition, the data structures employed in the processes are updated. Particularly interesting is the updating of the guide's recent memory. In previous algorithm, we named the value stored in the guide's recent memory for the concept c, $recentMemory(c)$. This $recentMemory(c)$ increases for the concepts c such that they appear in the just-created storyboard, that is the guide reinforces those recollections. On the contrary, $recentMemory(c)$ decreases for the concepts c such that they do not appear in the just-created storyboard. Metaphorically, the guide forgets those concepts a bit.

6 Navigation algorithm

According to Jul & Furnas (1997), the process of navigating throughout an environment can be considered to be composed of two tasks and two tactics. The tasks are searching (looking for a known target) and browsing (seeing what is available in the world). The tactics are querying (submitting a description of the object being sought to a search engine, which will return relevant content or information) and navigation (moving oneself sequentially around an environment, deciding at each step where to go next based on the task and the parts of the environment seen

so far). A combination of a task with a tactic produces a navigation process. Our guide is supposed to know the environment she inhabits. All the information she needs to navigate the virtual world (as a metaphor of her memories) is stored in the knowledge base and accessible to her. Therefore, in our case, the task is searching and the tactic is querying, that is, our virtual guide will search a space by asking the knowledge base a question about the location of a place or an object.

The algorithm selects the more suitable navigational node for the target location. Then the guide navigates towards that node (by successively navigating to the nodes composing the minimal path from the current location to the target node). Once the guide reaches the target node, she gets nearer the target location if the node is farther away from the location than a particular threshold. Finally, the appropriate action is carried out. If the agent is telling stories, the guide will start to perform the storyboard prepared for the reached location.

7 Implementation

As pointed out by Laird (2002), although the development of realistic VEs is an expensive and time-consuming enterprise requiring expertise in many areas, computer games provide us with a source of cheap, reliable and flexible technology for developing our own VE for research. During the last few years, AI researchers have been using several computer game engines (Descent 3, Quake II, etc.) to check their algorithms.

A current trend is to use the Unreal Tournament (UT) engine as the development platform. UT is a cheap game that provides access to level editors to define the environment, a scripting language (UnrealScript) to define the physics of the world and the way objects interact, and the ability to import your own objects into the game. UT is currently being used by several groups to develop storytelling systems (Cavazza *et al.*, 2001; Young, 2001; Robertson & Oberlander, 2002), and AI-based bots in general (Laird, 2002).

Gamebots (Adobbati *et al.*, 2001) is a modification to UT that allows characters in the game to be controlled via network sockets connected to other programs. For a detailed description of Gamebots, see Kaminka *et al.* (2002).

We chose UT as the platform on which our virtual worlds run. In fact, the system was intended to provide a virtual storyteller for virtual models of the Mayan Cities of Palenque and Calakmul, which are being developed as UT worlds. As we wished our system to be open and portable, we decided to use Gamebots to connect our virtual guide to UT. In addition, we chose a second VEs platform to check that our guide is easily portable. In particular, we selected DeepMatrix (Reitmayr *et al.*, 1999), as we had already developed an agent to connect to DeepMatrix servers.

We developed the proposed system. Figure 6 shows a diagram of the current implementation. The core of the virtual guide is a Java application that is able to connect to both UT worlds (through Gamebots) and VRML worlds (running on a DeepMatrix Server). This Java application controls the movement and animations of the guide in the world as well as the presentation of texts, which show the generated narratives.

The current version uses a MySQL (2004) database to store the knowledge base containing the information about the environment as well as the narrative information. The Java application accesses these data through JDBC. The developed system uses Jess to carry out inferences on the information, that is, to reasoning. Jess (2004) is a rule engine and scripting environment written entirely in Java language, and originally inspired by the CLIPS expert system shell.

In addition, we have developed a library that allows exporting the narrative knowledge base as a graph where the story elements are nodes connected with one another in terms of cause–effect and subject–object relations. The library generates the graph in the Pajek format. Thus, the narrative knowledge bases can be analysed and visually represented by using Pajek (Batagelj & Mrvar, 2003). In fact, Figures 10 and 11 have been generated in this way.

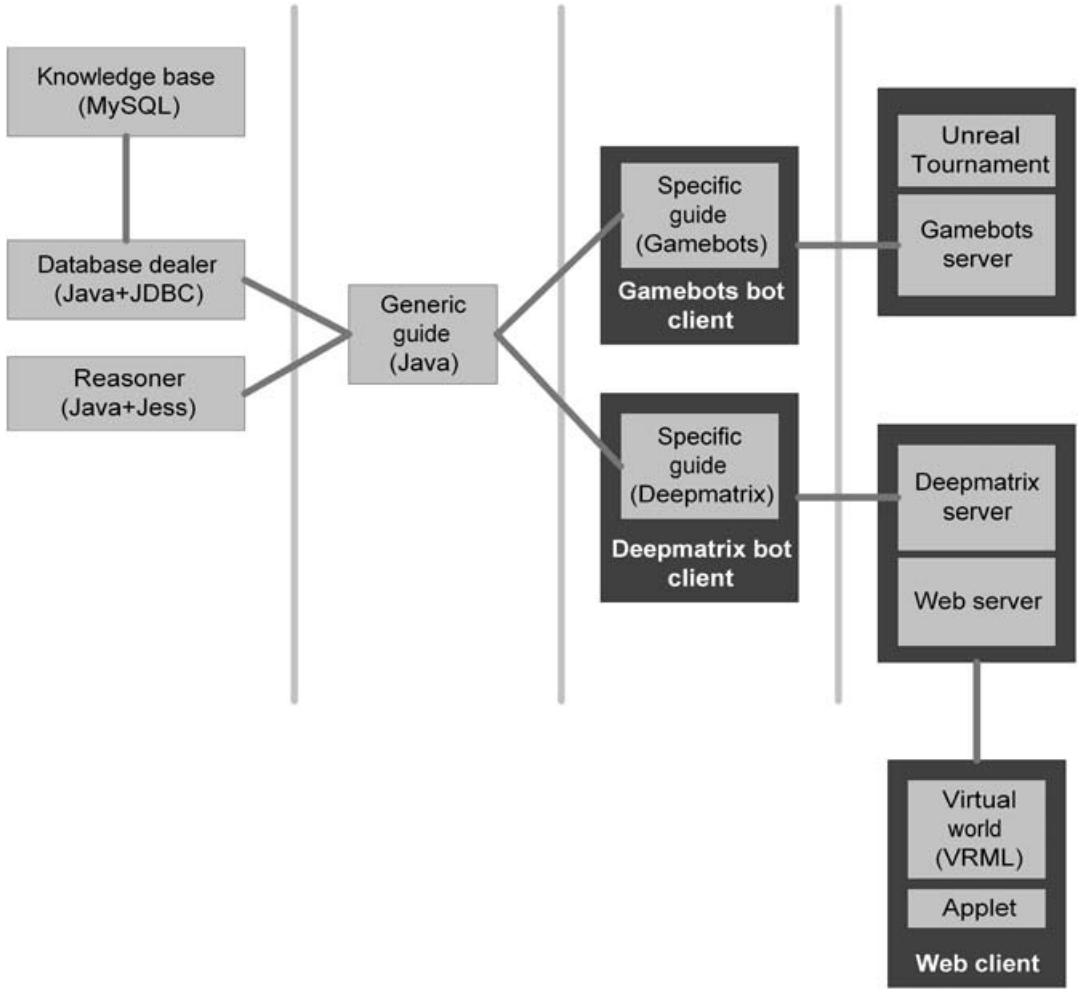

Figure 6 Implementation diagram of the storytelling system

7.1 Interface

We developed a user interface that allows the user to control and trace the storytelling processes. Figure 7 shows the interface. Next, we list and briefly describe the different components of the interface (note that they are labelled as in Figure 7):

