User Model User-Adap Inter (2008) 18:417–453
DOI 10.1007/s11257-008-9053-1

ORIGINAL PAPER

A stroll with Carletto: adaptation in drama-based tours with virtual characters

Rossana Damiano · Cristina Gena · Vincenzo Lombardo · Fabrizio Nunnari · Antonio Pizzo

Received: 10 May 2007 / Accepted in revised form: 1 July 2008 / Published online: 22 August 2008

Abstract In this paper, we present an application for character-based guided tours on mobile devices. The application is based on the Dramatour methodology for information presentation, which incorporates a dramatic attitude in character-based presentations. The application has been developed for a historical site and is based on a virtual character, "Carletto", a spider with an anthropomorphic aspect, who engages in a dramatized presentation of the site. Content items are delivered in a location-aware fashion, relying on a wireless network infrastructure, with visitors who can stroll freely. The selection of contents keeps track of user location and of the interaction history, in order to deliver the appropriate type and quantity of informative items, and to manage the given/new distinction in discourse. The communicative strategy of the character is designed to keep it believable along the interaction with the user, while enforcing dramatization effects. The design of the communicative strategy relies on

R. Damiano (✉) · C. Gena · V. Lombardo
Dipartimento di Informatica, Università di Torino, Torino, Italy
e-mail: rossana@di.unito.it

C. Gena
e-mail: cgena@di.unito.it

V. Lombardo
e-mail: vincenzo@di.unito.it

A. Pizzo
Dipartimento DAMS, Università di Torino, Torino, Italy
e-mail: antonio.pizzo@unito.it

R. Damiano · C. Gena · V. Lombardo · A. Pizzo
CIRMA, Centro Interdipartimentale di Ricerca per la Multimedialità e l'Audiovisivo, Università di Torino, Torino, Italy

F. Nunnari
Virtual Reality and Multi Media Park, Torino, Italy
e-mail: fnunnari@vrmmp.it

the fact that the units of the presentation are tagged with metadata concerning their content and communicative function. The description of the application is accompanied by an evaluation study based on a sample of about 300 visitors, carried out in April 2006, when the installation was open to the public for 1 week.

Keywords Interactive drama · Virtual characters · Cultural heritage · Mobile computing · Adaptive applications

1 Introduction

Access to different forms of cultural heritage requires different personalization formats. In this respect, it is enlightening to think over the dichotomy between museums and historical locations.

In museums, collections often derive from heterogeneous sources; so, they require the intervention of experts, the museum curators, to provide some continuity to the exhibition by arranging the spatial layout and the physical position of the items. No matter what is the underlying principle for organizing the visit—the 'glue' may be a historical or artistic period—the goal of an exhibition is to provide an interpretative mediation between the items of a collection and the visitor. Since it is unlikely that the same mediation works for the totality of the visitors, personalization in museums attempts to fill this gap by setting up 'different exhibitions' to match the interests and expertise of different categories of visitors—see Pechenizkiy and Calders (2007) for a survey of the methodologies of museum visit personalization.

On the contrary, historical locations basically reflect the social and political functions they once played. In general, the historical reasons that have determined the appearance of a location are normally unknown to the majority of visitors. So, even in presence of strong underlying organizing principles, the visitor is not able to assign a meaning to the location without the help of an expert, e.g. a guide. This situation has been effectively depicted by Giles Waterfield's annotation about the feeling of emptiness suggested by "clocks that don't tell the time" and "vases that contain no flowers" (Waterfield 1999). The adaption of guidance systems to personal interests requires a design that takes into account the original spatial layout (so mobile communication infrastructures are a must); the absence of items or their inert condition orients adaptation towards interpretive and reviving strategies.

Orthogonal to the dichotomy between museum and historical location, are the following design approaches that have been put at work (also through integration) for the presentation delivery in the field of cultural heritage.

1.1 Storytelling

According to story theorists like Bruner (1991), stories provide a framework for making sense of events and their meaning (Mateas and Senger 1999). Stories can convey large amounts of information in a compact format, which is easily assimilated by the user (Gershon and Page 2001). In cultural heritage access, storytelling has been applied

to virtual tours in exhibitions (Hoffmann and Herczeg 2003), museums (Springer et al. 2004) and historical locations (Johnson 2004).

1.2 Characters

The use of characters improves the naturalness of the interaction with the user, making the system appear more responsive and cooperative. The framework of Embodied Conversational Agents (Cassel et al. 2000) addresses the capability of engaging in a natural face-to-face dialogue in which verbal and non-verbal modalities coordinate to build an empathic relationship with the user. These properties have promoted the use of embodied agents in educational applications (Lester et al. 2000; Ieronutti and Chittaro 2007) and in cultural heritage applications, where embodied agents have been proposed as virtual guides in contexts that range from city guides (Gustafson et al. 1999; Dybkjær and Dybkjær 2004) to exhibitions (Kopp et al. 2005).

1.3 Multimodality

Thanks to increasing high-quality graphics and sound, multimodal applications nowadays support a multiplicity of interaction paradigms—see, e.g. (Callaway et al. 2005; Solon et al. 2004)—that are particularly required to cope with variable environment conditions (e.g. the use of sound instead of scarcely luminous small screens in the daylight, or the opposite for noisy environments) and the limitations of computational resources (memory, battery, unreliable wireless connection, screen).

1.4 Mobile devices and wireless infrastructures for mobility

Mobility is an obvious solution for the personalization of the access to cultural heritage and the field is mature enough to promote debates on the application values on both the viewpoints of the target audience (Proctor 2005) and the designers/implementors (Raptis et al. 2005). Wireless applications are relevant for historical locations for two reasons: the first, as mentioned above, is that it is not possible to arrange variable spatial layouts for enhancing presentations; the second is to guarantee a continuity in the presentation through a unique device that is bound to an individual visitor.

This paper presents an application for a mobile device, "Carletto the spider", developed for a historical location, where the presentation is enacted by a virtual character who engages in a dramatized interaction with the visitor. The design of the application "literally" follows Giles Waterfield's suggestion that the visit of a historical location could be transformed into an immersive experience by adding "a touch of theatre" to the visitor's experience (Waterfield 1999). The virtual character is an anthropomorphized spider, Carletto, who appears on the screen of a portable device that the visitor carries with her during the visit. To deliver his presentation, Carletto engages in a *dramatic monologue* (rather than being a plain storyteller); the audiovisual output of his performance is displayed on the screen of the portable device, together with further information materials (schemata, images, etc.); the application is aware of the visitor

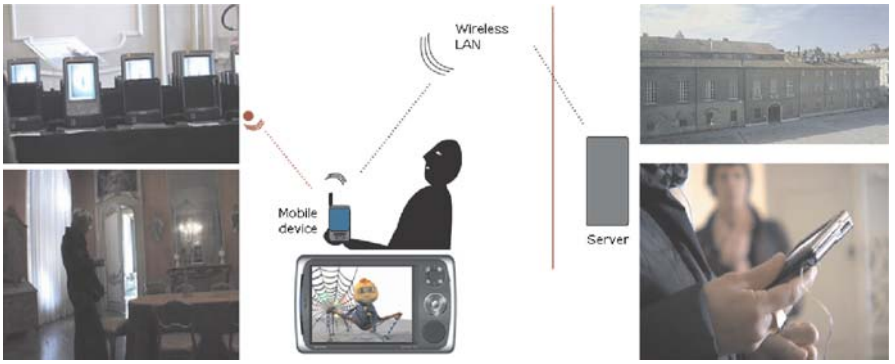

Fig. 1 The scenario of "Carletto the spider"

location, so the character adapts the type of information to the visitor location, and the degree of specificity to the time spent in the location by the visitor (Fig. 1).

The application is based on the Dramatour methodology for information presentation (Damiano et al. 2006b), that incorporates principles of drama in character-based presentations. The working assumption of the Dramatour methodology is that a character, who acts in first person and shares the visitor's present time and space, yields a powerful effect of physical and emotional presence. This assumption relies on the use of a certain type of media for the presentation delivery, the so-called *dramatic media*. According to Esslin, dramatic media, like television or cinema, add emotional qualities to the content they convey, for the very same reason of displaying live action (Esslin 1987). Emotional properties improve the effectiveness of content reception (Laurel 1993) and the user engagement, bringing an affective element to her experience (Picard 1997).

The Dramatour methodology is briefly described in the next section. In Sect. 3, we describe the application, with the representations it relies on—with emphasis on metadata annotation and communicative model—and its architecture. Section 4 describes the evaluation conducted to assess the effectiveness of the application, with a concern for the drama-based methodology. The comparison with related work (Sect. 5) and Conclusions (Sect. 6) end the paper.

2 The Dramatour methodology

The Dramatour methodology provides a framework for developing dramatized guided tours.

The methodology relies on three major tenets. First, the dramatized presentation, acted by a character, is authored in a specific layout for mobile devices, and is accompanied by the design of a specific strategy of interaction with the visitor. Second, the presentation is factorized into semantically tagged audiovisual units. Third, these units are edited on-the-fly during the visit, in a way that accounts for adaptation and interactivity. As the work of Hardman (2005) points out, the standardization of the work flow in multimedia production is very relevant for the development of and the

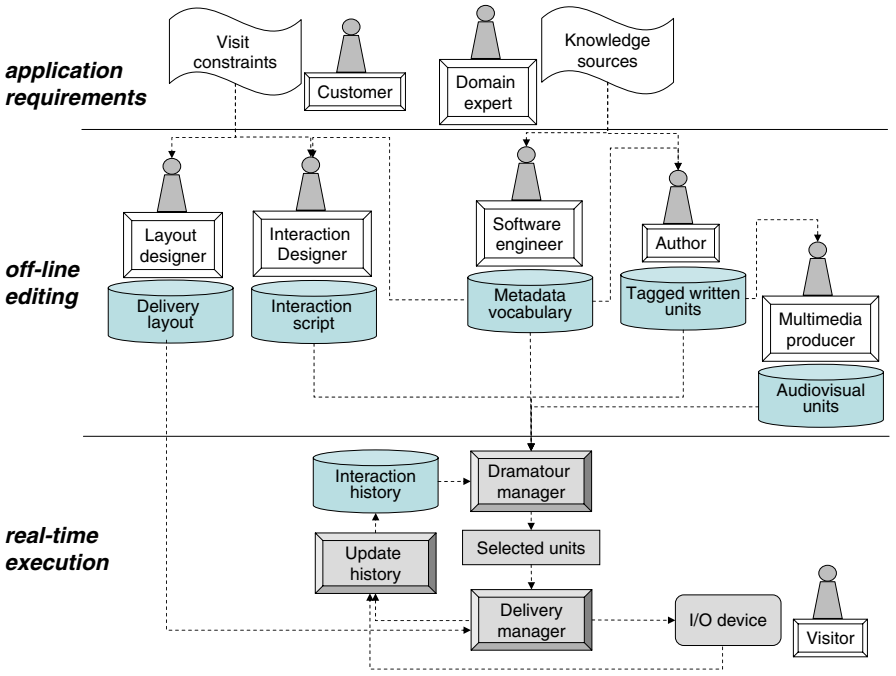

Fig. 2 The production pipeline of Dramatour applications. Boxes represent roles (gray boxes are computer-filled roles), cylinders represent knowledge bases. Dashed line represent the knowledge flow among roles

interoperability between applications. In the following, we summarize the design and production pipeline underpinning the methodology; details can be found in Nunnari et al. (2008). We will refer to Fig. 2 to support the exposition.

The top of the figure represents the first phase, in which the requirements of the application to be developed are elicited and collected (*application requirements*). In this phase, the management staff of the location, who plays the role of the *Customer*, provides the relevant constraints for the guided visit, like the duration of the visit, the profile of the target visitors, the organization of the visit (e.g. in groups or individuals), etc. A *Domain expert*, e.g. the curator of the location, provides the contents that are to be included in the presentation.

The second phase corresponds to the offline creation of the units in which the presentation is factorized (*off-line editing*, center of the figure). Given the contents provided by the domain expert, the *Software engineer* develops the domain-specific metadata according to which the units of the presentation will be semantically tagged. The *Author* writes the presentation by segmenting it into independent units; redundancy is required to give the presentation a certain degree of variability. Writing and tagging are strictly related, since tags provide the conceptual framework the author must refer to during the writing process. It is up to the author to create the character, and to enforce the dramatization in the text of the presentation units; from this point of view, the Dramatour author is a "procedural" author, in the sense established by

Murray (1998). Units become then the input to the work of the *Multimedia producer*, who manages the generation of the tagged audiovisual units that constitute the knowledge base of the application. This task can range from Flash animation to complex 3D graphics with sound and overlay text.

