Virtual tours across different media in DramaTour project

Rossana Damiano, Cristina Galia, Vincenzo Lombardo [1]

Abstract.

In recent times, human-computer interaction and information presentation have evolved towards sophisticated approaches that involve multi-modal aspects and character-based mediation.

This paper presents the application of the DramaTour methodology to the design and implementation of a guided tour for an historical site across different media.

The Dramatour methodology is based on a dramatization of the exposition in two respects. On the one side, the author plots a character's monologue that exhibits an engaging inner conflict; on the other side, the system architecture dynamically assembles the elementary units of the plot scripted by the author by implementing a tension between contrasting presentation styles.

1 Introduction

This paper proposes the application of the DramaTour methodology to the design and implementation of a guided tour for an historical site across different media. The DramaTour methodology [3] merges the use of storytelling techniques with the technology of virtual characters to create interactive dramatized presentations.

Characters are inspired to paradigm of the embodied conversational agents (ECA) paradigm [2] and the whole structure of the information presentation follows the interactive storytelling paradigm [11]. These two paradigms have significantly contributed, in the last decade, to the spreading of AI techniques in practical approaches to human-computer interaction (see [6, 15, 13, 14]).

The DramaTour methodology applies to the design of information presentation systems. Its working assumption is that information is presented through a dramatic narration operated by virtual characters, who act in first person and share the user's present time and space, yielding a powerful effect of physical and emotional presence. Interactional and communicative strategies of the characters are explicitly driven by the notion of *drama*: the presentation delivered to the user features the inner tension and the sense of direction that are typical of dramatic performance. According to Esslin's notion of *dramatic media* [5], user engagement increases when the system performance is conveyed through actions on an audiovisual display, resulting in greater effectiveness of content reception by the user [10].

In order to yield an effective computational model from the methodology and test the effectiveness of the approach, we limit the characters to one dramatic "storyteller", thus generating a dramatic monologue, and implement the generation of the character behavior through the real–time assemblage of pre–defined audiovisual behavior units that respond to user input. Dramatization applies to both the production of dramatic elementary units (from their writing to their interpretation by the virtual character through animation) and the editing operated by the system in delivering the content to the user. The applicative domain in which we are currently testing this methodology consists of guided tours in an historical site, during which the visitor is accompanied by the dramatic character. In this paper, we describe both a PDA-based version of the application, and a web-based version. The server that supports the visit, in fact, is designed to comply with different media: in this way, visitors who experience both modalities of visit are not required to become familiar with two different presentation styles, apart from the intrinsic differences implied by the web-based virtual tour with respect to the actual tour in the physical space.

The structure of the paper is the following. In Section 2, we describe the DramaTour methodology as a set of constraints on the system design concerning both system architecture and data definition for building practical applications. Then, we introduce the notion of dramatization within the methodological framework (Section 3) and present the example application (Section 4) developed for PDA and for the web (Section 5). Conclusions end the paper.

2 The DramaTour methodology

The DramaTour methodology defines a conceptual framework for organizing the behavior units of the system, and a system architecture that assembles these units in real time in response to user's inputs.

Behavior units are categorized through *meta-data*, that serve the function of identifying their content and their interactional and visual properties for real-time, contextual editing. *Topic* meta-data are used by to structure the information presentation according to one or more narrative lines, *interactional* meta-data account for engaging the interaction with the user, *editing* meta-data account for the audiovisual properties of the content that are relevant for the delivery.

The system architecture has a modular structure: the handling of the interaction with the user is mapped onto the *interaction manager*; the content organization is mapped onto the *presentation manager*; the ultimate delivery to the user in a well-edited, audiovisual continuum is handled by the *delivery manager*.

2.1 System architecture

The input to the system is given by the interaction history and the user input. The system is reactive, i.e., it responds to the user input by displaying an appropriate social and communicative behavior. The system executes a *decision-execution-sensing* loop. Decision addresses the selection of the next behavior unit; execution concerns the delivery of the unit; sensing concerns the processing of the user input.

