Luciana Bordoni, Francesco Mele and Antonio Sorgente (Eds.)

Proceedings of the

AI*CH 2016

The 10th workshop on Artificial Intelligence for Cultural Heritage

Workshop co-located with AI*IA 2016

Genoa, Italy, November 28, 2016

http://smcm.isasi.cnr.it/AIxCH2016

Copyright © 2016 for the individual papers by the papers' authors. Copying permitted for private and academic purposes. This volume is published and copyrighted by its editors.

Editors' addresses:

ENEA-DTE
Energy Technologies Department
Lungotevere Thaon di Revel 76
00196 Rome, Italy

luciana.bordoni@enea.it (Luciana Bordoni)

CNR-ISASI National Research Council
Institute of Applied Sciences and Intelligent Systems
Via Campi Flegrei, 34 - Comprensorio "A. Olivetti", Ed. 70
80078 Pozzuoli (Naples) - Italy

f.mele@isasi.cnr.it (Francesco Mele)
a.sorgente@isasi.cnr.it (Antonio Sorgente)

Preface

Today, there are some challenges that have to be addressed in the Cultural Heritage (CH) domain: management of large amounts of daily growing data; new interaction paradigms for Cultural Heritage promotion; research of fast and minimally invasive techniques for the assets protection. New applications in the area of Artificial Intelligence (AI) produced innovative services and technology to address such questions in CH.

The 10th Artificial Intelligence for Cultural Heritage workshop (AI*CH 2016) is the meeting point between Artificial Intelligence (AI) areas and Cultural Heritage (CH) domain. During the workshop the latest academic and industrial results in the application of AI methods and techniques to improve the valorization, conservation and promotion of the CH have been discussed and investigated. AI*CH 2016 is inherently interdisciplinary and integrates computer and information sciences with any aspect of the Cultural Heritage sector.

Each proposal has been reviewed by three members of the Program Committee of the Workshop, and on the basis of their recommendations 7 documents have been selected for publication and presentation at AI*CH 2016. In addition, the workshop was enriched by the participation of Sara Di Giorgio of the Progetto Portale della Cultura Italiana - ICCU Rome as invited speaker.

We thank all Program Committee members, our invited speaker, and all authors of accepted papers for making the realisation of the AI*CH 2016 possible.

December 2016 *Luciana Bordoni, Francesco Mele and Antonio Sorgente*

Workshop Organization

Chairs

Luciana Bordoni:	ENEA, Roma, Italy
Francesco Mele:	ISASI - CNR, Napoli, Italy
Antonio Sorgente:	ISASI - CNR, Napoli, Italy

Program Commitee

Liliana Ardissono	University of Turin, Italy
Luciana Bordoni	ENEA, Italy
Antonio Calabrese	ISASI-CNR, Italy
Francesco Cutugno	University of Naples, Italy
Cristina Gena	University of Turin, Italy
Tsvi Kuflik	University of Haifa, Israel
Francesco Mele	ISASI-CNR, Italy
Paul Mulholland	Open University, UK
Cataldo Musto	University of Bari, Italy
Antonio Origlia	University of Naples, Italy
Giovanni Semeraro	University of Bari, Italy
Antonio Sorgente	ISASI-CNR, Italy
Giuseppe Vettigli	Cambridge Coding Academy, UK

Contents

Invited Talk

Developing European Research Infrastructures in the Cultural Heritage Sector - ICCU's Contribution to PARTHENOS Project
Sara Di Giorgio 1

Regular Papers

Listen to What You Look at: Combining an Audio Guide with a Mobile Eye Tracker on the Go
Moayad Mokatren, Tsvi Kuflik, Ilan Shimshoni 2
Steps Towards Accessing Digital Libraries Using Narratives
Carlo Meghini, Valentina Bartalesi, Daniele Metilli 10
Why so Serious? Raising Curiosity Towards Cultural Heritage with Playful Games
Antonio Origlia, Maria Laura Chiacchio, Dario Di Mauro, Francesco Cutugno 18
Designing Interactive Experiences to Explore Artwork Collections: a Multimedia Dialogue System Supporting Visits in Museum Exhibits
Antonio Origlia, Enrico Leone, Antonio Sorgente, Paolo Vanacore, Maria Parascandolo, Francesco Mele, Francesco Cutugno 26
Multi Attributes Approach for Tourist Trips Design
Ilaria Baffo, Pasquale Carotenuto, Antonella Petrillo, Fabio De Felice 34
Integrating Archaeological Datasets: the ARIADNE Portal
Paola Ronzino, Achille Felicetti, Sara Di Giorgio 42
Syncretic Text Composition in Artificial Museum Guides
Antonio Sorgente, Antonio Calabrese, Gianluca Coda, Paolo Vanacore, Francesco Mele 50

Developing European Research Infrastructures in the Cultural Heritage Sector ICCUs Contribution to PARTHENOS Project

Sara Di Giorgio

Progetto Portale della Cultura Italiana - ICCU
sara.digiorgio@beniculturali.it

Abstract

PARTHENOS stands for Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies. It provides a thematic cluster of European Research Infrastructures, integrating initiatives, e-infrastructures and other world-class infrastructures, and builds bridges and provides common solutions to similar problems, between different, tightly interrelated fields. The project, funded by the EC in the framework of H2020 Programme, is creating a common ecosystem to support Humanities, Language, Cultural Heritage, History and Archaeology Research. The project is driving interoperability of tools, applicability of common standards, shared policies on IPR, Open data and Open Access. The ICCU has the task to identify the user requirements of the many communities involved, such as History, Language Studies, Social Science, Cultural Heritage, Archaeology, and related fields across the (Digital) Humanities, about data policies and to implement common policies and strategies for accessing research data, taking into account issues on IPR, Open Data and Open Access. It also contributes to the definition of a common framework for Interoperability and semantics and as content provider, in delivering data from cultural institutions.

Listen to What You Look at: Combining an Audio Guide with a Mobile Eye Tracker on the Go

Moayad Mokatren, Tsvi Kuflik and Ilan Shimshoni
The University of Haifa, Mount Carmel, Haifa, 31905
mmokat03@campus.haifa.ac.il, tsvikak@is.haifa.ac.il, ishimshoni@mis.haifa.ac.il

Abstract. The paper presents work in progress about integrating a mobile eye-tracker into a museum visitors' guide system, so to relive the visitor from explicitly requesting information about objects of interest. The novel and most challenging aspects of the study are the image based positioning and the identification of the visitor's focus of attention, while using a commercially available mobile eye tracker. A prototype system has been developed and it will be evaluated in a user study in a realistic setting. The focus of this paper is possible solutions for real-time efficient image based positioning, the overall system design and the planned evaluation.