1. **A:** Space, which shows the evolution of the data structures employed by the processes and algorithms. It allows the user to trace the performance of the virtual guide.
2. **B:** Sliders, which allow the user to parameterize the narrative construction by varying the value of several features.
3. **C:** Space, where the discourse generated by the guide is displayed. Note that in addition, the text generation process is shown by inserting comments indicating the special actions (for instance, a pronoun substitution) it carries out.
4. **D:** Widgets, which allow the user to select between two modes of execution of the system: *manual* (which allows the user to control when the guide should initiate every step) and *automatic* (the user does not control).

8 Example 1: Family home

This section shows an example of application of our system. The simple scenario is a home (see Figure 8) inhabited by part of a family: father, mother, son and daughter. In addition, another son

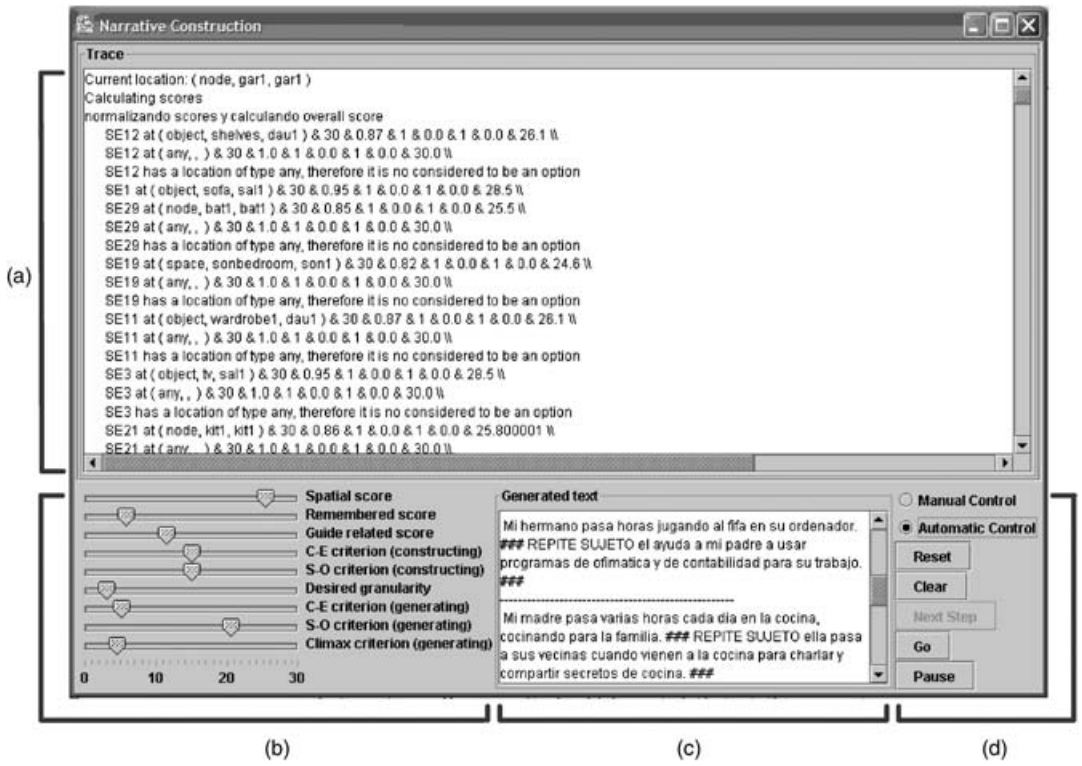

Figure 7 Interface to control and trace the storytelling processes

Figure 8 Snapshot of the scenario with the virtual guide

and another daughter live on their own. You visit the house invited by one of the latter children. The house is populated by objects that are associated with memories related to the members of the family. In principle, all of the children are supposed to have almost the same shared memories. However, each child has a well-defined profile, different from those of her brothers and sisters. This is the first time you visit the place, therefore your host shows you the house while speaking to you about its objects and inhabitants from her own perspective.

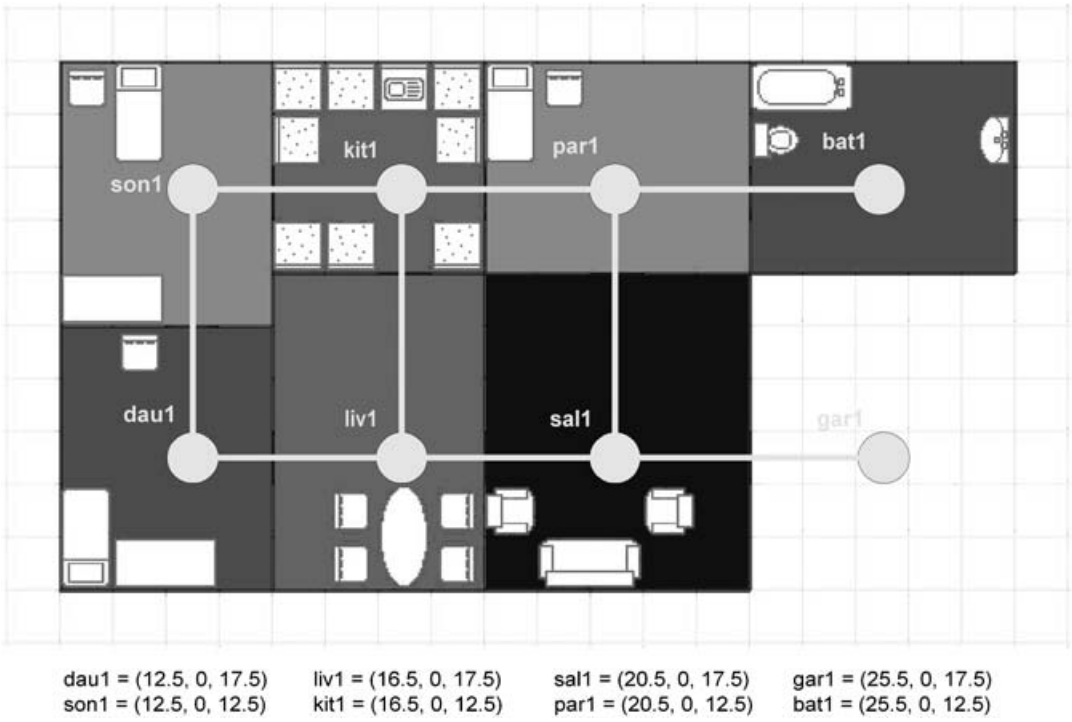

Figure 9 Navigation network of the family house environment

Table 3 Story elements: texts (extract)

Name	Text
SE1	When my parents are away, <*s*> my brother </*s*> usually has affairs with girls on this sofa.
SE2	You'll see <*s*> my mother </*s*> sitting on the sofa sewing if you come in here in the afternoon.
SE3	<*s*> My sister </*s*> watches love movies on tv when she has a while.
SE4	<*s*> The television </*s*> is fantastic, 35 inches and the last generation evolving sound.