[1] Dipartimento di Informatica, Università di Torino and CIRMA, Turin, Italy

Meta-data sets	Type	Values	
TOPIC	Tuple of ontological references	$\{<O_1,r_1>,...,<O_n,r_n>\}$	
	Tuple of flags $f \in \{given, new, null\}$	$\{<O_1,f_1>,...,<O_n,f_n>\}$	
COMMUNICATIVE	Communicative function	Presentational	
		Interactional	*Social*
			Directive
			Phatic
EDITING	Audio-visual features	Shot type	
		Objects	
		Angle	

Figure 1. The sets of meta-data in the DramaTour methodology

The system views the presentation as the realization of some *communicative function* [9]. Although the main goal of the system is presentational, the artificial character who delivers the presentation must account for some basic interactional functions in order to qualify itself as a believable agent and thus promote the user engagement.

The *informative* function is the primary task of the system, i.e., providing the user with useful and contextually relevant information during the visit. The execution of this function is assigned to the presentation manager, which handles the selection and the organization of the conveyed content through a sequence of behavior units. This module is responsible for realizing different presentation styles, each following a different narrative line, in order to obtain an alternation of styles that enforces the principle of dramatization (see next section).

Interactional functions, handled by the interaction manager, are divided into *social*, *directive* and *phatic*. Since the virtual character should qualify itself as a social agent, the system must perform some basic *social* behavior units that implement basic social protocols, like greeting the user at the beginning of the visit, or introducing itself. These behaviors have priority over the informative function (e.g. the character must introduce itself before providing any information); in many cases, they must take place at specific times along the interaction.

The *directive* function includes all the actions that the character performs in the attempt to force the user behavior in some way, like signalling conditions that may require the user to perform some action (for example, executing maintenance actions on the device on which the presentation is run, when prescribed by the visiting protocol in which the guided tour is embedded). In general, directive functions have no priority over basic social aspects (in order to enforce the notion of personification and autonomy of the virtual character) but have priority over the informative function.

The *phatic* function is activated when no other functions are applicable. Its purpose consists of signalling to the user that the character is active and willing to receive input. For example, it may be activated when the character has requested the user to perform an action of any kind - necessary for the prosecution of the interaction, and has not received any input after a given time interval.

2.2 Content organization

Behavior units constitute the knowledge base of the system. They contain multimedia content (an audiovisual clip with 3D animation and sound) and are tagged with the meta-information that the system uses to generate the interactional and presentational behavior of the character (summarized in Figure 1).

- *Communicative meta-data* describe the communicative function accomplished by the unit. Each behavior unit accomplishes one of the four communicative functions described above (informative, social interaction, directional, and phatic). It is up to the author to make sure that at least one unit matches each of the communicative functions acknowledged by the system. Moreover, each unit realizes only one communicative function.
- *Topic meta-data* contain the description of the informative content of the units. The informative content is classified with respect to an ontological representation of the domain that constitutes the domain of the presentation. Topic description is necessary for the presentation module to shape a coherent selection and exposition of the content of the presentation along one or more narrative lines.
  The presentation module relies on the ontological representation of the domain information to select the content to be conveyed to the user and to structure it in a coherent way. The presentation module follows a general strategy inspired by the focussing rules stated by Grosz and Sidner [7]. Since Grosz and Sidner's focussing heuristics have been elaborated for task-related discourse, in this methodology they have been adapted to the presentation of a set of domain facts, by mapping task decomposition relations onto sub-topic relations.
  Following sub-topic relations in an ontology according to the focussing heuristics corresponds to structuring the presentation along a certain dimension of meaning, corresponding to a narrative line. Since the ontology is hierarchical, the focussing heuristics determine a depth-first visit of the ontology.
  In principle, several meaning dimensions may be proposed to structure the same domain, corresponding to different presentation modalities. For example, the facts about an historical site may be "described" according to a topological dimension or "narrated" following a chronological dimension. In order to enforce the dramatization principle incorporated in the methodology, the author must encode domain knowledge according to at least two different ontologies. This requirement serves the function of establishing a dialectic conflict between presentation modalities. presentation, determining the active narrative line.
  Since the presentation order cannot be predicted in advance, each behavior unit whose communicative function is tagged as presentational is further tagged according to a *given/new* distinction [8]. When a new referent is introduced in the discourse, the methodology prescribes that a unit in which it is marked as new is selected, and enforces the principle according to which an item already in the interaction history cannot be introduced again by a unit in which it is marked as new. The methodology admits that a null value is employed in the given/new distinction: items marked in this way can be selected at any time without any restriction.
- *Editing meta-data* contain the information needed for assembling the audiovisual clip with the adjacent ones by interposing an audiovisual segment (called a *Transition Unit*) between them, to obtain visual fluency [1]. The system incorporates a set of *editing rules*, that implement a number of editing techniques, e.g., graphic qualities (including framing, mise-en-scene, etc.) and spatial continuity. Transition units, like behavior units, are selected by the delivery manager from a repository according to the editing rules.