The *Interaction script* is a high-level guidance for the dramatic presentation and is developed by the *Interaction designer* taking into account the visit constraints. The visit constraints are the input also for the *Layout designer*, who provides the visual layout that will frame the presentation on the mobile device (*Delivery layout*). The interaction script defines the communicative strategy adopted by the character during the presentation, taking the input of the visitor into account. This step is crucial to the creation of an effective character: the believability of the character depends, to a large degree, on the display of a coherent and contextually appropriate communicative behavior along the interaction with the visitor.

Both the author and the interaction designer contribute to the development of the annotation metadata. In particular, the author contributes to the annotation for the informative content, while the interaction designer contributes to the annotation of the interactional content. In fact, the presentation units are the elements through which both the interaction strategy and the information delivery are implemented. So, the author, who writes and annotates the units, should be aware of the interactional and informative contribution of each unit. The software engineer provides the author and the interaction designer with the ultimate formalization of the metadata, by taking into account also the insights from the domain experts.

The result of the offline editing is a segmented presentation in a dramatic format, i.e., a format in which the personification of an artificial character, who acts at present time and space—according to the paradigm of dramatic media (Esslin 1987)—is accompanied by specific properties, among which the most relevant is the capability of displaying some emotional conflicts concerning the character's values (McKee 1997; Seger 1990). Conflict is intended here as a tool to help the visitor to build a bond with the character (hence the author should avoid any conflicts between the visitor and the character). The visitor is part of the process of dramatic achievement, in the sense that, while visiting the site, she provides feedbacks that are used to shape the character-enacted presentation.

The principle of drama "first-person, present-time action" is achieved through the authoring of the presentation units, which is in charge of making the character's personality, values and conflict emerge (McKee 1997; Seger 1990). The character functions as a mediator between cultural contents and the visitor, acting more as a quality-improving interface rather than a simple information supplier. The character must show a space-time consistency, i.e., it must be related to the current space and time, avoiding the "ghost-from-the-past" effect. The author has to design both his explicit-external motivations (the character must show a desire to tell a story, a sense of duty that forces him to be there with the visitor) and his implicit inner motivation (the character must show a personality consistent with his wills and duties). In other words, the character needs an official reason to be there and, at the same time, he must have a personality that helps him to carry out the duty properly.

At the execution time (third phase, *real-time execution*, Fig. 2, bottom), the tagged repository of the audiovisual units and the interaction script become the input to the

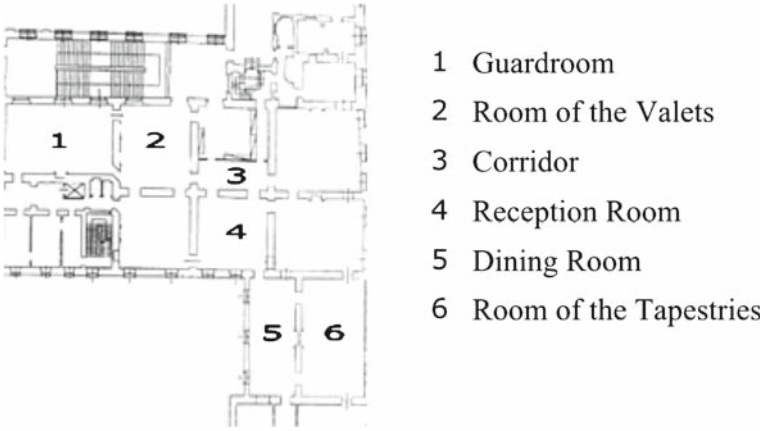

Fig. 3 The map of the Royal Apartments in Palazzo Chiablese, in Torino. The visit started from the Guardroom (room 1), but the visitors were free to stroll in the apartments without following a predefined order

software process that manages the application, the *Dramatour manager*. The Dramatour manager selects the units to deliver based on the interaction script. The interaction script must take into account the input provided by the visitor and for the past system/visitor interaction, incrementally recorded in the interaction history after the delivery of each unit.

3 The application "Carletto the spider"

The application "Carletto the spider" was developed for the temporary opening of a historical location in Turin, Palazzo Chiablese. This location, a 16th Century baroque palace, hosts the former Royal Apartments of the Savoy family, opened to the public during the Italian "Week of the Culture" in April 2006. The visit took place in the apartments, which include five rooms: the Guardroom, the Room of the Valets, the Reception Room, the Dining Room, and the Room of the Tapestries. A map of the apartments can be found in Fig. 3. Visitors were free to stroll inside the apartments, encouraged by the fact that the rooms are not pipelined. As they moved from a room to another, the anthropomorphized spider Carletto followed them as observing them from a webcam, adapting his presentation to the topics that are relevant to the current location.

The following section provides an overview of the application, with an eye on the pipeline illustrated in the previous section (reported in Fig. 2). Off-line editing is described extensively in Sects. 3.2 and 3.3, while the architecture for real-time execution is described in Sect. 3.4.

3.1 Overview of the application

The peculiarities of the location have influenced the design of "Carletto the spider" in many ways (*Visit constraints* in Dramatour pipeline). A number of facilities usually

Fig. 4 "Carletto the spider"

involved in exhibitions were not available for this location. The apartments, in fact, are generally not open to the public. For this reason, they are not predisposed for the visit in any way and the accidental visitor is not given any help to identify the rooms and the object displayed. So a mobile application was an obvious solution: an existing wireless network allowed the localization of the visitor at the room level; however, once in a room, it was not possible to further identify the objects the visitor was standing by.

So, the design of "Carletto the spider" was inspired by two main goals. On the one side, as a mobile application, it was intended to bridge the lack of any visit infrastructure (labels, marked paths, etc.): the device (and the character on it) follows the visitor through her changes of location, and compensates for the missing facilities by providing the visitor with contextual information. On the other side, Carletto accounted for Waterfield's observation about the "touch of theatre" to be added to the visit of historical locations, by delivering a dramatized presentation aimed at giving the visitor the impression of being a guest in an inhabited place.

The presentation given by Carletto was written by a script writer (the *Author* in Dramatour pipeline), with the support of an expert in the historical and artistic aspects of the location (the *Domain expert*). Carletto is a teenage spider whose family has inhabited the palace from ages (Fig. 4). Carletto not only knows the history of the palace in detail, but also knows a lot of funny anecdotes about the people who have lived there through the centuries, and is striving to tell them to the visitors. From time to time, Carletto consults his web, which functions as an archive of his memories. The conflict between the role of a "guide" (or an "audio-guide"), who exposes facts orderly and plainly according to the topology of the location, and the desire to be a "landlord" of the palace who recounts all the trivia and the anecdotes he knows—most of which see him or his family personally involved—meets the requirement of centering the presentation on an internal conflict of the character to gain the emotional engagement of the visitors. Moreover, Carletto engages in an external conflict with the cleaners, who would like to kick him out of the place. After some time in a room, Carletto

becomes uneasy, and tries to induce the visitor to move to another room, claiming that the cleaners may trap him.

The interaction strategy[1] of Carletto is inspired by a well-known, recognizable model, "the guided visit", in which the character drives the interaction, though leaving some room to the initiative of the visitor. Carletto keeps the control of the interaction, politely directing the visitor's attention to relevant aspects of the environment and instructing her in a dramatized style. However, in order to respect the basics of human-computer interaction, the visitor can take control at any time, by changing location or by pausing or stopping the presentation.[2] At the same time, the metaphor of the guided visit is the means through which a relevant part of the drama, Carletto's inner conflict between guide and storyteller, is put on stage.

The application follows a location-adaptive, general-to-specific presentation pattern. The choice to adopt a simple interaction metaphor and a personalization strategy based on implicit input (i.e., localization) was mainly due to the analysis of the target visitors (from the *Application requirements* in the pipeline), who were expected to be mostly elderly people not very familiar with mobile technologies (the opening took place in office hours from Wednesday to Saturday).

The character of "Carletto the spider" was graphically designed and realized by a 3D production team, by following the specifications given by the author. The metadata used for tagging the presentation units (*Metadata vocabulary*) have been developed by the *Interaction designer* and the *Author*, with the help of a *Software engineer*, who predisposed an ad hoc interface for tagging. Metadata encode the topic and the communicative function of the presentation units, needed to structure the character's interaction, and their visual features. By following the requirement of an average duration of the visit of about 25 min, presentation units were produced for an overall duration of 50 min, in order to improve the variability of Carletto's presentation. Due to random aspects of the interaction design, and to the subjective variability of the visitors' behavior, this ratio was sufficient to ensure partially different presentations for all visitors. Finally, a software engineer implemented the software that realizes the real-time execution of the application (the *Dramatour manager* and the *Delivery Manager*).

During the implementation, a "Wizard of Oz" study was carried out to evaluate the cognitive load imposed by a so engaging interaction. This experiment allowed the author and the interaction designer to tune the duration of presentation units, the density and timing of content exposition, and the alternation of talk and silence. Finally, before the opening of Palazzo Chiablese, a set of pilot studies were run with real users simulating the visit in the real setting. Twenty-two subjects were involved in this preliminary tests, and from their visit experience and feedback, the interface was partially re-designed: the "play/resume" and the "stop" buttons were added to give more control on the system; the name of the current room was inserted both on the screen and in the rooms themselves to provide more feedback about the success rate

[1] We use this term for uniformity with the Dramatour methodology (Nunnari et al. 2008), although it would be more appropriate to speak of communicative strategy given the limited interaction of this application.

[2] If the visitor takes the initiative, however, the responsiveness of the system is constrained to the completion of the character's "dialog turn", corresponding to the duration of the current presentation unit.

of the localization process; finally, the quantity of provided contents per room was reduced.

3.2 Metadata framework

In "Carletto the spider", the presentation units are tagged with the information that the system uses to generate the behavior of the character. Metadata are divided into two sets: *communicative* metadata, i.e., the communicative function accomplished by the unit, and *topic* metadata, i.e., the description of the informative content of the unit, necessary to select the units in a contextually-aware, adaptive fashion.

3.2.1 Communicative metadata

Each presentation unit accomplishes a communicative function; units must be self-contained, i.e., each of them accomplishes by itself some communicative function. Communicative functions hinge on the conversational agents theories, generalizing the specific term of 'communicative acts' over the variety of verbal and non-verbal modalities trough which the acts can be performed (Cassel et al. 2000).

Although 'Carletto the spider' is not a dialogue system, we refer to dialogue and conversation studies as a source of inspiration to realize a colloquial style of presentation. Conversation is a highly cooperative, structured activity, and a relevant part of this structure is devoted to the management of the conversation itself (Schegloff and Sacks 1973; Goodwin 1981). In the field of dialogue systems, the actions undertaken to establish and maintain the "common ground" (Clark and Schaefer 1989) are explained as a part of the effort to achieve the communicative goals, within a model of dialogue as a rational, cooperative activity (Grosz and Sidner 1990; Cohen and Levesque 1990). More recently, the attempts to imitate human interaction patterns in embodied conversational agents have drawn new attention on the importance of non-verbal behavior in face-to-face conversation (Cassel et al. 2000; Pelachaud 2005), and to the role of social dialogue to build the empathy between the participants to a conversation (Bickmore and Cassel 2005).

In "Carletto the spider", communicative functions are organized in a bipartite taxonomy, that reflects the informative and interactional functions achieved by human communication:

- The *informative* function, i.e., the task of providing the visitor with useful and relevant information during the visit, is the primary task of the system. The presentation units that realize the informative function are further labeled according to their topic to allow for the adaptation to the location of the visitor.
- The *interactional function* encompasses all the non-informative functions of the interaction, namely the *social*, *directive* and *phatic* functions.
  - The *social* function corresponds to the performance of the conventional rituals according to which the interaction is structured. In order to gain believability, the character must perform some basic social behaviors, so this function includes more specific functions for opening the interaction (greeting the visitor, introducing himself and establishing the goal of giving information to the visitor) and

closing it (saying goodbye to the visitor, recommending the visitor to come back in the future).
- The *directive* function includes all the actions that the character performs in the attempt to affect the visitor's behavior in some way (Searle 1969). It includes specific functions for explicitly requiring the visitor to perform some action (like changing position when the location system is not able to detect the current location, or moving to another room for time reasons).
- The *phatic* function is activated when all other functions are unapplicable. Its purpose is to signal the character's presence and its willingness to interact with the visitor (Jakobson 1990). For example, it is activated when the character leaves some time to the user to stroll in a room without swamping her with too much information.