Figure 9 displays the navigation network defined for the virtual house. Table 3 shows a few examples of text associated with some story elements for the family house. Table 4 exposes features (type, attributes, subjects, objects and locations) of the story elements that are relevant for the described examples. Figure 10 shows a graph that represents the story elements connected in terms of subject–object relations.

We are not using basic concepts in this case. Concepts are more appropriate for virtual heritage environments where concepts such as the names of kings, civilizations, historical eras, etc. can be annotated so that they are described the first time they appear. A familiar house is such a common context that it is difficult to find a new concept in it. Note that the guide's memory relies on the annotated concepts. Thus, if we wanted to force the memory to work in this case, we could employ the attributes of the story elements (instead of the concepts) as triggers of recollections.

The set of rules aimed at extending the spot are used. In addition, we employ a set of rules to translate the story elements and a set of commonsense rules to enhance the storyboard showing the guide's perspective. Next we show the translation rules that are relevant for the example we are exposing. Conditions to fire these rules include the type of the story element that is potentially translatable and the guide profile.

IF the story element type is **romantic** AND the guide profile is **romantic** THEN the story element type is changed to **romantic-for-romantic** and the guide attitude is set to **joy**

Table 4 Story elements: relevant features for the examples (extract)

Name	Relevant features for the examples	
SE1	type:	affair
	attributes:	affair 1
	subjects:	brother
	objects:	sofa
	locations:	object sofa 1
SE2	type:	women-sew
	attributes:	work 1
	subjects:	mother
	objects:	sofa
	locations:	object sofa 1, any 0.5
SE3	type:	romantic
	attributes:	romantic 1, culture 1
	subjects:	sister
	objects:	tv, movies
	locations:	object tv 1, any 0.5
SE4	type:	ultimate-technology-description
	attributes:	technology 1
	subjects:	tv
	objects:	
	locations:	object tv 1

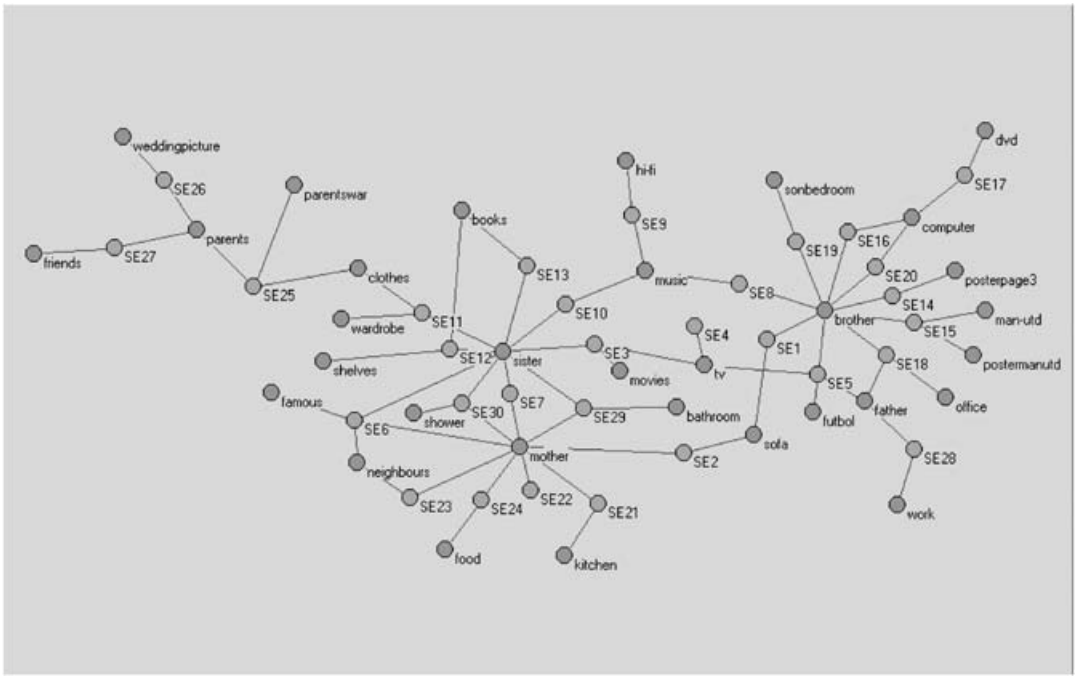

Figure 10 Graph that represents the story elements (of the family house example) connected in terms of subject–object relations

IF the story element type is **romantic** AND the guide profile is **non-romantic** THEN the story element type is changed to **romantic-for-non-romantic** and the guide attitude is set to **disgust**

IF the story element type is **alternative-culture** AND the guide profile is **commercial** THEN the story element type is changed to **alternative-for-commercial** and the guide attitude is set to **disagreement**

IF the story element type is **commercial-culture** AND the guide profile is **alternative** THEN the story element type is changed to **commercial-for-alternative** and the guide attitude is set to **disagreement**

IF the story element type is **ultimate-technology-description** AND the guide profile is **technology** THEN the story element type is changed to **marvellous-technology** and the guide attitude is set to **enthusiasm**

IF the story element type is **women-many-clothes** AND the guide profile is **male** THEN the story element type is changed to **women-worried-about-look** and the guide attitude is set to **ridiculing**

IF the story element type is **women-making-up** AND the guide profile is **male** THEN the story element type is changed to **women-worried-about-look** and the guide attitude is set to **ridiculing**

After the execution of the translation rules, the commonsense rules are employed. Some of these rules are triggered by certain types of translated story elements. These rules generate keywords as consequences, which in turn can trigger other rules, and so on. When the execution of the rules finishes, the obtained keywords are mapped into text sentences that show the guide's perspective. We usually employ several alternative sentences for each possible keyword. Thus, the guide randomly chooses one of the available sentences for each obtained keyword. In fact, the guide only employs a percentage of all the obtained sentences to show her point of view. This percentage is specified by the parameter *frequencyShowingPerspective* and it allows controlling how much the guide is involved in the discourse. In order to simplify the presentation, we show the possible final alternative sentences for each particular type of story element.

For **romantic-for-romantic:**
- Mmmmm… love is really beautiful.
- Ah… love, love!
- How lovey-dovey.

For **romantic-for-non-romantic:**
- How cheesy!
- Love… is so boring, I don't get it.
- Nonsense.

For **alternative-for-commercial:**
- The alternative thing is weird.
- It can't be good if only a few people follow them.

For **commercial-for-alternative:**
- The commercial thing is so boring.
- Masses consume directed by mass media.