The system assumes that behavior units are self-contained, i.e. that each of them accomplishes the execution of some specific communicative function, or, if its function is tagged as presentational, that it conveys a unit of meaning according at least one of the ontologi-

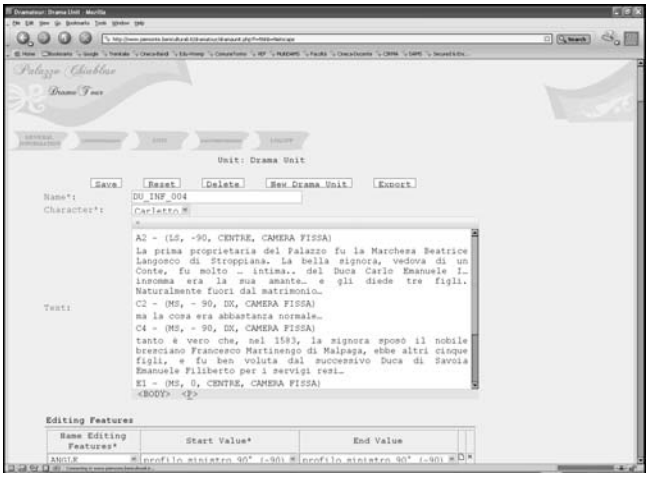

Figure 2. Web-based authoring interface

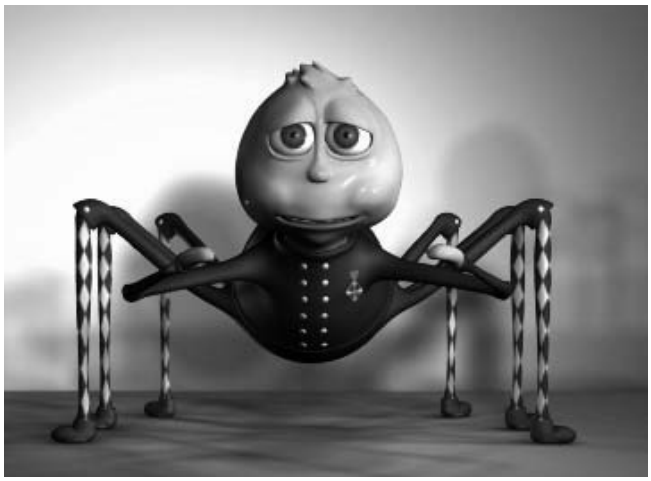

Figure 3. Carletto the spider.

cal representations. Clearly, the coordination between the system designer and the procedural author who develops the data of an application is the key to consistent scriptwriting with the system design. Moreover, it is up to the author to individuate and dramatize the character through the use of scriptwriting techniques in the authoring of behavior units.

In order to assist the authoring task, a web-based authoring interface has been created to enter the application data (behavior units and transition units) and to define the meta-data according to which they are classified by the system (Figure 2).