3.2.2 Topic metadata

The description of the informative content of the units, necessary to structure the presentation in a way that adapts to the behavior of the user, is encoded in *topic* metadata. The informative content is classified with respect to an ontological representation of the domain of the presentation, Palazzo Chiablese—described in Damiano et al. (2008). The organization of the presentation (see next section) relies on this representation to select the content to be conveyed to the visitor and structure it in a coherent and adaptive way (Geurts et al. 2003). In the field of cultural heritage, semantic annotation lends itself to the development of various types of applications, as exemplified by Aroyo et al. (2007) and van Ossenbruggen et al. (2007).

Since the location of Palazzo Chiablese is a complex domain, that can be described according to several semantic dimensions (like history, art or topology), the representation of the domain is subdivided into five specialized ontologies (see Fig. 5). The topic of the presentation units is described as a tuple of references to these ontologies (if a unit does not refer at all to a certain ontology, the reference to that ontology is set to the null value). In this way, topic metadata account for the fact that the same unit possibly concerns more than one topics, each described by a different ontology.

- The *topological ontology* describes the topology of location, centered on notions like rooms and room parts. According to the practice normally followed by human guides, topology provides a dimension of primary importance to organize the domain information.
- The *historical ontology* describes the historical facts related to the location. This ontology includes two main branches, describing, respectively the historical characters who lived in the palace and those who worked in it (further subdivided into professional roles, like painters, architects and craftsmen).
- The *ontology of objects* systematizes the variety of pieces of furniture and other items located in the apartments, most of which are awkwardly termed and unknown to standard visitors.
- The *chronological ontology* is an ontology of time intervals that serves the purpose of providing a temporal framework for locating the historical events.

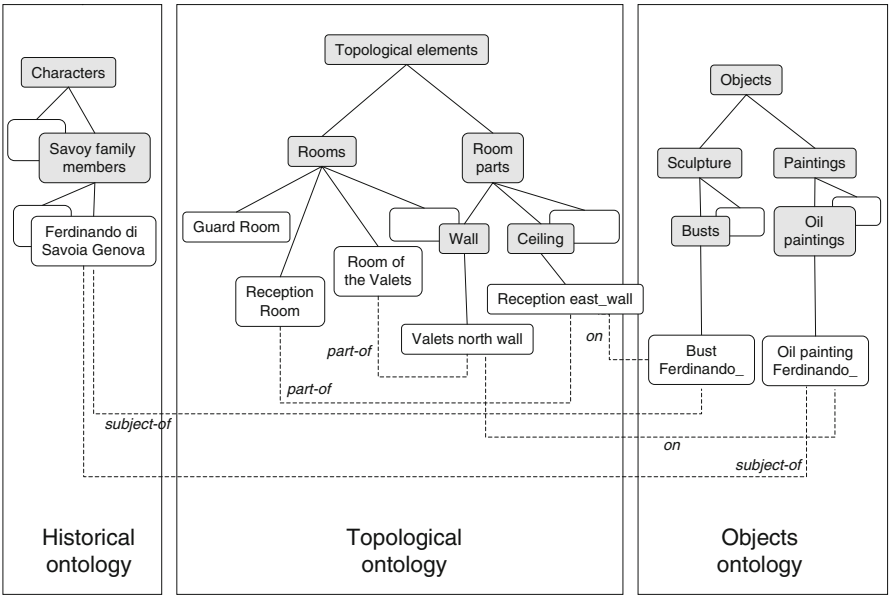

Fig. 5 A fragment of the domain ontology concerning Palazzo Chiablese, including portions of three specific ontologies (historical, topological and object ontologies). Grey boxes represent classes, white boxes represent instances. Subclass and instance relations are represented by solid lines, non-taxonomic relations are represented by dashed lines (for the sake of clarity, relations here are attached to the instances)

– The *symbolic ontology* describes the concepts (reigns, battles, marriages) that are celebrated by the art objects, located in the palace (paintings, statues, etc.).

In these ontologies, concepts are connected by subsumption relations in which each concept has only one ancestor, so all the ontologies are taxonomies. To simplify the description, the same concept cannot appear in more than one taxonomy. In order to represent the non-taxonomic relations among concepts, orthogonal relations (different than subsumption) have been added to the domain description to connect concepts within the same ontology or across different ontologies. For instance, the topological ontology contains a taxonomy of topological concepts (like rooms and room types) and the *part-of* relations according to which these entities are related, specifying, for instance, that a room contains a set of walls, a ceiling and a floor. As an example of relations spanning across different ontologies, consider the *subject of* a painting (object ontology), *located in* a room (topological ontology) and *painted by* an artist (historical ontology) to *celebrate* an event (symbolic ontology). The domain ontologies have been developed with the Protégé ontology editor; orthogonal relations have been implemented as Protégé slots (Gennari et al. 2003).

At every moment of the presentation—given the compliance with the visitor location—the selection of the topics is driven by only one of the ontologies, according to a predefined alternation encoded in the interaction strategy. In particular, the alternation between the topological and the historical ontology realizes Carletto's inner conflict between his official role of a guide and its personal desire to tell the visitor about the

interesting characters and episodes concerning the palace, as detailed out in Sect. 3.3. Tagging each unit with respect to all the ontologies allows designing strategies aimed at avoiding abrupt transitions when the ontologies are alternated.

Since "Carletto the spider" shows an adaptive behavior, the exact presentation order cannot be predicted in advance. So, each presentation unit is further tagged according to a *given/new* distinction of the discourse referents mentioned therein (Hajicova et al. 1998). Each unit is tagged with a tuple of referents marked as given or new according to how they are instantiated in the unit: a referent introduced by the unit is marked "new" ("The first owner of the Palace was Beatrice Langosco di Stroppiana"); a referent assumed as already introduced in some unit is marked "given" ("Lady Beatrice was involved in a few love affairs"). The system maintains an history of referents introduced. When a unit to be delivered contains a "given" referent, this must already be in the referents' history; viceversa, an item already in the referents' history cannot be introduced again by a unit in which it is marked as "new". Units containing unmarked items can be selected at any time without restrictions.

3.3 Interaction strategy

The interaction script—that encodes the strategy according to which the system selects the next communicative function to be achieved, thus realizing the communicative behavior of the character—is based on a set of specific protocols for addressing the different aspects of the interaction. Conversational protocols, inspired to the theory of dialogue games (Mann 1988), attempt to formalize conventional sequences of communicative acts representing stereotypical conversations—or portions of conversations—see Maudet and Chaib-draa (2002) for a survey of speech-act based versus protocol-based approaches to dialogue modelling.

In our application, each protocol is constituted by a sequence of communicative functions. For the sake of flexibility, a protocol does not specify which units will realize each communicative function, leaving the task of matching the functions with appropriate presentation units to a dedicated module, the Interaction Manager (see next section for the description of the system architecture). The activation of a protocol depends on the communicative context, given by the visitor input, the interaction history and the current location. From the point of view of the development of applications, the use of protocols responds to the requirements of modularity and abstraction, allowing the interaction designer to manipulate the system's communicative strategy as a set of abstract constructs during the phases of design and testing.

Figure 6 reports the interaction script. This script deals with all the special situations triggered by the user input, the wireless communication, and the termination conditions of the visit, leaving to specific protocols the task of dealing with these situations in a contextually appropriate way (including the presentation task, managed by a dedicated script). The visitor's actions of switching the PDA on or off (lines 2–3 and 4–5) have the highest priority and activate the social protocols for opening and closing the interaction (*opening protocol* and *closing protocol*, respectively). The termination of the visit (defined by the general visit requirements posed by the management of the historical site) has the immediately following priority, and results in the same effect of

```
1  PROCEDURE interaction_script (input, location, interaction history)
2     IF visitor switched on THEN
3        play opening_protocol
4     ELSEIF visitor switched off THEN
5       play closing_protocol
6     ELSEIF visit finished THEN
7       play closing_protocol
8     ELSEIF location is unknown THEN
9       play directive_protocol
10    ELSEIF delivery ratio is reached THEN
11      play waiting_protocol
12    ELSE call presentation_protocol (location, interaction history)
13    ENDIF
14 END PROCEDURE
```

Fig. 6 The interaction script. Implementation details, like the exact formulation of the delivery ratio, are explained in the text

activating the closing protocol (see the "visit finished" at line 6–7). If the location of the visitor is unknown, the system activates a *directive protocol* (lines 8–9), in which the character signals to the visitor that he cannot 'see' her anymore, and invites her to move in order to regain visibility (according to the webcam metaphor underlying the interaction with Carletto).

Finally, if the system detects that the visitor has remained in the current room for too long (this was measured with the amount of information delivered by Carletto about the current room—line 10–11 "delivery ratio is reached"), it enters a *waiting protocol* in which the character invites the visitor to a different room (e.g. polishes his medals and says impatiently "what about changing room?"), and starts waiting for a reaction (a sequence of three phatic units is played, unless the visitor has moved in the meantime). This behavior is motivated by two main reasons. On the one side, the need to keep some information about the room in case the visitor returns to it, is a circumstance that is permitted by the fact that the visitors can move freely across the apartments. On the other side, the attempts of Carletto to attract the visitor to a different room are always motivated by his fear of being trapped by the cleaners. In this way, the performance of a function (the directive one) dictated by the guided visit metaphor becomes, at the same time, the means of putting on stage a dramatic conflict of the character, namely the interpersonal conflict with the cleaners. The threshold that implements the notion of remaining in a room for too long is computed on the set of available presentation units for the room in the system's knowledge base—a set that becomes smaller every time the visitor returns to the room. So, at each return, the visitor will get a smaller number of presentation units about the room.

In all the special cases described so far, Carletto's behavior follows a fixed strategy. If none of these cases occurs, he switches to the location-adaptive *presentation protocol* (line 12–13 of the interaction protocol), synthesized in Fig. 7. First of all, the system checks whether the visitor has moved to a new room (line 2). If so, the system sets the ontology to the topological one (line 3) and sets the focus to the current room (the node of the ontology that corresponds to the room, line 4), so that a presentation unit about the room in general is selected (line 5). Notice that the expression 'select unit' stands for 'randomly select one unit from the pool of candidate units'.

```
1  PROCEDURE presentation_protocol (location, interaction history)
2     IF room has changed THEN
3        set current ontology to topological
4        set focus to room
5        select unit
6     ELSEIF focus set to room THEN
7        set focus to first room part
8        select unit
9     ELSEIF related unit on history exists THEN
10       set current ontology to historical
11       select unit
12       set current ontology to topological
13    ELSEIF delivery ratio of current room part is reached THEN
14       move focus to next room part
15    ELSE
16       select unit
17    ENDIF
18 ENDPROCEDURE
```

Fig. 7 The presentation protocol, sketched in pseudocode

On the contrary, if the visitor is in the same room as the previous selection cycle, the system either starts or continues the description of the content of the room. By doing so, it is driven by a topological principle, i.e., by the order according to which items are positioned with respect to the four walls of the room and the ceiling, according to the standard practice followed by human guides. The topological ontology provides an effective way to refer to the elements of the room: rooms, in fact, contain objects, furniture and decoration elements that are not common nowadays, and that are often termed awkwardly for the non-specialists. Describing the content of the rooms according to a topological order, i.e., wall by wall, makes the identification of the referred elements easier for the visitor, as confirmed by the results of the evaluation in Sect. 4. If the focus was previously set at the room level (line 6), the system moves the focus to some room part, according to the representation encoded in the topological ontology (line 8), and selects a unit accordingly (again, if several units are available, one is selected randomly).

When the focus is on room parts, the system attempts to realize an alternation between the topological ontology and the historical ontology. This alternation, as explained in Sect. 3.1, is a key aspect of the dramatization process, as it constitutes the means of letting the character's inner conflict emerge between his roles of guide and storyteller. Before switching to the historical ontology, the system checks if, at least one appropriate presentation unit is available (line 9, 'related unit on history exists'). If so, it switches to the historical ontology (line 10) and selects a unit before setting the ontology back to the topological one (lines 11–12), thus triggering a short historical detour. The switch depends on the availability of a unit that minimizes the transition between the two ontologies, in order to avoid abrupt focus shifts that would threaten the discourse coherence. Therefore the systems looks for the units that share the same topic as the current one in the historical and topological ontology. If no such units are available, the systems looks for the units whose topic is more general than the current one in the topological ontology, by consulting the part-of relations encoded in it. In case of several candidate units, the remaining domain ontologies (symbolic,

object and chronological, in this order) are exploited to rank the candidates (in case of equally ranked units, one is randomly selected). If no unit at all is available, the system skips the historical detour.