1 Introduction

Vision is our main sense for gathering information. When we want to gather information about something in our environment, we first look at it. Moreover, when we express interest in something, we look at it. However, the only information we get in this way is what we see: Size, shape, color, distance etc. Nowadays, a lot of additional information about the objects that we see is available online and can be easily accessible when one searches for it. Theoretically, it is available, just a click away, just a query away, or just by activating the mobile device, writing the query, submitting it, scrolling through the results list, selecting the relevant one and accessing the relevant page. This is a bit complicated set of actions in a mobile scenario, when an immediate, personalized and context-aware information is desired. Current technology offers a variety of ways for information delivery to mobile users. Context awareness is the general term describing the attempt to deliver relevant information at the relevant time and place to the user. What is usually common to most context aware services nowadays is that they make use of the communication and computational power (and sensors) of the users' mobile devices (e.g. mostly smartphones). In addition, they interact with their users mainly by their mobile device's touch screens, which has one major limitation: they are limited in size, the users have to look at them during the interaction, and use a keyboard or select icons etc. Even though voice commands can be used for activating applications, this option is still very limited.

A major challenge in the mobile scenario is to know exactly what the user is interested in. In classical human-computer interaction, the users use a pointing device, most commonly a mouse or by touching a touch screen. However, this is becoming a

major challenge in the mobile setting as noted by Calvo and Perugini [2014], who surveyed novel pointing approaches for wearable computing. The user's position is the best hint, accompanied by the user's orientation. Still, there are many possibly interesting objects near and around the user. If we know what the user is looking at, and what the specific user's gazing profile is, then we can narrow down the possibly relevant objects of interest and we can better serve the user with relevant service/information when needed.

As we move towards "Cognition-aware computing" [Bulling and Zander 2014], it becomes clearer that eye-gaze based interaction should and will play a major role in human-computer interaction before/until brain computer interaction methods will become a reality [Bulling et al. 2012]. With the advent of mobile and ubiquitous computing, it is time to explore the potential of mobile eye tracking technology for natural, intelligent interaction of users with their smart environment, not only in specific tasks and uses, but for a more ambitious goal of integrating eye tracking into the process of inferring mobile users' interests and preferences for providing them with relevant services and information, an area that received little attention so far.

Cultural heritage (CH) is a traditional domain for experimentation with novel computing technology. An intelligent mobile museum visitors' guide is a canonical case of a context-aware mobile system. Museum visitors move in the museum, looking for interesting exhibits, and wish to acquire information to deepen their knowledge and satisfy their interests. A smart context-aware mobile guide may provide the visitor with personalized relevant information from the vast amount of content available at the museum, adapted for his or her personal needs. Mokatren et al. [2016] already presented a novel image based positioning technique using mobile eye tracker for a museum visit, where the position of the visitor is identified in a predefined museum layout, and once is determined an object of interest can be inferred. In this work we aim at developing an audio guide system using a mobile eye tracker on the go as a positioning system and as an implicit pointing device for natural interaction with the system using gesture recognition.

2 Background and Related Work

2.1 Requirements for Museum Audio Visitor's Guide

The museum environment has many limitations, such as the restriction not to make noise, not to talk loudly, not to touch anything, etc. It is obvious that museum visitor's mobile guides should not be a replacement for traditional interpretation methods, but rather complement them [Economou, 1998]. Under these limitations Cheverst et al. [2000] have mentioned two key requirements for such guides, the first of which is Flexibility. The system is expected to be sufficiently flexible to enable visitors to explore, and learn about, a museum in their own way, including controlling their own pace of interaction with the system. The second requirement is that the system will be context aware, meaning that the information presented to the visitors should be

tailored to their personal context. The personal context includes, among other things, the visitor's interests, the visitor's current location and exhibits already visited.

2.2 Image Based Positioning

Consider a device consisting of a forward looking camera and an eye tracker. The device takes a picture while the user is fixating on a certain position within the image. The challenge is to recognize the object in the scene in order to deliver content related to this object to the user.

When an image taken by the front camera of the device, it can be matched to a set of existing images, where the goal is to find which of the images shows the same scene as the test image. The matching algorithm should work in cluttered scenes (scenes from which objects have been removed or added), where the images were not taken from the same pose and with varying illumination. For this to work local image features were developed that are unaffected by nearby clutter or partial occlusion. The features are at least partially invariant to illumination, 3D projective transforms, and common object variations. On the other hand, the features must also be sufficiently distinctive to identify specific objects among many alternatives. Several types of local features have been developed. The most popular type of feature is SIFT [Lowe 1999] but others also exist.

Location-awareness procedure using image matching works as follows:

1. A set of images of the exhibits should be taken, each image may contain one or more objects. For each object that appears in an image, a distinct label value and size of region around the object should be given (in terms of width and height – rectangular shape)
2. Eye-tracker scene camera frame is taken and image-to-image matching procedure is applied using SIFT features. The result is an image with labeled regions in the current scene's frame. A pair of images will be marked as matched if the percentage of the matched feature points (as presented by [Lowe 1999]) is larger than some threshold value (the threshold is determined by case study evaluation).
3. Fixation mapping transformation. The fixation point is transformed from the eye tracker scene camera to a suitable/matched region in the image that we got in step one (image from the data-set with labeled regions)

The result of the above procedure is a location id (or an exhibit id in a museum visit) and point/object of interest (specific object in the exhibit that the visitor looked at).

2.3 Pupil-Dev Mobile Eye Tracker

Pupil eye tracker [Kassner et al. 2014] that is presented in Figure 1, is an accessible, affordable, and extensible open source platform for pervasive eye tracking and gaze-based interaction. It comprises a light-weight eye tracking headset, an open source software framework for mobile eye tracking, as well as a graphical user interface to

playback and visualize video and gaze data. Pupil features high-resolution scene and eye cameras for monocular and binocular gaze estimation.

Figure 1. Pupil eye-tracker (http://pupil-labs.com/pupil)

2.4 Mobile Eye Tracker as a Pointing Device

Eye tracking is an active area of research, where significant progress is continuously made over a long time. Recently, Yousefi, et al. [2015] surveyed a large variety of mobile eye tracking applications and technologies, including aviation, marketing, learning, medicine and more. Furthermore, as predicted (and surveyed by [Yousefi, 2015]), relatively inexpensive, easy to use mobile eye trackers are appearing. Usually, they are experimented in specific areas of applications and tasks. Mokatren et al. [2016] proposed a tool for location awareness, interest detection and focus of attention using computer vision techniques and mobile eye-tracking technology, the focus was on a museum visit. The proposed tool is based on image based positioning technique, for that a set of images that represents the layout of the museum should be taken and stored for image to image comparison.

3 Research Goal and Questions

Our goal is to examine the potential of integrating the eye tracking technology as a natural interaction device into mobile audio guide system (e.g. using the eye-tracker as a natural pointing device in a smart environment). Using it as a pointing device that enables systems to reason unobtrusively about the user's focus of attention and suggest relevant information about the focus of attention as needed.

Our focus is on developing a framework for museum's audio guide that extends the work of Mokatren et al. [2016] for information delivery based on eye gaze detection and image based positioning. We will answer the following question: **How can we integrate the mobile eye tracker as a pointing device in a system that delivers audio information to the visitor?**

For that we have developed a prototype of a system that runs on a laptop and uses Pupil Dev [Kassner et al. 2014] mobile eye tracker for identifying objects of interest and delivering informative content to the users.