3 Dramatic issues put at work

In line with the notion of drama formalized in [4], we see drama as the combination of two main features: the fact that drama displays action at present time and the fact that it enacts a relevant conflict related to an emotional-dramatic value concerning the characters. Drama moves toward the solution of this conflict, yielding the typical impression of movement, and does it through a sequence of elementary units, called beats [12]. Beats are pure actional units formed by a action-reaction pair.

The principle of first-person, present-time action must be enforced by the authoring of the behavior units. The character's behavior, in fact, is not synthesized from an abstract specification of the character, its personality, its will. So, the methodology poses some constraints on the form of data: the data encoded by the author must bear an explicit description of their informative content and of their interactional function, that the system can rely on to sequence the units according to a consistent communicative and interactional strategy.

The dialectics between different presentation modalities substantiates a dramatic conflict. The emergent behavior of the presentation system should resemble as much as possible to a carefully authored monologue, in which an internal conflict of the character is exposed to the emotional response of the audience. For example, in the application described below (a guide to a historical site), describing objects and narrating stories about the site are put in a dialectical opposition, in which the descriptive task of the character, as the visit progresses, leaves the way to an historical line of narration, thus realizing a shift of the character from "guide" to "storyteller".

The interaction with the user regulates the dialectic alternation between presentation styles. The methodology specifies a meta-theatrical schema according to which user input along the interaction is interpreted as a positive or negative clue of user engagement and determines the presentational behavior of the character.

When a new application is designed, the author and the system designer must identify the exact nature of the dramatic conflict in the presentation domain, and define the interactional strategies according to which the system tries to account for the user reactions – conveyed through the specific modalities of the application framework. Going back to the museum guide example, by moving to a different location, the user may implicitly communicate interest or lack of interest for the presentation. If the user likes the current presentation style, the character maintains it until the related ontology has been completely explored (or explored to a sufficient degree), then switches to a different ontology.

4 A Dramatour application: Carletto, the virtual guide

The application of the methodology presented here has been tested in the historical location of a former residence of the Savoy family in Turin, Palazzo Chiablese. The application consists of an interactive guided tour on a mobile device enacted by a teenage spider, "Carletto", whose family has inhabited the palace from ages. Carletto not only knows the history of the palace in detail, but 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. The application is run on a mobile device, and the user input consist of pen pointing on the graphical interface and localization through the use of wireless infrastructure.

The conflict between the role of an "audioguide", who exposes facts orderly and plainly according to the topology of the location, and the desire to recount all the trivia and the anecdotes he knows from an historical perspective - most of which see him or his family personally involved - meets the methodology guideline of centering the presentation on an internal conflict of the character to gain the attentional and the emotional engagement of the users. Following the author guideline according to which the character itself must be carefully dramatized in the scripting units, Carletto engages in a continuous fight with the janitors, who would like to kick him out of the

place.

The visit is structured along a topological dimension, according to which the palace is modeled as a set of rooms. The system starts by enforcing the greetings by Carletto, i.e., it plays audiovisual clips where Carletto welcomes the visitors and introduces himself. Then, Carletto starts talking about the room in which the visitor is currently situated. Each time the user enters a room, Carletto starts (or resumes) the presentation of the objects (furniture, artworks) in the room. The system proceeds by delivering about the 80% of the material related to each room.

The idea of delivering the 80% of the whole material about each item along a certain ontological dimension pervades the whole architecture. For each subtopic of a room, the system delivers 80% of the material about the subtopic, then delivers one unit that has a direct connection with the last unit in the historical dimension, thus performing a short historical digression. For each return to a room, the system delivers 50% of the remaining material to be delivered: in this way, it sees to it that something relevant is left to be said about the room.

When all the subtopics of a room have been explored (with the rule of the 80% of the material) and the visitor remains in the same room for some time, we consider two cases. In case the system is still able to localize the visitor, it delivers some directional clips, that advise the visitor to proceed to some other room; in case of missing localization, Carletto switches to a historic presentation style for a while, i.e. the system selects and delivers units following the historical dimension, starting from the top of the historical ontology, then gently starts inviting the user to regain the connection by moving. In both cases, if the user does not seem to react to the system, Carletto finally activates the phatic function, by playing funny games and gazing to the user from time to time.