For **marvellous-technology:**
- That's incredible.
- It's amazing!
- Fantastic.
- Grand.

For **women-worried-about-look:**
- Typical!
- How much do these lassies waste?
- I'm lucky not to be a lassy.
- The look is too important for women.
- Women are so worried about her appearance…
- Women are fashion victims.

In this example, the guide is initially located in the garden. Note that we give more importance to the spatial score (*spatialScoreWeight*) than to the other two scores for the selection of the spot. That is because we people seem to show the rooms sequentially, following a spatial order. Note

also that we define *desiredGranularity*=2 and *desiredGranularityInSpace*=2. That is, we expect the guide to tell the story in pieces of granularity 2 and we expect her not to tell more than granularity 2 in each space of the house. We use these small values as we expect the guide to be quick showing the house. On the other hand, we give the selection order criterion the highest priority to order the story elements in the storyboard. The reason is that we think we talk in a spontaneous way when showing a house.

8.1 First performance

The first performance corresponds to the daughter who has the following profile: female, romantic, culture, alternative, social. The described algorithms are applied. For a detailed step-by-step description of the flow the work of Ibáñez (2004). As a result, the following performance is obtained:

[**Salon**] My sister watches love movies on tv when she has a while [*showing joy*]. Ah… love, love! My father and my brother never miss a football game on telly.

[**Dinning room**] My sister is a music lover and she usually listens to the most recent alternative rock. My brother listens to the top forty [*showing disagreement*]. Masses consume directed by mass media.

[**Sister's bedroom**] My sister has shelves full of love novels [*showing joy*]. She has three editions of the novel Sense and Sensibility, which is her favourite one [*showing joy*]. How lovey-dovey.

[**Brother's bedroom**] My brother spends ages in his room. He helps my father using office and accounting software for his job.

[**Kitchen**] My mother asks her neighbours to come to the kitchen to chat and share recipes. She knows everything about cooking, she is able to cook anything.

[**Parent's bedroom**] My parents have plenty of friends, they go out a lot. They have this wardrobe which contains mainly my mother's dresses.

[**Bathroom**] My mother and my sister spend half an hour in the shower every morning, so there's always a queue. They are easily found in the bathroom at any time, either making them up or combing their hair.

8.2 Second performance

The second performance corresponds to the son, who has the following profile: male, sport, technology, commercial, non-romantic. The different processes are applied. As a result, the following performance is obtained:

[**Salon**] The television is fantastic, 35 inches and the last generation evolving sound [*showing enthusiasm*]. Grand. My father and my brother never miss a football game on telly.

[**Dinning room**] The hi-fi system plays mini-discs and CDs with mp3 files [*showing enthusiasm*]. It's amazing. My brother listens to the top forty.

[**Sister's bedroom**] My sister has shelves full of love novels [*showing disgust*]. She has three editions of the novel Sense and Sensibility, which is her favourite one [*showing disgust*]. How cheesy!

[**Brother's bedroom**] My brother spends hours playing FIFA with his computer. He helps my father using office and accounting software for his job.

[**Kitchen**] My mother spends several hours every day cooking for the family. She asks her neighbours to come to the kitchen to chat and share recipes.

[**Parent's bedroom**] My parents have plenty of friends, they go out a lot. They have this wardrobe which contains mainly my mother's dresses [*showing ridiculing*].

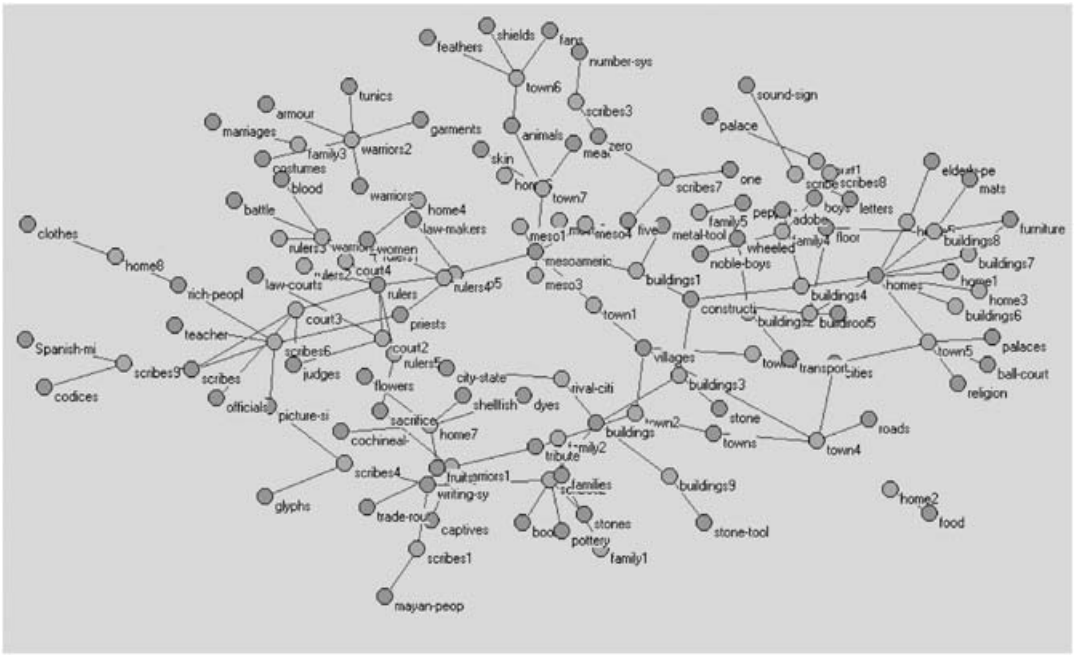

Figure 11 Graph that represents part of the story elements (of the Mayan cities example) connected in terms of subject–object relations

[**Bathroom**] My mother and my sister are easily found in the bathroom at any time, either making them up or combing their hair [*showing ridiculing*]. Typical. My mother and my sister spend half an hour in the shower every morning, so there's always a queue.

9 Example 2: Mayan cities

In addition, we also evaluated the system with a larger knowledge base containing information about Mayan culture and the Mayan city of Palenque. In fact, the system was initially conceived to be used as a guide in the virtual recreations of the Mayan cities of Palenque and Calakmul, in the context of a wider project[1]. In the virtual city of Palenque, the guide tells stories about Mayan culture, religion, science, architecture, etc. Figure 11 shows a graph that represents part of the story elements that compose the knowledge base. These story elements are connected in terms of subject–object and cause–effect relations. Next, we show an extract from the text generated by an occidental female virtual guide in Palenque. In particular, the text was generated by the guide when pointing at a typical Mayan home.

Mesoamerican homes were not just safe places to eat and sleep, they were workplaces too. There were no refrigerators or household appliances, so women had to work hard preparing food for the day's meals or for winter storage. It is always the same story, women sacrificing themselves for the family! Vegetables were cleaned and chopped with stone knives, as there were no metal ones. Beans and chillies were spread out in the Sun to dry, and maize kernels were ground into flour.

Women and girls spent long hours spinning thread and weaving it into cloth, then sewing it into tunics and cloaks for the family. Some women wove cloth to sell or to give to the government as a tax payment. Women slaving away by sewing! Homes were also where most sick or elderly people were cared for.