In our study we have considered different factors and constrains, the real environment lighting conditions (scenes varies in different day time, e.g. direct sunlight, see figure 2 for example) that can greatly affect the process of image based positioning. For that, different dataset images were taken at different times to ensure successful positioning procedure. Another aspect was the position of the objects relative to the eye tracker holder, since the eye tracker device is head-mounted as this is constrained by the environment layout.

Figure 2. Same exhibit at different day time.

4 Context-aware, Mobile Audio Guide Framework

A key challenge in using mobile technology for supporting museum visitors' is figuring out what they are interested in. This may be achieved by tracking where the visitors are and the time they spend there [Yalowitz and Bronnenkant, 2009]. A more challenging aspect is finding out what exactly they are looking at [Falk and Dierking, 2000]. Given todays' mobile devices, we should be able to gain access seamlessly to information of interest, without the need to take pictures or submit queries and look for results, which are the prevailing interaction methods with our mobile devices.

Lanir et al. [2013] discussed the influence of location-aware mobile guide museum visitors' behavior. Their results indicate that visitors' behavior was altered considerably when using a mobile guide. Visitors using a mobile guide visited the museum longer and were attracted to and spent more time at exhibits where they could get information from the guide. Moreover, they argued that "While having many potential benefits, a mobile guide can also have some disadvantages. It may focus the visitor's attention on the mobile device rather than on the museum artifacts".

In this section we describe the implementation of the audio guide framework that will address the above-mentioned two challenges – it will identify users' focus of attention accurately and it will do that unobtrusively. The system uses Pupil Dev [Kassner et al. 2014] mobile eye tracker (as a pinpoint device for inferring object of interest), laptop (that serves as a computational power) and earphones (for audio information delivery). The system extends the image based positioning technique that was presented by Mokatren et al. [2016] to deliver audio information about exhibits in the museum. A visitor wears the mobile eye tracker which is connected to a laptop (carried on back bag) enters the museum, when he looks steadily for approximately three seconds at an exhibit, the image based positioning procedure starts and location/position and point of interest is identified.

We have implemented two versions of audio mobile guide:

1. Reactive: After identifying the position of the visitor and point/object of interest, "beep" sound is played and immediately after that audio information about the exhibit is delivered (see Figure 3).
2. Proactive: After identifying the position of the visitor and point/object of interest, "beep" sound is played, and the system wait for mid-air gestural action (stop sign). After performing the gestural action, audio information is delivered (see Figure 4).

For both systems we have added an option to stop the audio information delivery at any time by performing mid-air gestural action (stop sign).

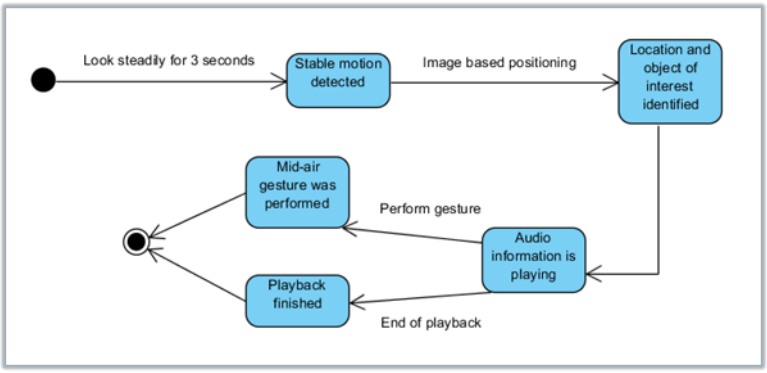
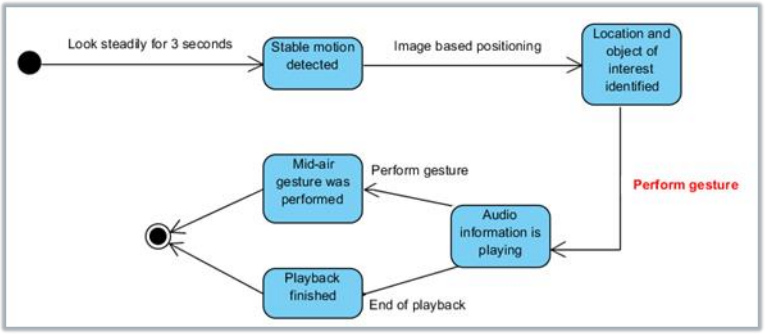


Figure 3. State machine diagram for the **reactive** version



Figre 4. State machine diagram for the **proactive** version

5 Experiment Design

The system will be evaluated in user studies; the participants will be students from University of Haifa. The study will be conducted in Hecht museum[1], which is a small museum, located at the University of Haifa that has both an archeological and art collections.

[1] http://mushecht.haifa.ac.il/

The experiment will be within-subject design that will compare the use of the audio guide with the two versions. The study will include an orientation about using the eye tracker, mid-air gestural interaction (one type of gesture – "stop sign") and the mobile guide, then taking a tour in the museum with the audio guide.

The exhibits will be divided into three categories: Small exhibits, large exhibits and showcases (vitrine shelves). Each case-study will include exhibits from the three categories, we will try to differentiate each case-study exhibits by choosing different exhibits from the same category to reduce as possible the effect of learning.

Data will be collected as follows: The students will be interviewed and asked about their visit experience, and will be asked to fill questionnaires regarding general questions such as if it is the first time that they have visited the museum, their gender and age, and more.

6 Discussion and Conclusions

In the CH setting, visitors' movement in space, time spent, information requested, vocal interaction and orientation were used for inferring users' interest in museum exhibits. Adding eye gaze as additional source may greatly enhance the ability to pinpoint the user's focus of attention and interest (e.g. on products or exhibits), hence improve the ability to model the user and better personalize the service offered to her/him (e.g., exhibit or product information, shopping assistance).

In this paper we presented a framework for a context-aware mobile museum audio guide that uses mobile eye tracking technology for identifying the location of the visitor and inferring his point/object of interest. The audio guide system framework consists of two versions: Reactive and proactive, in the reactive version audio information is delivered immediately once the point of interest is identified, in contrast to the proactive version where the visitor needs to perform a mid-air gestural action to start the audio delivery. The system has not been evaluated yet.

In the image based positioning technique, there is overhead time in matching the camera scene image with every image from the dataset. If the visitor stands at a fixed point and a little time has passed since the last match procedure, then we can search for a matched image from the physical nearest neighbors only. For that we need to represent the data set using a graph, each node will represent the exhibit image/label and the arc value represent the physical distance.

Future work will focus on optimizing the image based positioning procedure by representing the museum layout using graph, and then evaluating the system in an experiment in a museum, in a realistic setting of a museum visit.

References

1. Bulling, A., Dachselt, R., Duchowski, A., Jacob, R., Stellmach, S., & Sundstedt, V. (2012). Gaze interaction in the post-WIMP world. In CHI'12 Extended Abstracts on Human Factors in Computing Systems, 1221-1224. ACM.