After the ontology alternation has been dealt with (meaning that the room has already been described in general and at least one room part has been commented on), the system checks whether there is any information left to say about the current room part (line 13); if the threshold on quantity has been reached, it moves the focus to the next room part (the next sibling of the wall node on the ontology, line 14). Notice that the described strategy complies with the focussing rules stated by Grosz and Sidner (1986) for task-oriented dialogue. According to Grosz and Sidner's rules, maintaining the focus on the current task has the priority over moving the focus to a subtask of the current task, that has the priority over moving the focus from the current subtask to a different subtask of the same task.[3] As at the room level, the quantity of information given for each room part is given by the ratio of the delivered units on the available units.

The described algorithm is adaptive since the system modifies its communicative behavior depending on the visitor location, thanks to the implicit feedback from the localization system, and on the duration of her stay in each room. This information is encoded in the interaction history, which contains a temporally indexed list of all the visit events (user input, location changes, unit delivery, etc.). When the visitor enters a room for the first time, she will receive general information about the room. As either the visitor spends a long period of time in a room, she will receive more detailed information; or if she comes back in the room, she will receive more detailed information as well. If the visitor remains in a room for a certain time, established by the procedural author, the character switches to role of a storyteller. This strategy, beside contributing to the dramatization, is normally employed by the human guides, as acknowledged also by other projects (Zimmermann and Lorenz 2008). In this way, thanks to the adaptation and to the random aspects incorporated in the system strategy, every visitor receives a personalized Carletto presentation, depending on the visit order, the time she spends in each room, and on how many times she comes back in each room.

In order to exemplify the interplay of the ontological framework and the interaction protocol, we resort to an example. Consider the fragment of the domain ontology represented in Fig. 5: the Duke Ferdinando di Savoia Genova (one of the Savoy family members), is the subject of two different artworks, an oil painting, situated on the north wall of the Room of the Valets, and a marble bust, situated on the east wall of the Reception Room.

Carletto is likely to talk about the Duke Ferdinando at different points in the presentation. When the system follows a topological order of presentation, the Duke Ferdinando may be mentioned when Carletto is talking about the locations in which the painting and the bust are situated. The Duke Ferdinando may subsequently become the subject of a historical digression centered on the Savoy family.

So, a plausible presentation flow may be constituted by the following sequence of presentation units (for brevity, null values are omitted from the topic representation),

[3] The actual implementation supports the focussing procedure on a taxonomy of any depth. However, here we report the algorithm only for the levels of the room and the room parts, since only these two were developed in the final version of the system.

assuming that the user location does not vary and that no user input is received by the system:

4 {PU_33 < Topology : Room_of_the_Valets >, }
{< History : Carlo_Felice >}
…
7 {PU_35 < Topology : Room_of_the_Valets_north_wall >,
< Objects : oil_painting_Ferdinando >,
< History : Ferdinando_di_Savoia >}
8 {PU_107 < Topology : Palazzo_Chiablese >,
< History : Ferdinando_di_Savoia,
< Chronology : 1850 >, < Symbols : Marriage > }

With PU_33, the system starts introducing the room of the Valets by giving general information about its function and the characters who gave it its current appearance (Carlo_Felice), then illustrates the room walls one by one (for brevity, we omit the units about the south and east walls, represented by the dots). After presenting the oil painting on the north wall of the room (PU_35), the system switches to the historical digression. So, Carletto starts talking about the Duke Ferdinando (PU_107), with a unit that talks about his marriage with an Austrian princess, that led some main renovations in the Palace.

Remember the system tries to minimize the transition between the ontologies by looking for a unit that not only has the same value (or a closely related value) as the previous one for the historical ontology, but also shares the most values concerning the remaining topic metadata: in this case, the closest match for the unit PU_35 is the unit PU_107, that has the same value for the historical ontology (see the item < History : Ferdinando_di_Savoia >); no unit being available that has the same value for the topological ontology (< Topology : Room_of_the_Valets >), PU_107 is selected, since its topic value for the topological ontology, 'Palazzo Chiablese' (< Topology : Palazzo_Chiablese >), is a more general concept than the concept of 'room'. Notice that, PU_107 is preferred to other units that talk about the Duke Ferdinando in relation with other rooms, since, here, the part—of relation that connects the current room with the Palace itself is considered more relevant than the relation that holds between the current room and the other rooms (although they belong to the same class).

Alternatively, the system may introduce the Duke Ferdinando when presenting the Reception room, which contains a bust that represents him. Or, he may be mentioned in a historical digression after Carletto talked about the Duke's marriage, since this event is explicitly mentioned in one of the units concerning the Reception Room.

3.4 System architecture

The system is structured according to a client-server schema. This schema is suitable for a number of possible deployments, provided that the notions of location awareness

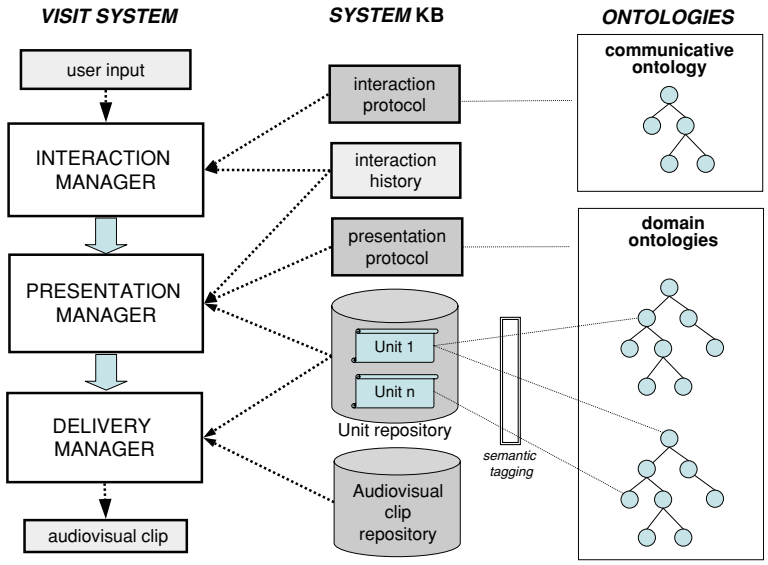

Fig. 8 The architecture of the Dramatour server. Solid lines represent control flow, dashed lines represent data flow

and content dimension (the information to be presented) are present. In fact, we have applied the same schema to on-site tours, with mobile devices connected to the server through a wireless network, and virtual tours, in which visitors explore a virtual replica of the location by using a point+click map interface on the web browser and connect to the server through the internet (Damiano et al. 2006a).

The Dramatour server executes a *decision-execution-sensing* loop: decision addresses the selection of the next presentation unit; execution concerns the delivery of the selected unit; sensing concerns the processing of new user input. The core of the server, which is responsible for the decision phase, has a modular structure: the *Interaction Manager* generates the presentation in terms of communicative functions by accounting for the visitor input; the *Presentation Manager* translates the communicative functions into presentation units; the *Delivery Manager* handles the delivery of the corresponding audiovisual units to the visitor in an audiovisual continuum. A sketch of the control and data flow among these modules is visible in Fig. 8.

The loop executed by the visit server is the following:

1. *Sensing* The Interaction Manager gathers the input from the visitor. The changes in the visitor location are actively monitored by the system (see Sect. 3.4.1) to identify the appropriate content to be delivered and the depth of presentation, which depends on the visitor location and on the duration of the permanence in that location.
   Some asynchronous inputs generated by the actions allowed by the interface, like pausing and restarting the presentation, do not affect the interaction strategy (they are only relevant for the client-server protocol), but require an immediate reaction. So, they cause the system to exit the loop to execute hard-wired behaviors;

2. *Decision* The Interaction Manager determines the next communicative function based on the interaction script, given the interaction history and the visitor input (localization and explicit input). The interaction script defines the succession of the communicative functions. The selection of a communicative function is accompanied by the specification of its topic (if the selected function is the informative one) or by the specification of its subtype (an optional specification that can be given for the other communicative functions).
   The selected communicative function is passed to the Presentation Manager. The Presentation Manager retrieves from the system's repository of presentation units a unit that matches the selected communicative function. If the selected function is the informative one, it consults the presentation script and the interaction history to set the topic in a way that complies with both the current context (timing, localization) and the dramatization process;
3. The Delivery Manager retrieves from the repository of audiovisual units the selected presentation unit. The Delivery Manager also takes care of the visual continuity of the clip with the previous one, by following the editing methodology described in (Lombardo et al. 2003);
4. *Execution* Finally, the selected unit is sent to the device for the delivery; when the commitment to play is confirmed, the interaction history storage is updated with the pair ⟨visitor input, selected presentation unit⟩, and this pair will contribute to the future decisions.

3.4.1 Implementation

The interaction between the client and the server follows a question-answer schema and is based on TCP connections. When the client contacts the server for the first time, the server releases an ID to identify the session, and creates a new interaction history. In the following interaction, the client provides its location and the server answers with a unit to be played. Built-in controls of the PDA allow the visitor to stop, pause and resume the presentation; so, they do not have any consequences in terms of the client server schema, with the exception of the stop command, that causes the interaction to be copied to the system log and destroyed on the server.

The PDA also perform the localization of the client, that provides the visitor current location, i.e., it identifies the room is which the visitor is currently situated. The localization system is based on measuring the signal strength of pre-installed 802.11 Wireless Access Points. We based our implementation on the technique used by Haeberlen et al. (2004), suitable to locate devices in macro-areas of indoor locations, such as rooms and short corridors. This approach, notwithstanding its limitations, has been already successfully used in many projects—see Borriello et al. (2005) for an overview.

The localization technique is based on two phases: an initial survey and the actual localization. To setup the survey, the target space is partitioned in areas. For each area, a survey was performed in order to record the list of visible access points and their signal strength variation. The survey of a room consisted in walking with the device throughout the whole room to collect scan data. We found that 60 scans were sufficient to calibrate the system. With a scan every 2 s, this turned out to 2 min walking for each

room. The collected data (average and variance of each Access Point) was stored on a datafile that was copied to all the devices employed in the visit.

During a visit, each client performs continuous scans of the visible Access Points. The actual location of the device is deduced by comparing the scan results with the statistics datafile. In choosing the scan ratio, it is necessary to consider that, even if a higher scan ratio yields to a faster localization, the communication of the device is interrupted during a scan process. We found 2 s to be a good compromise to obtain an acceptable localization deduction while leaving enough free band for data communication. As we know from Haeberlen et al. (2004), the probability to correctly deduce the location with one scan is very low (ca. 75%). In order to increase the probability of a right deduction, several scans must be performed in a row. In particular, with 5 scans the accuracy raises to 95%. In Haeberlen et al. (2004), the purpose is to minimize the probability to deduce an incorrect location. In our system, it is preferred to perform a faster localization by allowing an "unknown" location (addressed by the interaction script, see Sect. 3.3).

The implementation of the architecture is based on common hardware available on the consumer market and mostly on open-source software. The Server, that follows the specifications described in Sect. 3.4, is implemented in Java (http://java.sun.com), while the data base system is mySQL (http://www.mysql.com). The web-based authoring interface has been developed in PHP (http://www.php.net). The PDA is an ASUS A636 (PocketPC series). Most of the client software is written in Java and is executed by a virtual machine.[4] The localization system is based on a C library which obtains from the operating system the signal strength of each visible Access Point. The video clips are encoded and played back with the Macromedia Shockwave Flash (http://www.macromedia.com), with the resolution of the target PDA (320×240 pixels). The overall bitrate of a video, including sound, is about 400 Kbit/sec.

Since the Dramatour methodology is media-independent, beside the PDA-based version of the virtual tour, a web-based guided virtual tour of the same location has been developed by using the same visit server. The web-based application, described in Damiano et al. (2006a) was intended as a pilot test for cross-media implementation, and was published after the end of the opening of Palazzo Chiablese.