The system exits when the visitor has walked through more than half of the rooms of the visits and is near the exit door.

5 Example applications for different media

The visit can be performed either on site, using the PDA device on which Carletto delivers its presentation, or in a virtual way, through the web site. By mapping the topological organization of the guided tour onto the elements of the web interface, it is possible to exploit the same server for the two media. The web interface simulates the tour in the virtual space, by proposing to the user a virtual representation of each room accompanied by the sequence of clips in which Carletto provides information about the room. The visitor, in the virtual tour as well as in the actual visit, is free to decide the order in which he/she wants to visit the rooms. However, it is important to stress that the presentation delivered to the user may vary even if the user follows the same order of visit, thanks to the non-deterministic aspects incorporated in the architecture of the system.

An important advantage of this approach is that the user who performs the two modalities of visit experiences the same presentation style, so that the web-based virtual tour can be either a preparation for the actual tour or a prosecution of it. In fact, the two applications, beside sharing the same server, also share the same data. Data share is made possible by the fact that the same player is supported by both the web client (any http browser) and by the operating system that runs on the PDA.

From the author's point of view, not only the same methodology is enforced by both applications - with an advantage for scriptwriting and meta-data editing - but, since data are reused, the amount of work required by setting up a version of the tour for a different media than the PDA is significantly reduced.

Figure 4. The PDA for the guided tour.

The visit server, that follows the specifications described in Section 2.1, is implemented in Java (http://java.sun.com), while the data base system is mySQL (http://www.mysql.com/). The web-based authoring interface used to insert and manage clips information has been developed in PHP (http://www.php.net/).

A different client has been developed for the PDA-based and the web-based application. However, the client-server protocol is the same, the only difference being the adaptation of the client to the specific platform provided by the PDA and the web browser.

5.1 The virtual guide on PDA

The current implementation is based on common hardware available on the consumer market and mostly on open-source software. The client is currently running on an ASUS A636 PDA (PocketPC series) and the video clips are encoded into Macromedia Shockwave Flash (http://www.macromedia.com). The client (written in Java) pilots the media player by sending text-based commands to localhost.

The PDA also supports a localization client, that provides the user's current location, i.e., it identifies the room is which the user is currently situated.

The graphical interface of the PDA interface is very simple, since it includes only the window of the player, in which clips are shown to the visitor (see Figure 4). In the bottom part of the screen, the visitor can read the current location identified by the localization system. This information is particularly relevant from the interactional point of view: since localization is only probabilistic, it may be incorrect from time to time. In this case, for the user to make sense of the information she/he receives, it is important to provide her/him with a feedback about her/his current location according to the localization system. When the location is unknown, the player plays error clips in which Carletto invites the visitor to move to the center of the room and wait for the system to localize her/him.

By using the PDA buttons, the user can stop, suspend and resume the visit. No other forms of interaction are allowed by neither by the graphical interface nor by the PDA build-in controls.

Figure 5. The web site of the virtual tour with Carletto.

5.2 The web-based virtual tour

Setting up the web-based version of the guided tour consists mainly in the tasks of designing and implementing the web interface, which is aimed at orienting the user in the reproduction of the physical space.

The layout of the interface is divided into four square areas, as shown in Figure 5. In the top row, the area on the left contains the multimedia player, while the area on the right contains a PMVR representation of the currently selected room. The multimedia clips and the room representation, in fact, constitute the primary information delivered to the user through this interface, and so they occupy the upper part of it. The bottom row contains a numbered list of the room names (on the left) and a map of the site (on the right), in which rooms are numbered consistently with the list on the left. The map permits the user to figure out the natural order of visit of the rooms by showing the constraints in the physical space, an order which is further enforced by the room numbers. However, the user is left free to move from any room to any other room, since the DramaTour methodology does not make any commitment to a sequential access to the set of rooms if not required by a specific location.

The navigation flow is very simple: at the beginning of the visit, the current room is, by default, the first room. The user can click on the map to move to a different room. In order to provide the user with feedback about the current room, the room name becomes highlighted in the room list. Differently from the guided tour in the physical space, the user location is always known, so the clip player itself does not report any information about the current location.