2. Bulling, A., & Zander, T. O. (2014). Cognition-aware computing. Pervasive Computing, IEEE,
3. Calvo, A. A., & Perugini, S. (2014). Pointing devices for wearable computers. Advances in Human-Computer Interaction, 2014.
4. Cheverst, K., Davies, N., Mitchell, K., & Friday, A. (2000, August). Experiences of developing and deploying a context-aware tourist guide: the GUIDE project. In Proceedings of the 6th annual international conference on Mobile computing and networking (pp. 20-31). ACM.
5. Economou, M. (1998). The evaluation of museum multimedia applications: lessons from research. Museum Management and Curatorship, 17(2), 173-187.
6. Lowe, David G. "Object recognition from local scale-invariant features." (1999). The proceedings of the seventh IEEE international conference on. Computer vision. Vol. 2 ,pp. 1150-1157
7. Kassner, M., Patera, W., & Bulling, A. (2014). Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. 1151-1160. ACM.
8. Lanir, J., Kuflik, T., Dim, E., Wecker, A. J., & Stock, O. (2013). The influence of a location-aware mobile guide on museum visitors' behavior. Interacting with Computers, 25(6), 443-460.
9. Mokatren, M., Kuflik, T. and Shimshoni, I. (2016) Exploring the potential contribution of mobile eye-tracking technology in enhancing the museum visit experience. Accepted to the workshop on Advanced Visual Interfaces in Cultural Heritages – AVI-CH 2016 – co-located with AVI 2016.
10. Yalowitz, S.S. and Bronnenkant, K. (2009) Timing and tracking: unlocking visitor behavior. Visit. Stud., 12, 47–64.
11. Yousefi, M. V., Karan, E. P., Mohammadpour, A., & Asadi, S. (2015). Implementing Eye Tracking Technology in the Construction Process. In 51st ASC Annual International Conference Proceedings.

Steps Towards Accessing Digital Libraries Using Narratives

Carlo Meghini, Valentina Bartalesi, Daniele Metilli

ISTI "Alessandro Faedo" – CNR, Pisa, Italy
{carlo.meghini,valentina.bartalesi,daniele.metilli}@isti.cnr.it

Abstract. One of the main problems of the current Digital Libraries (DLs) is the limitation of the informative services offered to the users. Indeed, DLs provide simple search functionalities which return a list of the information objects contained in them. No semantic relation among the returned objects is usually reported which can help the user in obtaining a more complete knowledge on the subject of the search. The introduction of the Semantic Web has the potential of improving the search functionalities of DLs. Many cultural institutions have represented their metadata into *formal descriptions* encoded by means of formal languages such as RDF and OWL. In this context, the aim of our research is to introduce the narrative as a new search functionality which does not only return a list of objects but presents a narrative, composed of events that are linked to the objects of the library and endowed with a set of semantic relations connecting these events into a meaningful semantic network. The paper presents the first theoretical achievements on a model for representing narratives.

Keywords: Narratology, Digital Libraries, Semantic Web, Ontologies

1 Introduction

Digital libraries (DLs) are information systems that offer services over large sets of digital objects [23]. The traditional search functionalities of DLs, such as Europeana[1], consider that users express their information need through a natural language query, and the digital library returns a ranked list of digital objects. This approach works well on the web, where the objects of the search are rich text pages with images and links to other pages, but it performs poorly on most DLs. The reason is that the digital objects contained in them (e.g. representations of books, photographs) are not meant to be read and navigated on the fly like web pages. As a result, the response to a web-like query on a digital library is typically a ranked list of metadata descriptors. In our study, we aim at overcoming this limitation of current DLs by introducing a new first-class search functionality: the *narrative*. The vision is that a user searching for a cultural heritage item like the Baptistery of Florence in Europeana would obtain in response not the ranked

[1] http://www.europeana.eu/portal/

list of objects concerning the Baptistery but rather a *narrative* about it, made up of a list of events that compose its history, linked to the objects of the digital library that contextualize them. Our study aims to introduce this new search functionality using Semantic Web technologies, and developing a formal ontology for representing narratives. In order to reach this goal, we studied the narratology and computational narratology literature to identify the basic structure of a narrative. We also reviewed the Artificial Intelligence (AI) literature to identify the logical components of the narrative structure (e.g. events, fluents, agents), and give their logical definitions. Then, we developed a conceptualization of the narrative structure, as derived from the above background. Finally, in order to develop an ontology to represent narratives, we evaluated the CIDOC CRM[2] as reference ontology [11]. The evaluation was based on the mapping between the logic components of narratives and the terms included in the CRM.

The paper is structured as follows: Section 2 presents the narratology and computational narratology background and Section 3 an overview of the related works. In Section 4 an analysis of the AI literature in order to identify the formal components of narratives is reported. Section 5 presents a conceptualization of the narrative structure. In Section 6, a mapping between the formal components of narratives and the CRM is reported. Section 7 presents a brief discussion of the results of the mapping. Finally, Section 8 reports our conclusions.

2 Background

2.1 Narratology

In literary theory, narratology is a discipline that studies the narrative structure and the logic, principles, and practices of its representation [24]. The earliest antecedent to modern narratology can be found in the *Poetics* of Aristotle, who defines a narrative as the imitation of real actions that forms an argument whose fundamental units, or events, can be arranged in a plot [4]. However, the theoretical principles of narratology derive from linguistic-centered approaches to literature defined by Russian formalists in the early 20th century. Russian formalism defines narratology as based on the idea of a universal pattern of codes that operates within the content of a work. A narrative can thus be conveyed through several different means of communication, e.g. speech, writing, gestures, music. In particular, Vladimir Propp proposed a model to represent folktales as combinations of basic building blocks, including thirty-one narrative functions and seven roles of the characters [27]. Russian formalism distinguishes between a *fabula*, defined as a series of events in chronological order, and a *syuzhet*, which is the particular way the story is narrated by its author [27]. The theory of narratology was further developed by mid-20th Century European structuralism. Claude Lévi-Strauss outlined a grammar of mythology [17]; Tzvetan Todorov was the first to coin the term *narratologie* [29]. In the Cognitive Narratology [14] perspective, narratology is considered a psychological phenomenon, and narrative

[2] Conceptual Reference Model

aspects have to be studied from a cognitive perspective. Currently, there is no universally accepted definition of narrative structure. For instance, Crawford [9] claims that a narrative is a high-level structure based on causality, but not on spatio-temporal relations. In addition to the fabula and syuzhet, Bal [5] defines a third level that constitutes the concrete representation of the content that is conveyed to the audience.

2.2 Computational Narratology

Computational narratology studies narratives from a computational perspective, focusing on "the algorithmic processes involved in creating and interpreting narratives, modelling narrative structure in terms of formal computable representations" [19]. Computational narratology is based on engineering disciplines aiming at developing AI systems for reproducing human-like narrative behaviour and intelligent interfaces for interacting with narratives [20]. In AI this term refers to storytelling systems, i.e., software generating stories described in natural language, implementing linguistic formalisms [7]. Some of the early storytelling systems are TALE-SPIN [22] abd UNIVERSE [16]. These are hybrid systems that implement a computer model of creativity in writing. Recently, ontologies were used to generate narratives. For example, it happened in the MAKEBELIEVE [18] and ProtoPropp [13] projects. In our research context, these systems are interesting because they use formal models that provide a symbolic representation of a narrative.