[1] See `http://www.azc.uam.mx/cyad/posgrado/Lab/index.htm`

10 Experiments

In this section, we describe a study that investigated whether the addition of the storyteller viewpoint is perceived as positive by the users. To do so, three different levels of storytelling construction were employed in the experiment. The impact of the various levels were compared and reported in the following sections.

10.1 Method

Twenty-nine participants between the age of 25 and 36 years took part in the study. Fourteen were men and 15 women. All of them had a degree. They were asked to rate their experience with both computers and VEs. Experience with computers ranged from moderate to very experienced, with a majority rating their experience as reasonable or better. Experience with VEs ranged from none to reasonable, with a majority having little or no experience.

Video sequences of virtual guides showing three different levels of storytelling construction were shown to participants. These video sequences were recorded from the same scenario that was employed in the example previously shown in Section 8. It is a home (see Figure 8) inhabited by part of a family: father, mother, son and daughter. In addition, another son and another daughter lived on their own. The participant is supposed to visit the house invited by one of the latter children. The house is populated by objects that are associated with memories related to the members of the family. In principle, all of the children are supposed to have almost the same shared memories. However, each child has a well-defined profile, different from those of her brothers and sisters. This is the first time the participant visits the place, therefore his host shows him the house while speaking to him about its objects and inhabitants from her own perspective.

Participants were shown six video sequences (each about 3 min in length) grouped in three pairs. Each pair was composed of two sequences. One of them showed the son's performance, whose profile is male, non-romantic and Manchester United supporter. The other one showed the daughter's performances, whose profile is female and romantic. Each pair of sequences corresponds to one of the three conditions listed below:

- Condition *A*: The guide does not show her point of view.
- Condition *B*: The guide shows her point of view by statements added to her discourse.
- Condition *C*: The guide shows her point of view by both statements added to her discourse and animations (body language) showing her emotional reaction.

The order in which the conditions were presented to the participants was randomized and distributed so that a similar number of participants started with each condition.

Upon completion of each condition, participants were asked to assign a subjective rating to 10 different aspects of their experience according to a 7-point Likert scale (Likert, 1932) (see Appendix A).

As we were interested in evaluating how participants perceive the addition of the point of view of the guide, we modified and constrained the knowledge base so that the only differences between the performances of both guides (son and daughter) are precisely those due to the addition of the guide perspective. Next, we show the performances of both guides, particularly the performances showing the storyteller point of view. The performance corresponding to the son is:

[**Salon**] If you come in here in the afternoon you'll see my mother sewing on the sofa. My father and brother usually watch football games on tv. My sister watches love movies on tv when she has a while. How cheesy!! When my parents are off, my brother usually has affairs with girls on this sofa.

[**Dinning room**] My mother and sister sit usually around this table to talk about neighbours and celebrities. You know, two or more women together… gossip!

[**Sister's bedroom**] My sister's wardrobe is not only big, but full. Typical!!

[**Brother's bedroom**] My brother has this page three girl poster on the wall since he was fourteen. This poster shows that my brother is a Man United supporter. The best team on earth, of course.

[**Kitchen**] My mother spends several hours every day cooking for the family in the kitchen.

[**Parent's bedroom**] This is my parent's wedding picture. My parents say they are very much in love and after so many years they still kiss each other every day in a lovely way. Love… is so boring, I can't understand how people… My parents' wardrobe contains mainly my mother's dresses. The importance of the look!

[**Bathroom**] You can easily find my mother and sister in the bathroom at any time, either making them up or combing their hair. Women are so worried about their appearance…

The following performance corresponds to the daughter:

[**Salon**] If you come in here in the afternoon you'll see my mother sewing on the sofa. Women slave away by sewing! My father and brother usually watch football games on tv. My sister watches love movies on tv when she has a while. Mmmmmmm…aaaahhhh… love is really beautiful. When my parents are off, my brother usually has affairs with girls on this sofa. Anyway… only love can make you happy.

[**Dinning room**] My mother and sister sit usually around this table to talk about neighbours and celebrities.

[**Sister's bedroom**] My sister's wardrobe is not only big, but full.

[**Brother's bedroom**] My brother has this page three girl poster on the wall since he was fourteen. You know, men are pigs! This poster shows that my brother is a Man United supporter.

[**Kitchen**] My mother spends several hours every day cooking for the family in the kitchen. It is always the same story, women sacrificing themselves for the family!

[**Parent's bedroom**] This is my parent's wedding picture. My parents say they are very much in love and after so many years they still kiss each other every day in a lovely way. Ah… love, love! My parents' wardrobe contains mainly my mother's dresses.

[**Bathroom**] You can easily find my mother and sister in the bathroom at any time, either making them up or combing their hair.

10.2 Statistical analysis

For each studied aspect, ordinal Likert scores were compared across conditions. Wilcoxon signed rank tests (Wilcoxon, 1945) were applied between conditions A and B, A and C, and B and C. All tests were two-tailed.

10.3 Results

Conditions A (no viewpoint) and B (guide's viewpoint shown through the guide's discourse) differed significantly in 8 out of 10 questions. Figure 12 compares participants' ratings for conditions A and B. Participants found the guide's discourse in condition B more believable ($Z = 4.291$, $P < 0.001$), natural ($Z = 4.549$, $P < 0.001$), intelligent ($Z = 4.269$, $P < 0.001$) and emotional ($Z = 4.728$, $P < 0.001$) than in condition A. They also found the guide's general behaviour in condition B more intelligent ($Z = 2.639$, $P < 0.01$) and emotional ($Z = 2.297$, $P < 0.05$). Furthermore, participants felt more engaged by the guide's general behaviour in condition B than in condition A ($Z = 4.137$, $P < 0.001$). They also thought that they got to know the guide's personality better in condition B ($Z = 4.438$, $P < 0.001$).

Conditions A (no viewpoint) and C (guide's viewpoint shown through both the guide's discourse and the guide's animation) differed significantly in all the questions. Figure 13 compares participants' ratings for conditions A and C. Participants found the guide's discourse in condition C more believable ($Z = 4.336$, $P < 0.001$), natural ($Z = 4.256$, $P < 0.001$), intelligent ($Z = 4.493$, $P < 0.001$) and emotional ($Z = 4.665$, $P < 0.001$) than in condition A. They also found the guide's general behaviour in condition C more believable ($Z = 4.244$, $P < 0.001$), natural ($Z = 4.731$, $P < 0.001$), intelligent ($Z = 4.658$, $P < 0.001$) and emotional ($Z = 4.745$ $P < 0.001$). Furthermore, participants felt more engaged by the guide's general behaviour in condition C than in condition