4 The visit and the evaluation settings

The system evaluation lasted three days during the opening, and consisted of a quantitative and qualitative analysis of a survey questionnaire given to the users at the end of the visit. The visit was free of charge and the visitors were welcomed by hostesses who gave them a headphone-equipped PDA, with written instructions advising that the system was a prototype under study, and some suggestions about the use of the guide (e.g. "You are free to stroll in the rooms", "Carletto is able to adapt its description depending on your position", "You can stop and resume the presentation by pressing the button illustrated below", and so on). Although instructions were likely to influ-

[4] The Mysaifu Java Virtual Machine for Windows Mobile PDAs (http://www2s.biglobe.ne.jp/~dat/java/project/jvm/index_en.html).

ence the behavior of the users somehow, thus interfering with the evaluation, this risk was traded off against the probability that they would take Carletto for a standard, non-adaptive guide, and limit their behavior accordingly, thus reducing the relevance of the evaluation.

After the visit, every user was asked to fill in a questionnaire of 17 questions. The topic areas proposed by the questionnaire were aimed at investigating the following aspects by means of ad hoc questions (summarized in brackets), which contained single and multiple choices, Likert scales,[5] and a free comment section:

- assumptions on the target visitors (age, level of education, acquaintance with media technology equipments);
- effectiveness of the communicative model (appropriateness of content quantity and quality, visitors' capability to orient themselves and to find items, request for more interactivity);
- visit experience
  - user acceptance of the drama-based approach (the personalities attributed to Carletto, the overall visit evaluation);
  - comparison with other more traditional ways of guiding museum visitors (appreciation for human and audio-guides);
  - global visit experience (ergonomics of the device, visual aspect of Carletto).

The questionnaires were autonomously filled by the users to avoid any possible interference; they were anonymous, voluntary, and introduced by a written presentation explaining the general research aim (collecting real and anonymous users data for a human-machine interaction research).

After having gathered all the questionnaires, we analyzed the collected data exploiting descriptive statistics, such as mean and standard deviation. Moreover, in order to find significant relationship between the variables under study, we also measured Pearson correlation, since scores showed a normal distribution, see Keppel et al. (1998) for details.

However, since statistical analysis can be misleading and too narrow (Dourish 2001), we carried out a qualitative study relying on insights from the users' observed behavior and reactions. More specifically, we carried out informal field observations and a qualitative analysis of the open answers collected by questionnaires (see Sect. 4.2).

4.1 Quantitative evaluation

All the 380 visitors compiled the questionnaire. However, we just considered 249 questionnaires since some were incomplete, and many users could not complete the visit because of technical problems.[6]

[5] A Likert scale is a type of survey question where users are asked to evaluate the level at which they agree or disagree with a given sentence (Sirkin 2005).

[6] During the visit, some users suffered from technical problems mainly caused by the maintenance of the devices and the wireless network. The first technical problem was due to the fact that 20 visits were

4.1.1 Evaluation regarding the a priori assumptions

During the system development, we made some a priori assumptions on target visitors in order to be guided in the design of the application. The main assumptions, related to a local survey on museum visitors,[7] were that most target visitors would probably be elderly, with medium/high education level, not particularly familiar with technology and PDA. These hypotheses were addressed by a set of questions regarding age, level of education, media technology equipments of the users that showed the following results, amply confirming our initial assumptions.

The visitors were 87 males and 161 females, with an age distribution ranging from younger than 15 years old to older than 64 years. In particular, the complete age distribution was: 20% >64, 21% 55–64, 15% 45–54, 10% 35–44, 22% 25–34, 10% 15–24, 2% <15.

Their level of education was high, as expected: 42% graduate, 44% high school, 12% junior high school, 2% primary. Concerning their media technology equipments, 90.36% declared to have a cellular phone, 67% a computer, 80% an Internet connection (home or work), and only 4.2% a palm. This distribution seems to reflect the commonly assumed diffusion of technology in Italy, as reported by the National Institute of Statistics (ISTAT).[8]

4.1.2 The effectiveness of the communicative model

The effectiveness of the presentation strategy incorporated in the behavior of the character was evaluated by analyzing the answers to a set of questions that directly addressed the presentation delivered by Carletto. The presentation strategy followed by Carletto is based on the visitor's location, more precisely on the room in which the visitor is currently situated. Given the current room, Carletto starts providing information according to a topological order, i.e., the items contained in a room are described according to their position with respect to the walls of the room (rather, for example, than by grouping them into categories; for the reasons of this choice, see Sects. 3.1 and 3.3). Since the system keeps the information concerning some items for the possibility that the visitor returns to the same room, some items were skipped on purpose by the presentation delivered by Carletto (according to a 80% delivery ratio).

The first two questions aimed at assessing the effectiveness of the communicative model regarded the quality and quantity of the content delivered. The analysis of results showed that visitors were quite satisfied with the *content quality* (average score

Footnote 6 continued
scheduled per half an hour, with the availability of 50 palm devices, and the palm battery sometimes took longer than half an hour to recharge, and some batteries were faulty. So, some user experienced incomplete or frequently interrupted visits, without the possibility of changing the device. The second source of problems concerned the wireless network of the building, which was usually used for other purposes (data exchange, Internet access).

[7] *Il Pubblico di Mostre e Musei in Piemonte*. Study from the Fitzcarraldo foundation, a local independent research center (http://www.fitzcarraldo.ithttp://www.fitzcarraldo.it), in Italian.

[8] Report available at http://www.istat.it/dati/catalogo/20071106_00/indice.pdf, in Italian.

of 3.80, with a Standard Deviation of 1.2), and with the *content quantity* (average score of 3.72 with a SD of 1.23).[9]

Due to the gross-grained localization at the room level, one of the challenges of the presentation was to address the ability of the visitors to find the described items (a relevant issue especially with elderly people): in fact, the terms used to indicate the objects and pieces of furniture contained in the rooms are often unknown or confusing for the visitor. So, a question specifically addressed the visitors' capability to orientate themselves and to find the items referred by Carletto. The results show that most visitors (83%) were able to orientate themselves and to find the items Carletto was talking about.

Finally, given the guided-tour metaphor, we tested the users' request for more interactivity during the presentation. As shown by the results, most of them would appreciate a more active interaction with the possibility of directly asking questions to Carletto (41%: yes, 39%: sometimes).

We have also correlated these values with the other variables under study. The evaluation of the conversational model did not show any significant correlation with sex, age and media technology equipments, but:

- *the ability to orientate oneself in the location and to find items* has a very significant positive correlation with the *overall visit evaluation* ($r = 0.332$, significant at the 0.01 level, 2-tailed), the preference for *audio-guides* ($r = 0.214$, significant at the 0.01 level, 2-tailed) and *the use of PC* ($r = 0.166$, significant at the 0.01 level, 2-tailed). The visitors who had a good experience of the visit were also satisfied with the presentation strategy. These visitors seems to be familiar with technology, since they appreciate audio-guide and they are PC users.
- *the request for more interactivity* has a significant negative correlation with the judgement about the content quantity ($r = -0.151$, significant at the 0.05 level, 2-tailed) and a positive significant correlation with *the use of PC* ($r = 0.142$, significant at the 0.05 level, 2-tailed): visitors who seem to be less satisfied with the visit would like not only to receive more contents, but also to have more interactivity. These visitors frequently use the PC.

4.1.3 Visit experience

Acceptance of the drama-based approach: In order to investigate the general acceptance of the drama-based approach, we directly ask the visitors to rate their visit experience: the average score for the *overall visit evaluation* is 3.75 with a standard deviation (SD) of 1.25.

In addition, we were also interested in how visitors perceived the dramatized presentation delivered by the virtual character through the assignment of a personality to Carletto (one of the tenets of the drama paradigm). Therefore, we designed a multi-choice question to ask visitors which personality they would associate with Carletto. The aim was to discover what visitors thought about Carletto's presentation, depending on the personality they associate to him. The choices provided to the visitor

[9] The possible scores were ranging in a Likert scale from 1 to 5.

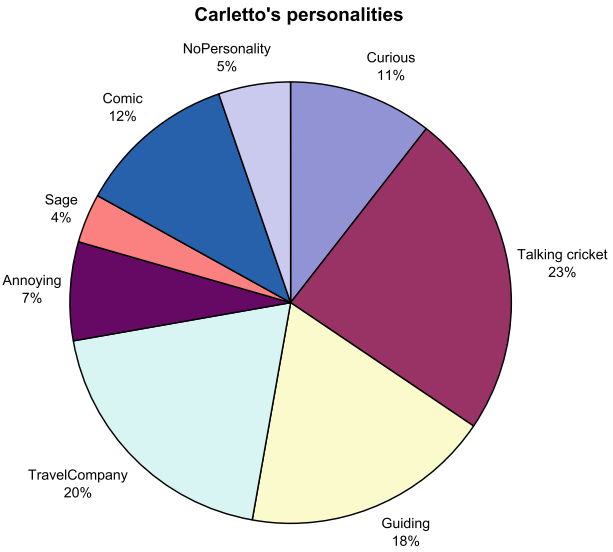

Fig. 9 The personalities of "Carletto the spider"

corresponded not only to the intended features of Carletto's personality, but also to the personalities that were at risk of being communicated by the behavior of the character. Every personality has a distinctive feature associated, thus the analysis of the visitors' choices is an indirect way of discovering their impression about the key features of the overall experience, as well as their positive or negative judgements. Visitors could explicitly choose to assign Carletto "no personality"; this choice was added to evaluate whether the dramatization paradigm was successful or not.

Figure 9 shows the results.[10] The mostly associated personality is the "*talking cricket*" (23%), the character of Collodi's "Pinocchio". The talking cricket is a little lively animal, who typically speaks too much, but it is also the metaphor of "the voice of the visitor's conscience", which suggests visitor what to explore next. From time to time, in fact, Carletto tries to influence the visitor's behavior by asking her to move to another room, as a way to enforce his role of a guide along the dramatization. Both the "*travel mate*" (20%) and the "*guide*" (18%) have received many preferences. These results show that the visitors perceived the character in a friendly way, and they are inclined to follow its advices, as demonstrated by the field observation described below. These opinions are always associated to positive judgements.

It is interesting to examine the significant correlations found between the "guide" and the "travel mate" and other variables. These correlations implicitly show that the "talking cricket", which was intended to match the visitors' perception of Carletto as a funny character (sometimes annoying, but lovely)—was not connoted as positively as the other two personalities:

[10] The results have been obtained by weighting the choices when the visitor gave more than one answer.

– *the visit evaluation and the guide character* ($r = 0.189$, significant at the 0.01 level, 2-tailed); when the appreciation for Carletto increases, visitors tend to identify him as a "guide";
– *the visit evaluation and the travel mate* ($r = 0.220$, significant at the 0.01 level, 2-tailed); when the appreciation for Carletto increases, visitors are likely to identify him as a "travel mate", in a positive and friendly way;
– *the age and the travel mate* ($r = -0.158$, significant at the 0.05 level, 2-tailed). When the age of visitors decreases, they are more inclined to identify him as a "travel mate", in a friendly perspective.

Fewer ratings were associated to the *comic character* (12%), the *prying one* (11%), the *annoying one* (7%), and the *sage* (4%). They were always associated to negative judgements. We found significant correlations between:

– *the visit evaluation and the annoying character* ($r = -0.219$, significant at the 0.01 level, 2-tailed). Quite obviously, visitors who dislike Carletto perceive him as annoying;
– *age and the prying character* ($r = -0.223$, significant at the 0.01 level, 2-tailed). Visitors are inclined to judge Carletto like a prying character, as long as their age decreases, thus young visitor find him "offensively curious or inquisitive", in a negative way. As shown by the free comments answers, they did not appreciate his gossip about the Savoy royal family;
– *education and the sage character* ($r = -0.284$, significant at the 0.01 level, 2-tailed). When the education level increases, visitors tend to not consider Carletto as a "sage", as well as they are not satisfied about the contents he conveys.

Comparison with other guides: The comparison with other, more traditional, ways of guiding visitors was evaluated by exploiting a set of comparative questions.

We compare the evaluation of the virtual guide with two more established approaches, namely, the human guide and the audio-guide. The choice to indicate these two paradigms was motivated by the fact that they both constituted—to different degrees—a source of inspiration for the presentation strategy of Carletto.