3 Related Works

Narratives have been recently proposed to enhance the information contents and functionalities of DLs, with special emphasis on information discovery and exploration. For example, Bletchley Park Text [26] is a semantic application that allows users to explore collections of museums. The semantic description of the resources is used to organize a collection into a personalized web site based on the chosen topics. In the PATHS project [12] a system that acts as an interactive personalized tour guide through existing digital library collections was created. In this system the events are linked by inherence relations. Within the CULTURA project [1] a tool to enrich the cultural heritage collections with guided paths in the form of short lessons called *narratives* was developed.

4 Components of Narratives in Artificial Intelligence

We conducted a study of the AI literature in order to identify the formal components of narratives. In this Section we report the logic definitions of the components of narratives as defined in the Event Calculus (EC) theory [15], with a brief mention also to the Situation Calculus (SC) [21] as related background. The SC is a logic language for representing and reasoning about dynamical domains. In dynamical domains the scenarios change because of the actions performed by the agents. The basic elements of the calculus are:

- *Situations* represent a sequence of actions. The situation is a state resulting from these actions. Sequences of actions are represented using the function symbol *do*, so that *do(a, s)* represents the new situation after that the action *a* was performed in situation *s*.
- *Fluents* are functions and predicates that vary over situations (e.g. location of the agent). Fluents are situation-dependent components used to describe the effects of actions.
- *Actions* are changes performed by agents from a situation to another in a dynamic world.

SC works well when there is a single agent performing instantaneous, discrete actions. When actions have duration and can overlap with each other the alternative formalism is the EC, which is used for reasoning on actions and changes and it is based on points rather than on situations. EC allows reasoning over intervals of time and fluents are time-dependent rather than situation-dependent. EC axioms define a fluent true at a point in time if "the fluent was initiated by an event at some time in the past and was not terminated by an intervening event" [28]. Davidson [10] defines *actions* as a subclass of *events*. In Davidson's opinion, the distinct sign between general events and actions is the intentionality of actions. Like SC, Event Calculus has actions. However, Davidson's distinction between events and actions is not present. In the EC actions are events. In the following list we report the logical definitions of the components of narratives.

- *Generalized events*. In the context in which actions and objects are aspects of a physical universe with a spatial and temporal dimension, a generalized event is a space-time chunk. This abstraction allows thinking to generalize event concepts like actions, locations, times, fluents and physical objects.
- *Mental events and mental objects*. The relations between an agent and "mental objects" like *believes* and *knows*, are called propositional attitudes, because they identify attitudes that agents can have towards a proposition [28]. Using the reification method it is possible to turn a proposition into an object that could become an argument of a sentence (because only terms and not sentences can be arguments of predicates).
- *Narrative*. A narrative is a possibly incomplete specification of a set of actual event occurrences [25]. The EC is narrative-based, unlike the standard SC in which an exact sequence of hypothetical actions is represented.

5 A Conceptualization of the Narrative Structure

Following the narratology theory and the components of narratives as defined in the AI literature, we envisage a *narrative* as consisting of three main elements:

1. the *fabula*, representing the fabula as defined by the Russian formalism, i.e. the sequence of the events that composes the story in chronological order;
2. the *narrations*, one or more texts that narrate the fabula, and that correspond to Bal's definition of *presentation*;
3. a *reference* function that connects the narrations to the fabula and allows deriving the *syuzhet* (or plot) as defined by the Russian formalism.

Fabula. The fabula is built on top of events, an event being an action or occurrence taking place at a certain time at a specific location. In a fabula, events are connected to each other by three kinds of relations:

- a *mereological* relation, relating events to other events that include them as parts, e.g. the birth of Dante Alighieri, the major Italian poet of the late Middle Ages, is part of the life of Dante;
- a *temporal occurrence* relation, associating each event with a time interval during which the event occurs. An event occurs before (or during, or after) another if and only if the period of occurrence of the former event is before (or during, or after) the period of occurrence of the latter. We formalize these relations between events using Allen's temporal logic [3];
- a *causal dependency* relation, relating events that in normal discourse are predicated to have a *cause-effect* relation in the narrator's opinion, e.g. the eruption of the Vesuvius caused the destruction of Pompeii. We are not interested in modeling the mechanical causal relationships that connect events in a physical process. We are rather interested in a more generic notion of causality, whereby the connected events may be years apart in time and the causal connection may be indirect.

Narrations. Each narration of a fabula consists of one or more narrators and a text, which is *authored by* the narrator(s).

Reference. The reference function connects each portion of text that narrates an event to the narrated event. In order to model reference we need to identify textual units, which we call *narrative fragments* (or simply fragments), each of which narrates one or more events. Notice that the reference function allows deriving the plot of the narrative. Indeed, by visiting the text of the narration in its natural order, it is possible to access the *narrative fragments* and, via these, the events in the fabula, *in the order established by the narrator.*

We provided a specification of the above conceptualization in mathematical terms. This allows us to concentrate on the proper capturing of the notions highlighted above, postponing any language consideration to a later stage, once the mathematical specification will have brought forward the required machinery. A detailed description of the mathematical specification is reported in [6].

6 CRM Mapping of Narratives Components

In order to develop a semantic model to represent narratives, on top of which to develop the new search functionality for DLs, we used the CIDOC CRM (CRM for short) as reference ontology [11]. This choice was determined by the fact that the CRM is an ISO standard ontology and promotes a shared understanding of cultural heritage information through a semantic framework that any cultural heritage organization can use to map its cultural objects. Furthermore, the CRM terms allow mapping the most important logic components of narratives as reported below. The following definitions are extracted from the CRM official documentation[3].

[3] `http://www.cidoc-crm.org/docs/cidoc\_crm\_version\_6.2.pdf`

- *Event*. In the CRM, the class *E5 Event* corresponds to the definition of *event* in the EC theory. This class "comprises changes of states in cultural, social or physical systems, regardless of scale, brought about by a series or group of coherent physical, cultural, technological or legal phenomena".
- *Action*. Actions identified by Davidson correspond to the class *E7 Activity* in the CRM. "This class comprises actions intentionally carried out by an actor that result in changes of state in the cultural, social, or physical systems documented".

In order to refine our mapping, we analysed the single types of generalized events that are useful to represent the components of events and we mapped them with the classes of the CRM.