The client/server schema is implemented in the following way: the web client, i.e., the web browser, only supports the clip player (Shockwave Flash Player 8) and the PMVR player. In order to account for users with different connection bandwidths, the latter may vary according to the user's choice between QuickTime VR and Flash. The visit client is a PHP program that runs on the http server. PHP receives the user's current location form the client and communicates with the visit server (implemented in Java, see previous section) via a socket.

6 Conclusions

The paper has presented the DramaTour methodology for the design and implementation of information presentation systems. The major issue of DramaTour is that information is conveyed through a dramatic narration process acted by a character. Narration items are created by an author, classified according to a well defined set of metadata into a database, interpreted through the animation of a computer graphics character, assembled on–the–fly on the basis of the user's input and the direction of the dramatic narrative. So, the system is half-way between fully intelligent, experimental systems and off–the–shelf scripted systems for practical applications.

We have developed two version of the same applications for two different media, a portable device for on-site guided visits to an historical location and a web-based guided virtual tour of the same location. The design and development of these applications has shown that the methodology is viable the and that it extends to cross-media design.

Both applications have been tested in the first week of April 2006 during the Italian *week of culture* in the historical site of Palazzo Chiablese in Turin. In this occasion, the historical site has been opened to the public and the mobile virtual guide has been offered to the users. The web site for the virtual tour has been online since then at the *url* http://dramatour.di.unito.it. Future work includes the analysis of the data gathered about the system behavior and the visitors' satisfaction, collected through interviews and written forms.

REFERENCES

[1] F. Biral, V. Lombardo, R. Damiano, and A. Pizzo, 'Cyrano goes to hollywood: a drama-based metaphor for information presentation', in *AIMS 2003*, pp. 17–24, (2003).

[2] J. Cassell, J. Sullivan, S. Prevost, and E. Churchill (eds.), *Embodied Conversational Agents*, The MIT Press, Cambridge, Massachussets, 2000.

[3] R. Damiano, V. Lombardo, F. Nunnari, and A. Pizzo, 'Dramatization meets narrative presentations', in Proceedings of ECAI 06, Riva del Garda, (2006).

[4] R. Damiano, V. Lombardo, and A. Pizzo, 'Formal encoding of drama ontology', in *LNCS 3805, Proc. of Virtual Storytelling 05*, pp. 95–104, (2005).

[5] M. Esslin, *The Field of Drama*, Methuen, London, 1988 (1987).

[6] N. Gershon and W. Page, 'What storytelling can do for information visualization', *Communications of the ACM*, **44(8)**, 31–37, (2001).

[7] B. J. Grosz and C. L. Sidner, 'Attention, intentions, and the structure of discourse', *Computational Linguistics*, **12**, 175–204, (1986).

[8] E. Hajicova, B. H. Partee, and P. Sgall, *Topic-Focus Articulation, Tripartite Structures, and Semantic Content*, Dordrecht: Kluwer, 1998.

[9] R. Jackobson, 'The speech event and the functions of language', *On Language*, (1990).

[10] B. Laurel, *Computer as Theater*, Addison Wesley Longman, Reading, MA, 1993.

[11] M. Mateas and A. Stern, 'Integrating plot, character and natural language processing in the interactive drama faade', in *1st International Conference on Technologies for Interactive Digital Storytelling and Entertainment TIDSE 03*, (2003).

[12] R. McKee, *Story*, Harper Collins, New York, 1997.

[13] A. Nijholt and D. Heylen, 'Multimodal communication in inhabited virtual environments.', *International Journal of Speech Technology*, 343–354, (2002).

[14] I. Poggi, C. Pelachaud, F. de Rosis, V. Carofiglio, and B. De Carolis, 'Greta. a believable embodied conversational agent', *Multimodal Intelligent Information Presentation*, (2005).

[15] O. Stock and M. Zancanaro (eds.), *Multimodal Intelligent Information Presentation*, Springer, 2005.