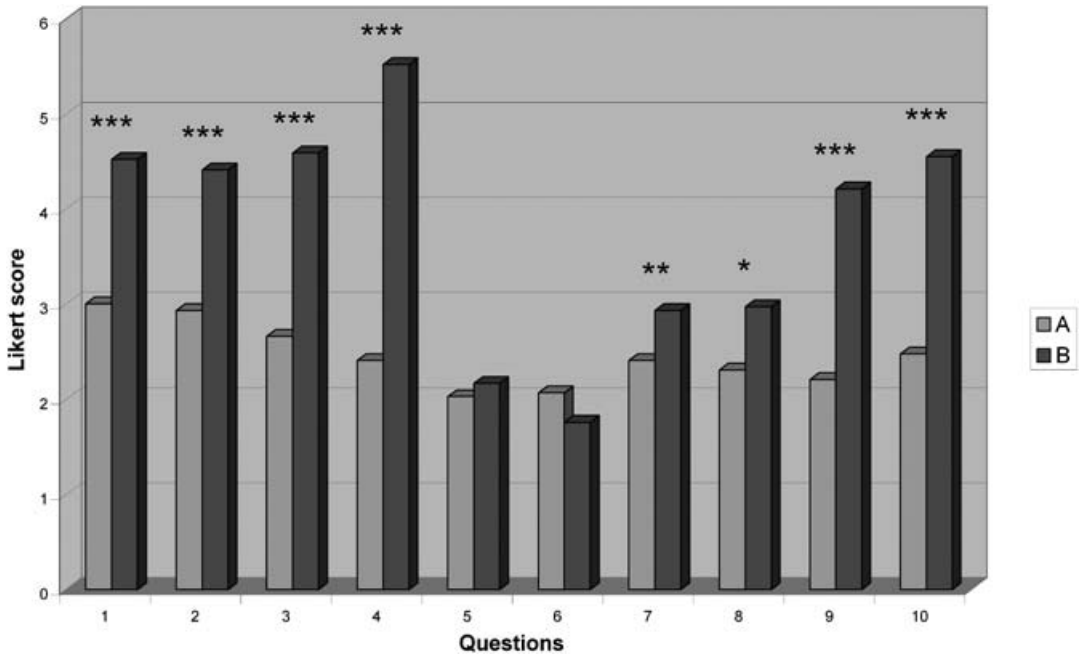

Figure 12 Significant differences in subjective ratings between the conditions: (A) no viewpoint and (B) viewpoint shown through discourse (*$P < 0.05$, **$P < 0.001$, ***$P < 0.001$)

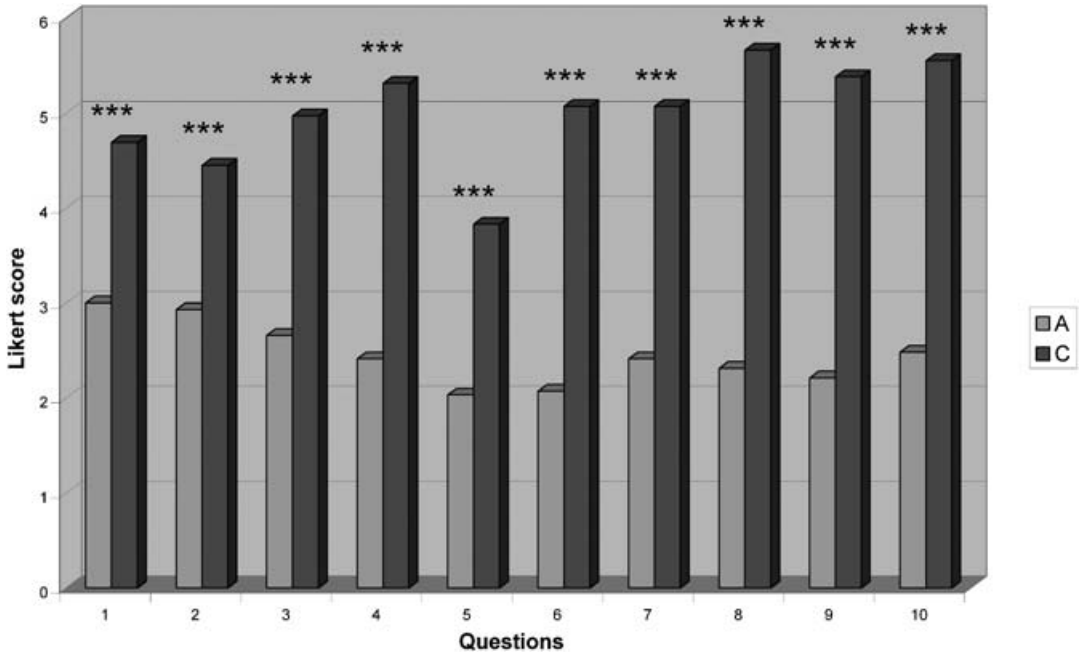

Figure 13 Significant differences in subjective ratings between the conditions: (A) no viewpoint and (C) viewpoint shown through both discourse and animations (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$)

A ($Z = 4.481$, $P < 0.001$). They also thought that they got to know the guide's personality better in condition *C* ($Z = 4.639$, $P < 0.001$).

Conditions *B* (guide's viewpoint shown through the guide's discourse) and *C* (guide's viewpoint shown through both the guide's discourse and the guide's animation) differed significantly in 7 out of 10 questions. Figure 14 compares participants' ratings for conditions *B* and *C*. Participants found the guide's discourse in condition *C* more intelligent ($Z = 2.202$, $P < 0.05$) than in condition *B*. They also found the guide's general behaviour in condition *C* more believable ($Z = 4.462$,

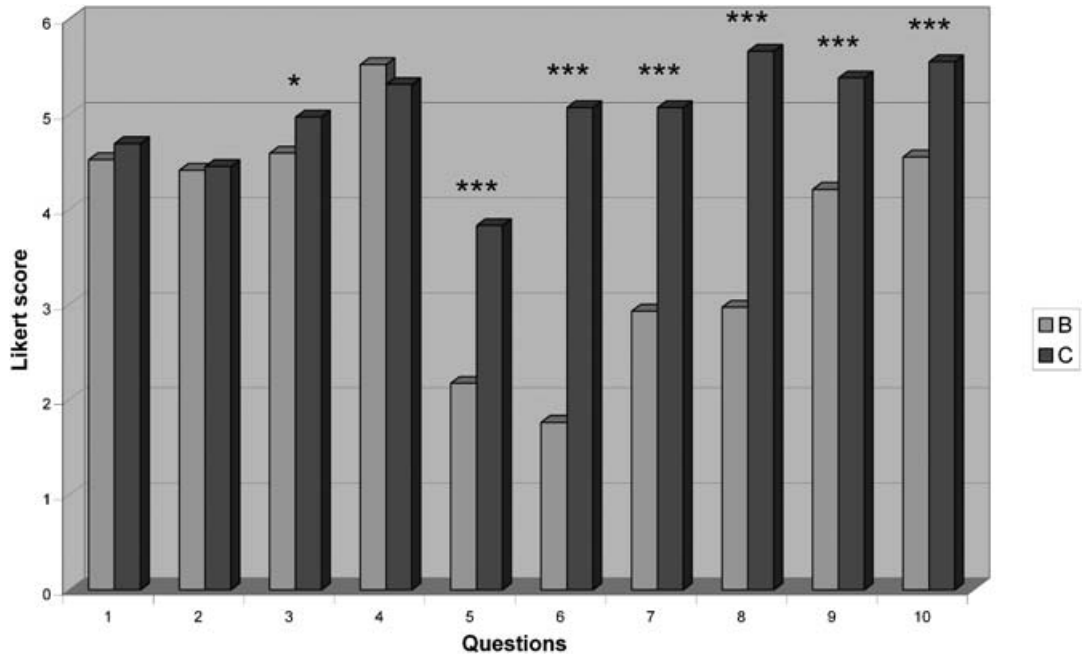

Figure 14 Significant differences in subjective ratings between the conditions: (B) viewpoint shown through discourse and (C) viewpoint shown through both discourse and animations (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$)