- *Agent*. The CRM uses the class *E39 Actor* to represent people, either individually or in groups, who have the potential to perform intentional actions of kinds for which someone may be held responsible.
- *Location*. This concept is represented in the CRM through the class *E53 Place*. "This class comprises extents in space, in particular on the surface of the earth [...] independent from temporal phenomena and matter".
- *Time*. CRM uses the class *E52 Time-Span* to represent this concept. "This class comprises abstract temporal extents, in the sense of Galilean physics, having a beginning, an end and a duration".
- *Physical Objects*. In the CRM the class *E18 Physical Thing* describes "all persistent physical items with a relatively stable form, man-made or natural".
- *Mental Objects*. In the CRM the class *E28 Conceptual Object* comprises "non-material products of our minds and other human produced data that have become objects of a discourse about their identity, circumstances of creation or historical implication".

The relations defined on the events (and actions) of the fabula, are expressed by the following CRM properties:

- The *mereological* relation is represented using the property *P9 consists of*, which associates an instance of *E4 Period* with another instance of *E4* that is defined by a subset of the phenomena that define the former. *E5 Event* is a subclass of *E4*, therefore *P9* can be used also as an event mereology.
- The *event occurrence* relation is represented by the CRM property *P4 has time-span*, which describes the temporal confinement of an instance of an *E2 Temporal Entity* and therefore of an event. Because the period of occurrence of an event may not be known, the CRM allows to directly relate events based on their occurrence time. To this end, it introduces seven properties (*P114-P120*) mirroring Allen's temporal logic [2].
- The *causality* relation is represented by introducing a new property of *causal dependency*. This is the only proposed extension to the CRM, whose only causal property *P17 was motivated by* cannot be used for narratives since it relates activities but not events. Indeed, CRMsci[4], an extension of CRM for science, defines a direct causality relation, via the property *O13 triggers,* which "associates an instance of *E5 Event* that triggers another instance of

[4] `http://www.ics.forth.gr/isl/CRMext/CRMsci/docs/CRMsci1.2.2.pdf`

E5 Event with the latter [...]; in that sense it is interpreted as the cause". However, this property is inadequate to the needs of narratives, whose events may be separated by possibly long periods of time.

7 Discussion

As result of the mapping, the logic components of narratives we identified can be defined using classes and properties of the CRM. Furthermore, the CRM provides several subclasses of *E5 Event* which identify types of event (e.g. *E63 Beginning of Existence, E65 Creation*). These subclasses are useful to establish a first categorization of events. Furthermore, another advantage of the use of the CRM is the existence of CRMinf[5], an extension of the CRM, which we are considering to describe the inference processes of the narrator. Indeed, in addition to the components of narratives, we considered to represent the inferential process of a narrator who reconstructs the events that compose a narrative starting from the study of the primary sources. Our model aims at describing the *knowledge provenance*, i.e. the process of tracing the origins of knowledge [8]. Reconstructing the inference process is important to evaluate the trustworthiness of the knowledge. A user can determine the quality of the knowledge based on its derivations. CRMinf aims at representing data attribution, scientific concepts of observation, inferences and beliefs. For these reasons, we have adopted the CRM as reference vocabulary to construct an ontology for representing narratives.

8 Conclusions and Future Work

In this paper we have presented the first theoretical achievements on a model for representing narratives. In particular, we have presented a review of the works on narratives and described our project relating it to developed models both in the cultural heritage and in the digital library fields. The final aim of our research is introducing narratives as as a new search functionality for DLs. As output of a query, this new search functionality should not return just a list of objects, as the current DLs report as output, but it should present a semantic network we called *narrative*, composed of the events of the narrated story. The events and their contextualizing components, e.g. digital objects, are connected by semantic relations meaningful to the user.

References

1. Agosti, M., Manfioletti, M., Orio, N., Ponchia, C.: Enhancing end user access to cultural heritage systems: Tailored narratives and human-centered computing. In: New Trends in Image Analysis and Processing 2013, pp. 278–287. Springer (2013)
2. Allen, J.F.: Maintaining knowledge about temporal intervals. Communications of the ACM 26(11), 832–843 (1983)

[5] `http://www.ics.forth.gr/isl/CRMext/CRMinf/docs/CRMinf-0.7.pdf`

3. Allen, J.F.: Towards a general theory of action and time. Artificial intelligence 23(2), 123–154 (1984)
4. Aristotele: Poetica. Laterza (1998)
5. Bal, M.: Narratology: Introduction to the theory of narrative. University of Toronto Press (1997)
6. Bartalesi, V., Meghini, C., Metilli, D.: Steps towards a formal ontology of narratives based on narratology. In: CMN 2016 (2016)
7. Cavazza, M., Pizzi, D.: Narratology for interactive storytelling: A critical introduction. In: Technologies for Interactive Digital Storytelling and Entertainment, pp. 72–83. Springer (2006)
8. Committee, P.E., et al.: Premis data dictionary for preservation metadata, version 2.0. Retrieved May 22, 2010 (2008)
9. Crawford, C.: Chris Crawford on interactive storytelling. New Riders (2012)
10. Davidson, D.: Essays on actions and events: Philosophical essays, vol. 1. Oxford University Press (2001)
11. Doerr, M.: The cidoc conceptual reference module: an ontological approach to semantic interoperability of metadata. AI magazine 24(3), 75 (2003)
12. Fernie, K., Griffiths, J., Archer, P., Chandrinos, K., de Polo, A., Stevenson, M., Clough, P., Goodale, P., Hall, M., Agirre, E., et al.: Paths: Personalising access to cultural heritage spaces. In: Virtual Systems and Multimedia (VSMM), 2012 18th International Conference on. pp. 469–474. IEEE (2012)
13. Gervás, P., Díaz-Agudo, B., Peinado, F., Hervás, R.: Story plot generation based on cbr. Knowledge-Based Systems 18(4), 235–242 (2005)
14. Herman, D.: Narratology as a cognitive science. image and narrative 1(1) (2000)
15. Kowalski, R., Sergot, M.: A logic-based calculus of events. In: Foundations of knowledge base management, pp. 23–55. Springer (1989)
16. Lebowitz, M.: Story-telling as planning and learning. Poetics 14(6), 483–502 (1985)
17. Levi-Strauss, C.: Structural analysis in linguistics and in anthropology. Semiotics-An Introductory Anthology pp. 110–128 (1963)
18. Liu, H., Singh, P.: Makebelieve: Using commonsense knowledge to generate stories. In: AAAI/IAAI. pp. 957–958 (2002)
19. Mani, I.: Computational modeling of narrative. Synthesis Lectures on Human Language Technologies 5(3), 1–142 (2012)
20. Mani, I.: Computational narratology. Handbook of narratology pp. 84–92 (2014)
21. McCarthy, J.: A basis for a mathematical theory of computation. Computer programming and formal systems 354 (1963)
22. Meehan, J.R.: Tale-spin, an interactive program that writes stories. In: IJCAI. pp. 91–98 (1977)
23. Meghini, C., Spyratos, N., Sugibuchi, T., Yang, J.: A model for digital libraries and its translation to rdf. Journal on Data Semantics 3(2), 107–139 (2014)
24. Meister, J.C.: Narratology. (2012), synthesis. American Museum of Natural History. Available at http://ncep.amnh.org.
25. Miller, R., Shanahan, M.: Narratives in the situation calculus. Journal of Logic and Computation 4(5), 513–530 (1994)
26. Mulholland, P., Collins, T.: Using digital narratives to support the collaborative learning and exploration of cultural heritage. In: Database and Expert Systems Applications. pp. 527–531. IEEE (2002)
27. Propp, V.: Morphology of the Folktale, vol. 9. University of Texas Press (1973)
28. Russell, S.J., Norvig, P.: Artificial intelligence: a modern approach. Prentice Hall (1995)
29. Todorov, T.: Grammaire du décaméron. Mouton The Hague (1969)