The visitors under study generally appreciate both human (average rate 3.8, range: 1–5) and audio-guides (average rate 3.5, range: 1–5). We found a very significant positive correlation between those who like audio-guides and the visit evaluation ($r = 0.486$, significant at the 0.01 level, 2-tailed); the appreciation of audio-guides has a very significant negative correlation with age ($r = -0.191$, significant at the 0.01 level, 2-tailed). Hence, visitors who appreciate audio-guides have also a positive attitude toward Carletto. It is possible that these visitors do not like a social dimension in the visit, and they are likely to be alone in experiencing their visit. Moreover, as the age level decreases, their preference toward such devices tends to increase.

Global visit experience: The user experience about the visit was evaluated by asking questions about the ergonomics of the device, the visual aspect of Carletto, and the easiness of use.

Concerning specifically the ergonomics of the device, 82% of the visitors found the device easy to handle, with significant positive correlation with the use of the

mobile phone ($r = 0.152$, significant at the 0.05 level, 2-tailed). Therefore, keeping the PDA from its shorter side was not problematic for visitors. However, after the field observations, see Sect. 4.2.1, we concluded that it would be a safer solution to put a cord on it, in order to let the visitors hang up the device to their neck.

The average score concerning the *visual aspect* of Carletto is 3.7 with a SD of 1.3, range 1–5; the average score concerning the *easiness of use* is 4.03 with a SD of 1.23, range: 1–5. Carletto's visual aspect was well accepted by visitors. This result was not obvious for a zoomorphic character, especially because Carletto recognizably belongs to a type of animal, the spider, that is held to be repugnant by most people. However, correlations showed that education and easiness of use have a significant negative correlation ($r = -0.192$, significant at the 0.05 level, 2-tailed). Thus, visitors with a high education level are critical about the interaction modality proposed by Carletto.

4.2 Qualitative evaluation

4.2.1 Field observations

As introduced in Sect. 4, we carried out some informal field observations, wherein we observed the behavior of the visitors by strolling in the site amidst them like stewards and hostesses that moved along rooms to take care of them (so observed visitors did not feel under study). Often visitors asked us for help or questions, that we tried to translate into insights about their experience with Carletto. The most interesting observations that emerged are:

- at the beginning, visitors tend to be on their own, isolated, and the use of the mobile device limits their ability to listen and to speak. After a while, some of them regain the group, and discuss among them about the presentation; in particular, they try to match what they see on each other's screens;
- the use of the device makes the visitors less inclined to observe the location and the objects around them;
- some visitors are completely directed by Carletto, waiting for his advice about what to see and where to go;
- sometimes, older visitors address Carletto with spontaneous comments; some very young visitors seem to think that Carletto is hiding somewhere in the ceilings.

These findings enrich the results described in the previous section. An interesting point is that visitors do not like to be on their own during the visit, thus the virtual guide should be designed to be used by more than one user.[11] The second observation suggests that the presentation must not capture the visitors' attention in an all-involving way: visitors need to be encouraged at looking around, instead of focusing on the device. Finally, the last two points somehow underline that some users immerse themselves in the dramatization process, since they tend to assign a place to the virtual character in the real world. This can be interpreted as an example of the Laurel's *interface as mimesis* (Laurel 1993).

[11] Only the 13% of visitors visit a museum alone. See *Il Pubblico di Mostre e Musei in Piemonte*: http://www.fitzcarraldo.it.

4.2.2 Grounded theory

The questionnaire contained a free comment section for the question "Some observation?". Of all the 249 visitors, 114 filled in the question. The qualitative analysis of visitors' free comments was inspired by the Grounded Theory, a well defined research methodology that emphasizes the generation of theory from data, in an inductive process of data analysis (Strauss and Corbin 1990). According to the Grounded Theory methodology, collected data may be qualitative or quantitative or a combination of both types, since an interplay between qualitative and quantitative methods is advocated. For this reason, we found very useful to integrate the free comments study with the analysis of quantitative data and of the field observations described before, in order to discover possible interconnections and make them emerge during the phases of the methodology.

In particular, three phases of data analysis are involved in the Grounded Theory methodology: *open coding*, the analytical process through which concepts are identified starting from the analysis of the collected material, and the properties and dimensions of these concepts are discovered; *axial coding*, the process of relating categories to their subcategories, termed "axial" because coding occurs around the axis of one category, linking categories at the level of properties and dimensions; *selective coding*, the process of integrating and refining the theory.

We started by examining the open answers, which were broken down into discrete parts. Then we closely examined data and compared them for similarities and differences, and we started to accumulate concepts. At the same time, we started the inductive process of the investigation and definition of main categories, subcategories and variables involved in the phenomenon under study. The main categories identified in the study were *positive evaluations without suggestions* (e.g. "cool", "I love it"), *negative evaluations without suggestions* (e.g. "ridiculous", "I hate it"), *reported technical problems*, *suggestions*. After having analyzed the material related to the main categories, we concentrated our analysis on the *suggestions*, which were the most interesting for the purposes of our evaluation, namely assessing the acceptance of the proposed approach and collecting feedback for future re-design. The following subcategories of suggestions emerged from the inductive analysis (presented with some example of their characterizing properties and comparison with the related results obtained during the quantitative analysis of data):

- *Information*: the need for more information, the possibility to go into more depth about the topic presented and to receive more information, etc. This subcategory is strongly related to the results obtained in the question about the quality and the quantity of contents proposed by the presentational strategy, see Sect. 4.1;
- *Interactivity*: the request for more interactivity, the possibility to choose the quantity of delivered information, the possibility to step back, the possibility to manually indicate the location, etc. This subcategory is strongly related to the results obtained in the question concerning the user request for more interactivity, see Sect. 4.1;
- *Context*: the need for contextual information, the display of the items Carletto is talking about, the highlight of contextual information, etc. These findings show that,

for some visitors, the presentation strategy should realize a more direct link to the surrounding environment;
– *Adaptation*: the adaptation of the guide to the age or the needs of the visitors ("Carletto is for children" or "Carletto is a child" express the need for a more serious guide). This subcategory highlights the fact that the dramatized presentation should be adapted to different categories of visitors.

Finally, we put together these findings with quantitative data and the analysis of the field observations, and we started relating categories to their subcategories along with their properties and dimensions to form more precise and complete explanations, until an acceptable level of saturation was reached. We have also tried to investigate conditions (e.g. when a visitor needs more interactivity, which kind of visitors need more interactivity), action and interactions (e.g, in which situation visitors experience the lack of interactivity) and their consequences by interplaying qualitative and quantitative data. The findings helped us to refine the core theory of the selective coding: visitors have highlighted the need for some kind of more *personalized interaction*, because they would like to be able to choose the modality and the type of presentation, the quality and the quantity of delivered information, the character itself, its personality, etc. Therefore, as discussed below, our future re-design will be aimed at taking into account these needs by boosting the adaptation features and adding user modeling to the system.

4.3 Summary of results

The visitors under study seem to reflect the general distribution of visitors in Piedmont,[12] the Italian region where Torino is located: the mostly represented age range is comprised between 45 and 64 years old; the majority of visitors are women; about the 40% of visitors has a degree, and a similar percentage has a high school diploma. Thus we could affirm that the sample of the survey seems to be quite representative of the museum population in our territory, under a socio-demographic point of view. However, it has been hard to find some distinctive socio-demographic features on the basis of which to model the visitor, confirming what pointed out by previous work (Petrelli et al. 1999a).

Significative correlations show that visitors with high education level are more critical about the experience with Carletto, in particular with the content quantity and the dramatized presentation. Visitors with high education level are not inclined to assign the personality of a "sage" to Carletto. Young visitors who appreciate Carletto are likely to identify him as a "travel mate", while young visitors who do not appreciate Carletto do not like his extreme curiosity. In general, young visitors are likely to assign a positive evaluation to a device-driven museum visit, like the audio-guide. Concerning media technology equipment, visitors who use a PC are likely to have a good visit experience in term of orientation; but, also, the visitors who demand more interactivity are likely to use the PC. Unfortunately, the visitors who wrote productive suggestions (e.g. many of them have underlined the need for a more serious guide pre-

[12] "*Il Pubblico di Mostre e Musei in Piemonte*", Fitzcarraldo foundation.

senting information in a more detached and objective way) did not exhibit distinctive features that let us model users.

As it can be noticed, it is very difficult to create stereotypical models of visitors according to these spurious correlations. In order to be more accurate in these findings, we also ran a K-means cluster analysis (Huang 1998) to identify relatively homogeneous groups of cases based on sex, age, education, media technology equipments, and the visit experience values. Only two clusters emerged:

– *Cluster 1* is characterized by the following features: 30 year old females, high school education level, average use of the web; they own a PC and a cell phone, they really liked the visit (the average visit experience value is 4) and sometimes they would like to ask Carletto something; they really like human and audio-guides;
– *Cluster 2* is characterized by the following features: 61 year old females, high school education level, unfrequent use of the web; they possess a PC and a cell phone, they really like the visit (the medium visit experience value is 4) and sometimes they would like to ask Carletto something; they really like human museum guides and like audio-guide;

These two clusters cover the population under study only partially, and in general the population of museum visitors. So probably, as other studies have already highlighted (Zancanaro et al. 2007), museum visitors who share similar visit habits can hardly be identified by similar socio and psycho-demographic features. On the basis of our analysis, we can hypothesize some coarse classifications, useful at introducing a more adaptive interaction and delivery of contents. Excluding visitors who like Carletto as it is, possible distinctive features are: visitors who like to be guided, and visitors who prefer to move freely in the museum; visitors who prefer to be in a group, and visitors who prefer to be alone; visitors who want a more controllable interface; visitors who want the possibility to decide the quantity and quality of information delivered; visitors who would like to interact with a more serious and detached character.

5 Comparison with related work and discussion

The Dramatour methodology for dramatized information presentation and the application "Carletto the spider" for character-based guidance to a historical site share many design issues with a number of projects active in the applied research on information presentation and cultural heritage access. We can single out three main aspects: multimodality, character-based mediation, and adaptivity.

Multimodality characterizes most of the applications developed for the field of cultural heritage (Callaway et al. 2005; Solon et al. 2004; Stock and Zancanaro 2005). Our approach does not push forward the notion of multimodality, although it includes multimodal input, with implicit and explicit input forms (Kobsa et al. 2001), and output, with sound, graphical visualizations introduced by Carletto, and textual feedback on the location detection. From the point of view of multimodality, important limitations of our application—as revealed by the evaluation—are the impossibility of requesting contents that interrupt the presentation flow, and the impossibility of having a social experience during the visit, which is fundamental for the people who

visit a site in group. In the applications for cultural heritage exploitation, it is worth noting the efforts toward the improvement of the cohesion of the social experience (Aoki et al. 2002), a very difficult task when dealing with "personal" devices.

The use of an anthropomorphic design helps people in establishing a relationship with virtual characters, and enhances the presentation of information in most cases. Visitors' assignment of a personality to Carletto supports such a claim. Improved sound and graphic processing have provided virtual characters with realistic facial expressions and lip-synch speech that display emotions for pedagogical aims (Leon and Fisher 2006). On a different and complementary perspective are the computer graphics animated characters (Maestri 1996, Clark 2002): they are carefully modeled and rendered but their interactivity is typically constrained within the limited boundaries of computer games. The character of Carletto is posited at the junction of embodied conversational agents and computer graphics characters. The communicative behavior of the character is not generated in real time, according to the methodology established by Cassel et al. (2000), but encoded in larger units. The animation task has been solved by identifying a number of elementary expressions that were composed to form the presentation units that contain the behavior of the character. This choice, mainly determined by time and budget limitations, is motivated by the empirical finding, reported by Bickmore and Cassel (2005), that the expectations set by an embodied agents are hardly met by the limitations of the state-of-the-art techniques for synchronizing multimodal behavior, thus possibly resulting in the frustration of the user.

In the field of conversational agents, the conversational agent *Max*, developed by Kopp et al. (2005), provides interesting elements for comparison. Max is a fully-fledged conversational agent who acts as a tourist guide, with the ability to respond to the verbal inputs of visitors. Its behavior, differently from Carletto, is not encoded in pre-built units, but is generated on the fly by an architecture that includes a hierarchical planner for determining the dialogue behavior, and specialized modules to provide a multimodal realization of the generated behavior. However, the design goals of Max are very different from our application. Max is not a mobile application and is not adaptive—its interaction metaphor is more of a receptionist than a guide; so the kind of relationship it builds with the user differs from the type of personal involvement that our application aims at.