$P < 0.001$), natural ($Z = 4.499$, $P < 0.001$), intelligent ($Z = 4.492$, $P < 0.001$) and emotional ($Z = 4.573$, $P < 0.001$). Furthermore, participants felt more engaged by the guide's general behaviour in condition *C* than in condition *B* ($Z = 3.892$, $P < 0.001$). They also thought that they got to know the guide's personality better in condition *C* ($Z = 3.356$, $P < 0.001$).

10.4 Discussion

The results show that the guide's discourse is perceived as more believable, natural, intelligent and emotional when it includes the guide's viewpoint. However, in that case, the guide's general behaviour is not perceived as more believable or natural, even though it is perceived as more intelligent and emotional. Why does it happen? On the one hand, maybe there is a natural trend to perceive an improvement in the general behaviour when one of its key components is improved. That would explain that the participants perceived the general behaviour as more intelligent and emotional when they perceived a clear improvement in the discourse. On the other hand, several participants commented that they perceived the guide's behaviour as less natural when the guide showed her viewpoint because in that case the divorce between the discourse and the guide's body language was more evident. That would explain that the participants did not perceive the general behaviour as more believable or natural. That would also justify that the significant differences in the intelligence and emotion of the behaviour were not very high ($P < 0.01$ and $P < 0.05$, respectively).

The guide's general behaviour is perceived as more believable, natural, intelligent and emotional when the guide's viewpoint is shown through both the guide's discourse and the guide's animation than when it is only shown through the guide's discourse. However, the guide's discourse is not perceived as more believable, natural or emotional, even though it is perceived as more intelligent (with a small significant difference $P < 0.05$). Thus, it seems that the addition of the guide's body language reinforcing the guide's discourse improves the perception of the general guide's behaviour, but not the perception of the guide's discourse.

Thus, from the experiment we can conclude that the addition of the guide's viewpoint to the virtual guide's discourse makes the discourse appear more believable, natural, intelligent and emotional. However, the guide's viewpoint must also accordingly be shown through the guide's animation so that the general behaviour of the guide appears as more believable and natural.

In spite of particular issues, the guide's general behaviour is seen as more engaging when the guide's discourse includes the guide's viewpoint than when it is not included. The guide's general behaviour is even more engaging when the guide's viewpoint is shown through both the guide's discourse and the guide's animation. Furthermore, the users get to know the guide's personality better when the guide's discourse includes the guide's viewpoint than when it is not included. They get to know it even better when the guide's viewpoint is shown through both the guide's discourse and the guide's animation. Thus, the experiments recommend the addition of the guide's viewpoint in both the discourse and body language of the guide.

11 Conclusion

We have proposed an open and reusable architecture for the construction of virtual guides who tell stories about the worlds they inhabit from their own perspective. It consists of two main components: the structure of a knowledge base to code the narrative knowledge about the worlds and a novel hybrid algorithm for the generation of narratives showing the storyteller's viewpoint. This system relies on the information stored in the designed knowledge base.

The storytelling process involves several mechanisms, which can be summarized as follows. With every step, the virtual guide decides (improvises) where to go and what to tell there, through a process that selects a location and a story element (this improvization tries to emulate what occurs in the mind of a real guide by taking into account the spatial structure of the world, her memories and her own interests). Then the story element is extended and translated from the virtual guide perspective and enhanced with information generated through simple commonsense rules. At the same time, a related guide attitude is induced. Next, the new story elements are used to generate text sentences and special effects, while the guide attitude is used to select suitable guide actions. Finally, the storytelling process is carried out, that is the guide is moved to the selected location, the environment is updated with the required special effects and the textual sentences are told while the guide shows her attitude through the selected actions.

The complete architecture for storytelling was implemented. The core of the virtual guide is an agent that is able to connect to Unreal Tournament worlds (through Gamebots) and VRML worlds (running on a DeepMatrix server).

We carried out a study with users in order to both validate the implemented storytelling system and test whether the users consider it significantly better or worse (in terms of several aspects) adding the virtual guide perspective (as both text and animations showing emotions) to the guide discourse. The participants in the study were between the age of 25 and 36 years, and all of them had a degree. The results show both that the guide's general behaviour is more engaging and that the users get to know her personality better when the guide's viewpoint is added. Furthermore, the results indicate that the addition of the guide's viewpoint to the virtual guide's discourse makes the discourse appear more believable, natural, intelligent and emotional. The guide's viewpoint must also accordingly be shown through the guide's animation (body language) so that the general behaviour of the guide appears more believable and natural.

We plan to work on several issues in the near future in order to extend and improve this research. In the real world, a guide takes into account not only his own profile but also the profile of the people being guided. In this sense, we plan to extend the algorithms so that the storytelling construction takes into account not only the virtual guide's profile but also a kind of combination of the guide's profile with the user's profile. For the system to work, it needs appropriate narrative information in the knowledge base. In order to ease the knowledge acquisition, we intend to define a suitable methodology, inspired in usual techniques from the knowledge engineering area. Finally, we will work towards the improvement of the NLG component of the system.

Acknowledgments

The authors would like to express their gratitude to Leticia Lipp for generous proofreading.

Appendix A: Questions used in the study

The participants in the storytelling experiment were asked the following questions, which were rated on a 7-point Likert scale.

1. Please rate how **believable** the guide's discourse is.
2. Please rate how **natural** the guide's discourse is.
3. Please rate how **intelligent** the guide's discourse is.
4. Please rate how **emotional** the guide's discourse is.
5. Please rate how **believable** the guide's general behaviour is.
6. Please rate how **natural** the guide's general behaviour is.
7. Please rate how **intelligent** the guide's general behaviour is.
8. Please rate how **emotional** the guide's general behaviour is.
9. Please rate how **engaged** you feel by the guide's general behaviour.
10. Please rate how much you think you **got to know the guide's personality**.

In addition, they were offered the opportunity to comment and suggest.

References

Adobbati, R., Marshall, A. N., Scholer, A., Tejada, S., Kaminka, G. A. & Schaffer, S. 2001. Gamebots: a 3D virtual world test-bed for multi-agent research. In *Second International Workshop on Infrastructure for Agents, MAS, and Scalable MAS*, Montreal, Canada.

Aylett, R. 1999. Narrative in virtual environments—towards emergent narrative. In *Narrative Intelligence Symposium*, AAAI 1999 Fall Symposium Series.

Aylett, R. & Cavazza, M. 2001. *Intelligent Virtual Environments—a State-of-the-Art Report*. State of Art Reports, Eurographics.