Why so Serious? Raising Curiosity Towards Cultural Heritage with Playful Games

Antonio Origlia[1,2], Maria Laura Chiacchio[1], Dario Di Mauro[1], Francesco Cutugno[1,2]

[1] University of Naples "Federico II"
[2] Inst. of Applied Sciences and Intelligent Systems of CNR
antonio.origlia@unina.it, marialaura.chiacchio@gmail.com, dario.dimauro@unina.it, cutugno@unina.it

Abstract. *Serious games* have an important role in supporting access to cultural heritage through storytelling and game mechanics. These games, however, are more suitable for learning environments: in order to stimulate people to look for cultural content, other means are necessary. In this paper we present our view on the role *playful games* may have in eliciting curiosity and how a specific gaming mechanics, customised characters building, may change the way technological systems contribute in attracting people to cultural sites.

1 Introduction

The term *gamification* has become very popular in the last years, as digital games are becoming more and more integrated with everyday life. The term indicates the process of adding a layer of mechanics, typically associated with games, to certain tasks in order to make them less imposing. Gamification approaches may also be designed to introduce a rewarding factor to the decision making process, so that people perform the task in a way that the designer considers advantageous, as in the case of points collection in supermarkets. People engaged in gamified tasks have a serious attitude and gamification results in minimally invasive mechanics that are well-integrated with the task at hand.

The gamification idea is sometimes abused and may give rise to misunderstandings. Following [14, p. 46], although players "[...] might be motivated for a while by shiny prizes, real engagement requires a much stronger lure. That means a deeper, more interesting system design must be developed". It is therefore important to understand that the goal of gamification is not to amuse people but to reduce the negative impact of due tasks. Using games to support learning has proven itself to be effective. Games designed with a main purpose other than pure entertainment are called *serious games*. Compared to gamified tasks, serious games do not integrate a pre-existing experience and make use of more complex mechanics, thus being completely independent objects. In these games,

the intended message is set in the forefront, users become rapidly aware that the main goal of the experience is not pure entertainment. Serious games find their natural application in teaching environments. In the context of cultural heritage, their use has been repeatedly tested with promising results concerning the *learning* experience. Some examples are "Icura" [13], "The battle of Thermopylae" [6] and "Thiatro" [11,12]. For a full review of serious games for cultural heritage, see [27].

Re-establishing the connection between people and cultural heritage is a topic that challenges modern museums as they struggle to find a place in the information age. The definition given by the International Council of Museums states that museums are institutions that should provide *education, study and enjoyment*. The museum is, therefore, not only a place to learn but also a place people may choose to look for enjoyment. One of the ways museums can provide enjoyment is by satisfying curiosity about cultural heritage. It is therefore necessary to elicit this curiosity in order to let people rediscover museums. Serious games, being designed for learning tasks, tend to be based on an extrinsic motivation provided by the learning environment or by rewards favouring task completion *as it is supposed to be done*. Extrinsic motivation is known, however, to have a detrimental effect on intrinsic motivation [9]. As curiosity is an intrinsic motivational force, it cannot be elicited with *due* tasks, regardless of their gamified looks. Intrinsic motivation comes from personal disposition of doing something *for its own sake*, like to have fun: it is through this path that curiosity can be activated.

If we relate differences in game mechanics complexity, separating gamified tasks from serious games with the specific goal of letting people *have fun*, we obtain puzzles and playful games. The former are defined in the dictionaries as toys, games or other contrivances to be solved by ingenuity or persistence. The latter are much more complex to define but a common feature is that "[...] a good game is a machine that generates stories when people play it" [31, p.300]. Such games are of interest for museums as storytelling has an important communicative function in cultural heritage [2,19].

In this work, we will examine the role playful games can have in generating curiosity towards cultural heritage through technological approaches. While there are many aspects of playful games that can help accomplish this goal, we will concentrate on one particular feature that, in our opinion, has also the potential to provide critical information to artificial intelligence systems for automated narrative adaptation and guided tours: customised characters creation.

2 Just for fun

Although the positive impact of games in serious activities is proven, the main reason why people play is simple: having fun. Gamified tasks and serious games act as the classic *spoonful of sugar*, making due tasks less imposing but their main goal is not to entertain people. On the other hand, entertaining media, like movies, have the power to influence people's opinions and motivation. Re-

sults presented in [5] showed that, seeing the "JFK" movie was associated with a significant decrease in viewers' reported intentions to vote or make political contributions. The study presented in [8] showed that "Malcolm X" significantly increased people awareness concerning discrimination issues. More recently, [1] showed that both "The rainmaker" and "As good as it gets" increased people's support for Obama's *Affordable Care Act*. It is not uncommon, today, to see digital games adopt complex themes like the risks involved in pervasive surveillance systems depicted in "Watchdogs" and the metaphor of racial hate used in "Deus-Ex: mankind divided". The line separating a serious game from a playful game, in the cultural heritage field, is subtle as both include fun and cultural components. Serious games appear to be more common and aim at improving the quality of the learning experience. Playful games should aim, instead, at increasing the general curiosity of people towards cultural heritage. The importance of storytelling to deliver such contents is well established and, nowadays, interactive applications are common. In the case of games, however, being interactive is not enough: player choices must have a clear impact on the story being told, thus evoking a sense of *agency*. Agency may be obtained in multiple ways, like with branching stories. It represents a fundamental component for successful games: professional designers highlight that "[...] a failure to provide a convincing sense of agency is frequently a reason that game scenes (or entire games) fall flat" [17, p. 106].

Although artificial intelligence has been traditionally used in digital games to act as the player's opponent and to control virtual allies, experimental approaches use it to monitor user choices to dynamically adjust the narrative. These challenge traditional theoretical frameworks of narrative description [22]. They may take into account, for example, the manually annotated tension of narrative events, as in "Façade", or the user inclination towards specific playing styles to predict emotional feedback [16]. These systems appear, however, not to be taking full advantage from mechanics traditionally used in the gaming world to support the emergence of engaging, collaboratively built narratives. Among these, customised character creation is the most relevant.

3 Customised characters

Character creation has multiple applications in the framework of cultural heritage enjoyment. In this section we examine the potential impact of this mechanics in the field of AI for cultural heritage.