Research topics related to Carletto also include location-aware adaptive guides and user-adaptive museum guides. Concerning the former one, we cite GUIDE (Cheverst et al. 2000), a tourist guide developed for the city of Lancaster in Great Britain. GUIDE is adaptive with respect to the location of the user, her walking speed, the places already visited, the time of the day, and the language and interests of the user. The latter information is acquired by a registration form and is not updated or refined during the interaction. The user interface of GUIDE exploits a browser metaphor enriched by a character with a friendly and polite personality. Differently from our system, the character is not intrusive and gives information in a textual format, since GUIDE is designed for outdoor pedestrian use. Lol@ (Pospischil et al. 2002) is a guide for the city of Vienna. Similarly to "Carletto the spider", Lol@ is based on a lightweight adaptation approach, but is adaptive also toward the device.

Regarding user-adaptive museum guides, user modeling has been used in different projects that aimed at producing more refined adaptation strategies. HyperAudio

(Petrelli et al. 1999b) is based on a static user model, dependent on an initial questionnaire; HIPS (Bianchi and Zancanaro 1999) and Museum Wearable (Sparacino 2002), contain a dynamic user model, that evolves as the visitor proceeds in the museum by leaving a trace of her physical movements. In HIPS, the physical movements within the museum were used to classify visitors. This dynamic classification utilized a non-intrusive user modeling approach, wherein the museum acted as an interface.

In the PEACH (personal experience with active cultural heritage) project (Kuflik et al. 2005), visitors were provided with a personalized experience made possible by user modeling and user adaptation. The user model gathers information about the visitor and guides the adaptation of information presented to the user. The main features of the user modeling in this project are that (i) the user modeling is "non-intrusive"; hence visitors are not required to provide any personal information and user modeling is based solely on users' behavior and on an enjoyment feedback; (ii) the user modeling component can support different applications, with different requirements.

The CHIP project (Aroyo et al. 2007) is focalized on recommender technologies and user modeling in the specific field of museums and cultural heritage. The authors exploit techniques from semantic web and user modeling to recommend guided tours. The user model is built by asking visitors to rate artworks on a five-stars scale, while using a knowledge base semantically enriched with metadata and vocabularies. At the moment the CHIP project has the Rijksmuseum Amsterdam as hosting partner.

Finally concerning user-adaptive guides in the cultural heritage domain, we mention iCITY (Carmagnola et al. 2008). iCITY is an adaptive, social, multi-device recommender guide that provides information about the cultural resources and events promoting the cultural heritage in the city of Torino. The most interesting feauture of the iCITY approach is that the user model also considers the tagging activity of the user. Thus the approach stresses social tagging as a new and powerful kind of feedback and as a way to infer knowledge about users.

All the mentioned user-adaptive guides differ from our application, mainly for the presence of user modeling component able to generate more precise adaptation. Considering the broad division between content-based recommendations, when the user is recommended on the basis of her own past preferences, and collaborative recommendations, when the user is recommended on the basis of the past preferences of other people with similar tastes (Adomavicius and Tuzhilin 2005), Dramatour only takes into account the current location of the visitor and the content already delivered, with the goal of not repeating the same presentation units in the presentation and tuning the quantity of contents to the visit advancement.

In our application, since visits were booked by phone, visitors could not be extensively interviewed to acquire information for the personalization. It is well known that adaptable systems requiring an additional user effort can cause the "paradox of the active user" (Carroll and Rosson 1987). For example, users often refuse to visit sites that impose to respond to an interview first because they would save time getting their immediate task done. In the case of Carletto, any user effort before the visit (i.e., filling in a form on the PDA, or to completing a questionnaire on the web site *before* the visit) would have been done without further advantages for the visitor: the adaption is performed *una tantum*, since Palazzo Chiablese is not normally open to the public. Asking visitors to fill form or complete questionnaires is not always the

appropriate strategy, as in the case of our system. As notice by Petrelli et al. (1999a), a relative dislike for technology suggests reducing the explicit interaction to the very minimum, possibly even eliminating it. At the same time, the use of a mobile device and the real time aspects of the application would not allow the use of transparency techniques—described by Cramer et al. (2008) to compensate for the efforts of the user.

The adaptive audio-augmented guide for art exhibitions developed within the LISTEN project (Zimmermann and Lorenz 2008) is, under many aspects, similar our application, although, being an audio guide, does not include multimodal output. Similarly to our application, the LISTEN project is inspired by the aims of creating an aesthetic experience for the user and of incorporating the point of view of the curators in the visit experience, thus mediating between the exhibits and the user in an effective way. The LISTEN project shares with ours some of the instruments exploited to create such mediation, like the use of a metadata framework that mixes topological information (the relation of sound entities with the artworks situated in the room) and content description, and the location-based adaptation. Differently from our application, LISTEN includes a user modeling component, that relies on a rich notion of visit context. This component makes the system adaptive not only to the user location, but also to her visit style, grasped by features like speed or motion style; this type of adaptation is made possible by a fine-grained localization of the user in a more circumscribed environment (a single large room).

Although the drama-based approach features a number of peculiarities, it is possible to update the system by incorporating notions of personalization and adaptation. Here we sketch some proposals for future re-design.

- In order to deliver a presentation in a style that is more appropriate for the visitor, the system may propose several characters (e.g. a comic and a more serious one, at least), and let the visitors choose, at the beginning of the visit, the most appropriate guide. Every guide should introduce her/himself to the visitor in a video clip, and showing her/his character and her/his visit style. The guide performs the self-introduction in a dramatized way, reflecting the peculiar features of her/his behavior. So, the visitor would not requested to specify a list of attributes she would like for her visit experience, but may choose the guide that best matches her visit style.
- Each guide may show an adaptive behavior depending on the visit style exhibited by the visitor, in the style of Bianchi and Zancanaro (1999) and Kuflik et al. (2005). When the visitor remains longer in a room, the system may select clips containing more detailed contents, presented in a more detached way, e.g. without dramatization. This would satisfy one requirement that emerged from the data analysis, i.e., that visitors who want more contents prefer a more serious and detached modality in the fruition of contents.
- A number of visitors reported they would like to ask Carletto for something specific, so to improve interactivity. In order to allow visitors to make questions, but at the same time to leave the control to Carletto and simplify the input recognition (free vocal messages could be difficult to recognize), we propose that, when speaking of some topic, Carletto may show a poster that concerns some specific item in the room; this poster may remain active for a given number of seconds, and eventually

either skipped or retrieved. When the visitor clicks on a particular item, Carletto focuses his presentation onto that item.

6 Conclusions

In this paper we have presented the Dramatour methodology for dramatized presentations in a location-based character-enacted modality. The application "Carletto the spider", based on the Dramatour methodology, has been developed for the visit guidance in a historical site, implemented for mobile devices. The paper has addressed the annotation of knowledge provided by a "procedural" scriptwriter and the design of a communicative strategy compliant with the dramatic features. In this application the adaptation of the presentation is based on implicit user input, given by the changes of the visitor's location, the duration of the stay in each location, and the history of the presentation. We have also carried on an evaluation of the system in a open-to-the-public installation with almost 400 visitors, and we have proposed a number of extensions for increasing the adaptation and personalization issues according to the requirements that emerged from the evaluation.

References

Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. IEEE Trans. Knowl. Data Eng. **17**(6), 734–749 (2005)

Aoki, P.M., Grinter, R.E., Hurst, A., Szymanski, M.H., Thornton, J.D., Woodruff, A.: Sotto voce: exploring the interplay of conversation and mobile audio spaces. In: Proceedings of the SIECHI Conference on Human Factors in Computing Systems, pp. 431–438. ACM, USA (2002)

Aroyo, L., Brussee, R., Gorgels, P., Rutledge, L., Stash, N., Wang, Y.: Personalized museum experience: the rijksmuseum use case. In: Trant, J., Bearman, D. (eds.) Proceedings of Museums on the Web 2007: Selected Papers from an International Conference. San Francisco, CA, USA. http://www.archimuse.com/mw2007/papers/aroyo/aroyo.html (2007). Accessed May, 2008

Bianchi, A., Zancanaro, M.: Tracking users' movements in an artistic physical space. In: Proceedings of the i3 Annual Conference, European network for intelligent information interfaces, pp. 103–106 (1999)

Bickmore, T., Cassel, J.: Social dialogue with embodied conversational agents. In: Advances in Natural Multimodal Dialogue Systems, vol. 30, pp. 23–54. Springer, Netherlands (2005)

Borriello, G., Chalmers, M., LaMarca, A., Nixon, P.: Delivering real-world ubiquitous location systems. Commun. ACM **48**(3), 36–41 (2005)

Bruner, J.: The narrative construction of reality. Crit. Inq. **18**(1), 1–21 (1991)

Callaway, C., Not, E., Novello, A., Rocchi, C., Stock, O., Zancanaro, M.: Automatic cinematography and multilingual NLG for generating video documentaries. Artif. Intell. **165**, 57–89 (2005)

Carmagnola, F., Cena, F., Console, L., Cortassa, O., Gena, C., Goy, A., Torre, I., Toso, A., Vernero, F.: Tag-based user models for social multi-device adaptive guides (2008, this issue)

Carroll, J., Rosson, M.: The paradox of the active user. In: Carrol, J. (ed.) Interfacing Thought: Cognitive Aspects of Human-Computer Interaction, pp. 80–111. MIT Press: Cambridge (1987)

Cassel, J., Bickmore, T., Campbell, L., Vilhjamlmsson, H., Yan, H.: Human conversation as a system framework: designing embodied converational agents. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (eds.) Embodied Conversational Agents, pp. 29–63. The MIT Press: Cambridge (2000)

Cheverst, K., Davies, N., Mitchell, K., Smith, P.: Providing tailored (context-aware) information to city visitors. In: Brusilovsky, P., Stock, O., Strapparava, C. (eds.) AH. Lecture Notes in Computer Science, vol. 1892, pp. 73–85 (2000)

Clark, K.: Inspired 3D Character Animation. Premier Press (2002)

Clark, H.H., Schaefer, E.F.: Contributing to discourse. Cogn. Sci. **13**, 259–294 (1989)