Aylett, R. & Luck, M. 2000. Applying artificial intelligence to virtual reality: intelligent virtual environments. *Applied Artificial Intelligence* **14**(1), 3–32.

Batagelj, V. & Mrvar, A. 2003. Pajek—analysis and visualization of large networks. In *Graph Drawing Software, Mathematics and Visualization*, Jünger, M. & Mutzel, P. (eds). Springer-Verlag.

Bates, J. 1994. The role of emotion in believable agents. *Communications of the ACM* **37**(7), 122–125.

Branigan, E. 1992. *Narrative Comprehension and Film*. Routledge.

Braun, N. 2002. Automated narration—the path to interactive storytelling. In *The 2nd International Workshop on Narrative and Interactive Learning Environments*, Edinburgh, Scotland.

Brooks, K. M. 1999. *Metalinear Cinematic Narrative: Theory, Process, and Tool*. PhD thesis, School of Architecture and Planning, MIT.

Cavazza, M., Charles, F. & Mead, S. J. 2001. Characters in search of an author: AI-based virtual storytelling. In *International Conference on Virtual Storytelling*, 145–154.

Domike, S., Mateas, M. & Vanouse, P. 2002. The recombinant history apparatus presents: Terminal Time. In *Narrative Intelligence*, Mateas, M. & Sengers, P. (eds). John Benjamins.

Friedman-Hill, E. 2004. Jess the rule engine for the Java platform. Available at `http://herzberg.ca.sandia.gov/jess/`

Hoorn, J. F., Konijn, E. A. & van der Veer, G. C. 2003. Virtual reality: do no augment realism, augment relevance. *Upgrade* **4**(1), 18–26.

Ibáñez, J. 2004. *An Intelligent Guide for Virtual Environments with Fuzzy Queries and Flexible Management of Stories*. PhD thesis, Departamento de Ingenieria de la Informacion y las Comunicaciones, Universidad de Murcia.

Isbister, K. & Doyle, P. 1999. Touring machines: Guide agents for sharing stories about digital places. In *Working Notes of the 1999 AAAI Fall Symposium on Narrative Intelligence*. `http://citeseer.ni.nec.com/285732.html`

Jul, S. & Furnas, G. W. 1997. Navigation in electronic worlds. *SIGCHI Bulletin* **29**(4), 44–49.

Kaminka, G. A., Veloso, M. M., Schaffer, S., Sollitto, C., Adobbati, R. & Marshall, A. N. 2002. Gamebots: a flexible test bed for multiagent team research. *Communications of the ACM* **45** (1 Special Issue: Game engines in scientific research), 43–45.

Knott, A., Mellish, C., Oberlander, J. & O'Donnell, M. 1996. Sources of flexibility in dynamic hypertext generation. In *Proceedings of the 8th International Workshop on Natural Language Generation*, Herstmonceux Castle, UK. `http://citeseer.ni.nec.com/knott96sources.html`

Laird, J. E. 2002. Research in human-level AI using computer games. *Communications of the ACM* **45** (1 Special Issue: Game engines in scientific research).

Lieberman, H. & Liu, H. 2002. Adaptive linking between text and photos using common sense reasoning. In *The 2nd International Conference on Adaptive Hypermedia and Adaptive Web Based Systems, AH2002*, Malaga, Spain.

Lieberman, H., Rosenzweig, E. & Singh, P. 2001. ARIA: an agent for annotating and retrieving images. *IEEE Computer*, 57–61.

Likert, R. 1932. A technique for the measurement of attitudes. *Archives of Psychology* **140**, 55.

Liu, H. & Lieberman, H. 2002. Robust photo retrieval using world semantics. In *The 3rd International Conference on Language Resources and Evaluation Workshop: Using Semantics for Information Retrieval and Filtering (LREC2002)*, Las Palmas, Canary Islands, Spain.

Liu, H. & Singh, P. 2002. Makebelieve: using commonsense to generate stories. In *Proceedings of the 20th National Conference on Artificial Intelligence (AAAI-02), Student Abstracts*, 957–958. Edmonton, Canada.

Louchart, S. & Aylett, R. 2002. Narrative theory and emergent interactive narrative. In *The 2nd International Workshop on Narrative and Interactive Learning Environments*, Edinburgh, Scotland.

Mateas, M. 1999. An Oz-centric review of interactive drama and believable agents. In *AI Today: Recent Trends and Developments*, Wooldridge, M. & Veloso, M. (eds). Lecture Notes in AI **1600**, 297–328. Springer. First appeared in 1997 as Technical report CMU-CS-97-156. Computer Science Department, Carnegie Mellon University.

Mateas, M. & Stern, A. 2000. Towards integrating plot and character for interactive drama. In *Working Notes of the Social Intelligent Agents: The Human in the Loop Symposium*. AAAI Fall Symposium Series, AAAI Press. A shorter version of this paper appears in Dautenhahn, K. (ed.), *Socially Intelligent Agents: The Human in the Loop*. Kluwer, 2002.

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

MySQL. 2004. MySQL: The world's most popular open source database. Available at `http://www.mysql.com/`

Oberlander, J., O'Donnell, M., Knott, A. & Mellish, C. 1998. Conversation in the museum: experiments in dynamic hypermedia with the intelligent labelling explorer. *New Review of Hypermedia and Multimedia* **4**, 11–32.

O'Donnell, M., Knott, A., Oberlander, J. & Mellish, C. 2000. Optimising text quality in generation from relational databases. In *Proceedings of the 1st International Natural Language Generation Conference*, Mitzpe Ramon, Israel.

Polti, G. 1922. *The 36 Dramatic Situations*. The Writer, Inc.

Propp, V. 1968. *Morphology of the Folktale*. University of Texas Press. First edition 1928.

Reitmayr, G., Carroll, S., Reitemeyer, A. & Wagner, M. G. 1999. Deepmatrix—an open technology based virtual environment system. *The Visual Computer* **15**(7/8), 395–412.

Robertson, J. & Oberlander, J. 2002. Ghostwriter: educational drama and presence in a virtual environment. *Journal of Computer Mediated Communication* **80**(1).

Singh, P. 2002. The public acquisition of commonsense knowledge. In *Proceedings of AAAI Spring Symposium on Acquiring (and Using) Linguistic (and World) Knowledge for Information Access*, Palo Alto, CA.

Spierling, U., Grasbon, D., Braun, N. & Iurgel, I. 2002. Setting the scene: playing digital director in interactive storytelling and creation. *Computers & Graphics* **26**, 31–44.

Thompson, S. 1955. *Motif Index of Folk Literature*. Indiana University Press.

Tozzi, V. 2000. Past reality and multiple interpretations in historical investigation. *Studies in Social and Political Thought* **2**, 41–57.

Wilcoxon, F. 1945. Individual comparisons by ranking methods. *Biometrics* **1**, 80–83.

Young, R. M. 2001. An overview of the Mimesis architecture: integrating intelligent narrative control into an existing gaming environment. In *The Working Notes of the AAAI Spring Symposium on Artificial Intelligence and Interactive Entertainment*, Stanford, CA.