Cohen, P.R., Levesque, H.J.: Rational interaction as the basis for communication. In: Cohen, P.R., Morgan, J., Pollack, M.E. (eds.) Intentions in communication, pp. 221–255. MIT Press (1990)
Cramer, H., Evers, V., Ramlal, S., van Someren, M., Wielinga, B., Rutledge, L., Stash, N., Aroyo, L.: The effects of transparency on trust in and acceptance of a content-based art recommender (2008, this issue)
Damiano, R., Galia, C., Lombardo, V.: Virtual tours across different media in DramaTour project. In: Proceedings of the First European Workshop on Intelligent Technologies for Cultural Heritage Exploitation (Ecai 2006). Riva del Garda, Italy, pp. 31–35 (2006a)
Damiano, R., Lombardo, V., Pizzo, A., Nunnari, F.: Dramatization meets narrative presentations. In: Brewka, G., Coradeschi, S., Perini, A., Traverso, P. (eds.) ECAI. Riva Del Garda, Italy, pp. 31–35 (2006b)
Damiano, R., Lombardo, V., Nunnari, F., Pizzo, A.: Ontological domain coding for cultural heritage mediation. In: Lesmo, L., Borgo, S. (eds.) Proceedings of FOMI 08. Frontiers in Artificial Intelligence and Applications, vol. 174, pp. 88–99. Turin, Italy (2008)
Dourish, P.: Where the action is: The Foundations of Embodied Interaction. MIT Press, Cambridge (2001)
Dybkjær, H., Dybkjær, L.: Design and first tests of a chatter. In: André, E., Dybkjær, L., Minker, W., Heisterkamp, P. (eds.) ADS. Lecture Notes in Computer Science, vol. 3068, pp. 166–177 (2004)
Esslin, M.: 1988 The field of drama. Methuen: London (1987)
Gennari, J., Musen, M., Fergerson, R., Grosso, W., Crubézy, M., Eriksson, H., Noy, N., Tu, S.: The evolution of Protégé: an environment for knowledge-based systems development. Int. J. Hum.-Comput. Stud. **58**(1), 89–123 (2003)
Gershon, N., Page, W.: What storytelling can do for information visualization. Commun. ACM **44**(8), 31–37 (2001)
Geurts, J., Bocconi, S., van Ossenbruggen, J., Hardman, L.: Towards ontology-driven discourse: from semantic graphs to multimedia presentations. In: Fensel, D., Sycara, K.P., Mylopoulos, J. (eds.) International Semantic Web Conference. Lecture Notes in Computer Science, vol. 2870, pp. 597–612 (2003)
Goodwin, C.: Conversational organization: Interaction between Speakers and Hearers. Academic Press, New York (1981)
Grosz, B.J., Sidner, C.L.: Attention, intentions, and the structure of discourse. Comput. Linguist. **12**, 175–204 (1986)
Grosz, B.J., Sidner, C.L.: Plans for discourse. In: Cohen, P.R., Morgan, J., Pollack, M.E. (eds.) Intentions in Communication, pp. 417–443. MIT Press (1990)
Gustafson, J., Lindberg, N., Lundeberg, M.: The August spoken dialogue system. In: Proceedings of Eurospeech'99, pp. 1211–1214. Budapest, Hungary (1999)
Haeberlen, A., Flannery, E., Ladd, A.M., Rudys, A., Wallach, D.S., Kavraki, L.E.: Practical robust localization over large-scale 802.11 wireless networks. In: MobiCom '04: Proceedings of the 10th Annual International Conference on Mobile Computing and Networking, pp. 70–84. New York, NY, USA (2004)
Hajicova, E., Partee, B.H., Sgall, P.: Topic-focus articulation, tripartite structures, and semantic content. Kluwer, Dordrecht (1998)
Hardman, L.: Canonical processes of media production. In: MHC '05: Proceedings of the ACM Workshop on Multimedia for Human Communication, pp. 1–6. New York, NY, USA (2005)
Hoffmann, P., Herczeg, M.: Distributed storytelling for narrative in spacious areas. In: Goebel, S., Braun, N., Spierling, U., Dechau, J., Diener, H. (eds.) Technologies for interactive digital storytelling and entertainment 2003, pp. 346–350. Frauenhofer: Stuttgart (2003)
Huang, Z.: Extensions to the K-means algorithm for clustering large datasets with categorical values. In: Data Mining and Knowledge Discovery, vol. 2, pp. 283–304 (1998)
Ieronutti, L., Chittaro, L.: Employing virtual humans for education and training in X3D/VRML worlds. Comput. Educ. **49**(1), 93–109 (2007)
Jakobson, R.: The speech event and the functions of language. In: Waugh, L., Monville-Burston, M. (eds.) On Language. Harvard University Press (1990)
Johnson, B.: Place-based storytelling tools: a new look at Monticello. In: Trant, J., Bearman, D. (eds.) Museums and the Web 2005: Proceedings. http://www.archimuse.com/mw2005/papers/johnsonB/johnsonB.html (2004). Accessed May, 2008
Keppel, G., Saufley, W.H., Tokunaga, H.: Introduction to design and analysis, a student's handbook. W. H. Freeman and Co. (1998)

Kobsa, A., Koenemann, J., Pohl, W.: Personalized hypermedia presentation techniques for improving online customer relationships. Knowl. Eng. Rev. **16**(2), 111–155 (2001)
Kopp, S., Gesellensetter, L., Krämer, N.C., Wachsmuth, I.: A conversational agent as museum guide—design and evaluation of a real-world application. In: Panayiotopoulos, T., Gratch, J., Aylett, R., Ballin, D., Olivier, P., Rist, T. (eds.) IVA. Lecture Notes in Computer Science, vol. 3661, pp. 329–343 (2005)
Kuflik, T., Callaway, C.B., Goren-Bar, D., Rocchi, C., Stock, O., Zancanaro, M.: Non-intrusive user modeling for a multimedia museum visitors guide system. In: Ardissono, L., Brna, P., Mitrovic, A. (eds.) User Modeling. Lecture Notes in Computer Science, vol. 3538, pp. 236–240 (2005)
Laurel, B.: Computer as Theater. Addison Wesley Longman, Reading, MA (1993)
Leon, J., Fisher, M.: The use of virtual characters to generate teachable moments. In: Trant, J., Bearman, D. (eds.) Museums and the Web 2006: Proceedings, Toronto. http://www.archimuse.com/mw2006/papers/leon/leon.html (2006). Accessed May, 2008
Lester, J.C., Towns, S.G., Callaway, C.B., Voerman, J.L., FitzGerald, P.J.: Deictic and emotive communication in animated pedagogical agents. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (eds.) Embodied Conversational Agents, pp. 123–154. The MIT Press: Cambridge (2000)
Lombardo, V., Biral, F., Damiano, R., Pizzo, A.: Cyrano goes to hollywood: a drama-based metaphor for information presentation. In: Krueger, A. (ed.) AIMS 03 Workshop (Artificial Intelligence in Mobile Systems), pp. 17–24. Seattle (2003)
Maestri, G.: Digital Character Animation. New Riders (1996)
Mann, W.C.: Dialogue games. Conventions of human interaction. Argumentation **2**, 511–532 (1988)
Mateas, M., Senger, P.: Narrative intelligence: an introduction to the NI symposium. In: Mateas, M., Senger, P. (eds.) Working notes of the Narrative Intelligence Symposium. http://www-2.cs.cmu.edu/afs/cs.cmu.edu/user/michaelm/www/nidocs/MateasSengers.pdf (1999). Accessed May, 2008
Maudet, N., Chaib-draa, B.: Trends in agent communication language. Comput. Intell. **18**(2), 89–101 (2002)
McKee, R.: Story. Harper Collins, New York (1997)
Murray, J.: Hamlet on the holodeck. The Future of Narrative in Cyberspace. The MIT Press, Cambridge, MA (1998)
Nunnari, F., Lombardo, V., Damiano, R., Pizzo, A., Gena, C.: The canonical processes of a dramtized approach to information presenation. In: Multimedia Systems. Special Issue on the Canonical Processes of Multimedia Production. (2008, Forthcoming)
Pechenizkiy, M., Calders, T.: A framework for guiding the museum tours personalization. In: Aroyo, L., Kuflik, T., Stock, O., Zancanaro, M. (eds.) Proceeding of PATCH Workshop on Personalized Access to Cultural Heritage. Corfu, Greece. http://www.iit.demokritos.gr/um2007/UM2007_WS3_PATCH.pdf (2007). Accessed May, 2008
Pelachaud, C.: Multimodal expressive embodied conversational agents. In: ACM Multimedia, pp. 683–689 (2005)
Petrelli, D., Angeli, A.D., Convertino, G.: A user-centered approach to user modeling. In: UM '99: Proceedings of the Seventh International Conference on User Modeling, pp. 255–264. Secaucus, NJ, USA (1999a)
Petrelli, D., Not, E., Sarini, M., Stock, O., Strapparava, C., Zancanaro, M.: HyperAudio: location-awareness + adaptivity. In: CHI '99: CHI '99 Extended Abstracts on Human Factors in Computing Systems, pp. 21–22. New York, NY, USA (1999b)
Picard, R.: Affective Computing. MIT Press (1997)
Pospischil, G., Umlauft, M., Michlmayr, E.: Designing LoL@, a mobile tourist guide for UMTS'. In: Paternò, F. (ed.) Mobile HCI. Lecture Notes in Computer Science, vol. 2411, pp. 140–154 (2002)
Proctor, N.: Off base or On target? Wireless and location-aware applications in the museum. In: Proceedings of International Cultural Heritage Informatics Meeting (ICHIM 05). Paris. http://ichim05.ichim.org/ichim05/contenu/PDF/S09-Proctor.pdf (2005). Accessed May, 2008
Raptis, D., Tselios, N., Avouris, N.: Context-based design of mobile applications for museums: a survey of existing practices. In: Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services Table of Contents, pp. 153–160. Salzburg, Austria (2005)
Schegloff, E.A., Sacks, H.: Opening up closings. Semiotica **4**(7), 289–327 (1973)
Searle, J.: Speech acts: an essay in the Philosophy of Language. Cambridge University Press, Cambridge (UK) (1969)
Seger, L.: Creating Unforgettable Characters. Henry Holt and Company, New York (1990)
Sirkin, M.: Statistics for the social sciences. Sage Publications (2005)

Solon, A., McKevitt, P., Curran, K.: Mobile multiModal presentation. In: Schulzrinne, H., Dimitrova, N., Sasse, M.A., Moon, S.B., Lienhart, R. (eds.) ACM Multimedia, pp. 440–443 (2004)
Sparacino, F.: The museum wearable: real-time sensor-driven understanding of visitors' interests for personalized visually-augmented museum experiences. In: Bearman, D., Trant, J., Sparacino, F. (eds.) Proceedings of Museums and the Web 2002, pp. 17–20. (2002)
Springer, J., Kajder, S., Brazas, J.B.: Digital storytelling at the national gallery of arts. In: Trant, J., Bearman, D. (eds.) Proceedings of Museums and the Web 2004. http://www.archimuse.com/mw2004/papers/springer/springer.html (2004). Accessed May, 2008
Stock, O., Zancanaro, M.: Multimodal intelligent information presentation. Springer (2005)
Strauss, A., Corbin, J.: Basics of qualitative research: grounded theory procedures and techniques. Sage Publications, Newbury Park, Calif. (1990)
van Ossenbruggen, J., Amin, A., Hardman, L., Hildebrand, M., van Assem, M., Omelayenko, B., Schreiber, G., Tordai, A., de Boer, V., Wielinga, B., Wielemaker, J., de Niet, M., Taekema, J., van Orsouw, M., Teesing, A.: Searching and annotating virtual heritage collections with semantic-web techniques. In: Trant, J., Bearman, D. (eds.) Museums on the web. San Francisco, CA. http://www.archimuse.com/mw2007/abstracts/prg_325000959.html (2007). Accessed May, 2008
Waterfield, G.: The modern visitor and the historical palace: is an understanding possible? In: Abitare la storia. Turin, Italy. Conference Papers (1999)
Zancanaro, M., Kuflik, T., Boger, Z., Goren-Bar, D., Goldwasser, D.: Analyzing museum visitors' behavior patterns. In: Conati, C., McCoy, K.F., Paliouras, G. (eds.) User modeling. Lecture Notes in Computer Science, vol. 4511, pp. 238–246 (2007)
Zimmermann, A., Lorenz, A.: LISTEN: a user-adaptive audio-augmented museum guide (2008, this issue)

Authors' vitae

Rossana Damiano is a Researcher at the Department of Computer Science of the University of Torino, where she works in the area of Natural Language Processing and Intelligent Agents, and Assistant Professor at the School of Multimedia and Arts, University of Torino. She obtained her PhD in Computer Science in 2002 from the University of Torino, with a dissertation on replanning techniques for normative agents. She has been active in computational models of communication in multi-agent systems, with a specific focus on dialogue pragmatics. In recent years, she has started a research line on foundational and practical aspects of artificial drama.

Cristina Gena is Researcher at the Department of Computer Science of the University of Torino, working in the area of intelligent user interfaces, and Assistant Professor at the School of Multimedia and Arts, University of Torino. She completed her PhD in Communication Science (University of Torino) in 2003, with a thesis on the evaluation of user-adaptive systems. Her current research activities address user modeling, adaptive web systems and their evaluation, context-aware systems, semantic web, web 2.0, usability and interaction design.

Vincenzo Lombardo graduated in Computer Science at the University of Turin (1987), and received his doctorate in Computer Science in the Turin-Milan University Consortium (1993). He heads the Centre for Research on Multimedia and Audiovisuals of the University of Torino and teaches at the School of Multimedia and Arts (www.multidams.unito.it), which he co-funded. He also leads applied research projects and training programmes at the Virtual Reality and Multi Media Park (www.vrmmp.it). His conducts multidisciplinary research on the computational modelling of cognitive processes, with reference to language, music and reasoning, multimodal environments for human-machine interaction, and applications of artificial intelligence in formalizing artistic expression.

Fabrizio Nunnari is researcher and developer at the ASA-Lab (Art and Science Allies Laboratory) of the Virtual Reality and Multi Media Park in Turin, Italy. He studied Computer Science at the University of Turin where he received a degree in 2001 and a PhD in 2005. His main interests lie in the fields of Virtual Reality and Multimedia Systems, where he investigates software architectures and languages for data visualization.

Antonio Pizzo is an Associate Professor and teaches Procedural Drama at University of Turin, Facoltà di Scienze della Formazione, and is one of founders of Officine Sintetiche, a work lab on performance and

new media in Turin. He got a degree in Modern Literature at University of Naples “L’Orientale” in 1991. In 1997, he attained the PhD in Modern and Contemporary Theatre in the conjoined programme between the universities of Naples and Salerno. He has published books on contemporary theatre and drama. Recently he has been studying the new area of interactive drama and artificial characters.