3.1 Collaborative narrative

AI-controlled narrative systems appear to start from a common assumption: the player is totally unknown to the system when the game begins. This, however, is rarely the case with playful games. First of all, it is mandatory for these to provide an invitation to play. When accepted, this testifies the user's will of entering

lusory attitude: the "curious state of affairs wherein one adopts rules which require one to employ worse rather than better means for reaching an end" [32, p. 23]. This element is critical for technological systems as it gives users a socially acceptable reason to contribute to narrative building, through the constraints established by the game. Professional game designers stress that "crafting this invitation to play, making it visceral and compelling to your target audience, is an important part of playcentric design" [14, p. 56]. Through the ways a player can contribute to the narrative, the system should try to adapt the story. Meaningful choices are a widely used element to involve people in shaping the narrative but game designers have also devised other means to accomplish this. Role playing games (RPGs), in particular, have explored the topic substantially with great success. While RPGs have been proposed for cultural heritage, there is still confusion about the relationship between the player and player controlled characters (if they are even present). One common misunderstanding lies in assuming that the player *is* the character. In RPGs, the player *creates* a character she would like to guide through the narrative by exploiting his capabilities and, as importantly, dealing with his deficiencies. While the classic "Dungeons and Dragons" (Gygax & Arneson, 1974) mechanics dealt most with defining action-oriented activity, there are other role-playing games that greatly focus on the character creation system by adding, for example, rules to balance virtues and disadvantages, as in the case of "Cyberpunk 2020" (Talsorian Games, 1988) and of the Storytelling system introduced by the more recent "World of Darkness" (White Wolf Gaming Studio, 2004). From the point of view of artificial storytelling, RPGs have the additional advantage of being designed to be played by small groups of people, which is of interest for cultural heritage [3,10]. The social component that comes from creating collaborating characters is a strong feature of this kind of game: the range of available abilities is wide and it is not possible for a single character to master them all. RPGs are designed in such a way that small groups of people can describe the role they would like to have in shaping the narrative together, which is also of interest for cultural heritage [30,21]. This, combined with social data coming from sources normally used for recommendation (see [28,29]) may yield critical insight to start adapting the narrative to the group even before the beginning of the experience.

3.2 Emotion elicitation

Playful games should provide fun to the players. Fun is a strongly emotional concept and can be defined as "pleasure with surprises" [31, p. 36]. Character creation is an activity that provides a lot of fun, as testified by the success of pen-and-paper RPGs but also by digital games like "the Sims". Attachment to the created characters is a powerful key to elicit players emotions, as shown by the "XCOM" series through the use of *permadeath*. Fun games, therefore, elicit emotions and characters are a strong mean to access the players' mind as they "[...] often wince in imagined pain upon seeing their avatar suffer a blow or sigh in relief upon seeing their avatar escaping physical harm." [31, p. 348]. The player builds strong empathic ties with her characters as she shapes them.

While empathy towards virtual characters is taken into account when designing virtual companions for guided tours, as in the case of "A stroll with Carletto" [7], customised characters have the potential to go beyond simple sympathy and activate the cognitive mechanisms related to emotions. This is the case of regret [18] as a consequence of agency because "[...] what matters for feeling regret is that the individual represents - even a posteriori - the situation as a choice" [26, p. 89]. In RPGs, players *gamble* the result of their creative effort, characters, by exposing them to a story that may damage, or even destroy, them. The menace of impending punishment on the character is one of the key components in making a game fun, as the risk of experiencing it creates endogenous value, provides excitement and increases challenge [31] by anticipating possible regret. Relief, on the other hand, "[...] increases with effort expenditure" [26, p. 106] which, in games, is provided by the challenges created through mechanics. Automatically balancing the challenge while keeping consistency with player expectations is an important task usually assigned to AI systems in digital games. From the player's point of view, preventing the threat of seeing *her* characters harmed creates the first component of fun: pleasure. Stories, on the other hand, are designed to provide the second component of fun: surprises. The relationship between surprise and curiosity is very strong and, in general, surprising events have the potential to generate curiosity. According to [24], curiosity reflects the desire to close inherently unpleasant information gaps. This desire, however, depends on the perceived likelihood that the gap will be closed by accessing information. [25]. Moreover, it has also been suggested that "[...] the amount of pre-existing knowledge in a particular domain may impact on the perceived likelihood of closure" [26, p. 57]. Specifically, the more information one already possesses on a certain domain, the more curious she becomes about the subject. This implies that curiosity may arise not only by increasing the amount of information available, but also by reducing the *perceived* size of the domain. Museums can be intimidating for the general public as information gaps may be perceived as less likely to be closed because of topics wideness. If the story is designed to leave small information gaps, however, the interest domain will be constrained to the story itself, which is more manageable and therefore likely to stimulate curiosity. Information gaps in such stories should lie in the background and left to be filled after the story has been told. Examples of how to accomplish this may be recurring jokes, historical background and casual conversations among non-playing characters.

3.3 Support the visit

In our view, as playful games are meant to stimulate curiosity towards cultural heritage, they must be conceived to be experienced off-site: the playful component must not interfere with on-site activities aimed at delivering the information. Customised characters may constitute an element of continuity between the different phases of cultural heritage experience described in [23]. They may accompany the transition from a playful activity, the game, to a serious one, the

visit, to the last one, summarisation. Their use may solve a number of problems interactive avatars have. First of all, when a user activates such a guide, a phase of acquaintance is necessary either to introduce the virtual character or simply the user interface. This implies a novelty effect that may interfere with the learning experience until it wears off, as shown by [20]. By allowing the automated guide system to assume the looks of the customised character, however, this effect may be reduced. Having a well known character associated with a playful experience also introduces the possibility of exploiting the power of the story in a completely new way. In museums, it has been observed that removing the object from its original context has the detrimental effect of atomisation. Stories, however, have the power to reinstate the context lost by atomisation while creating meaning, relevance and empathy [15]. The idea of putting stories before collections for museums has been explored in [4]. By observing the engagement readers displayed towards journalistic material combining rigorous research with vivid storytelling, the authors tested the same approach in museums. Specifically, they observed that people resisted to *cognitive kickouts* aimed at providing deeper contents and chose not abandon the story. This problem may arise because the two moments were mixed up: the engaging power of the narrative may prevent the user to access deeper contents while the story is being told. In order to separate the two phases, it may be necessary to provide a strong element of continuity to make people perceive information gaps as more likely to be closed. Customised characters, using their empathic value, can appeal to a powerful contextualisation strategy: recall of a shared experience. Specifically, the character can insert references to the story she lived with the player so that context is not built *during* the visit but is first provided off-site and then evoked on-site.

4 Conclusions

Numerous digital games have been designed to support cultural heritage but the majority of these appear to fall into the serious games category, as teaching is their main goal. Switching to a view that considers pure entertainment as the objective, however, we hypothesise that it is possible to recover concepts and mechanics coming from the digital gaming industry to stimulate curiosity towards cultural contents. We highlighted how the use of customised characters can support long-term interaction with cultural contents, social enjoyment of museums and AI controlled narrative development.

Future work will consist in producing prototypes to evaluate the impact of this particular mechanics and to explore the potential of other strategies used in playful games to stimulate curiosity towards cultural heritage.

References

1. Adkins, T., Castle, J.J.: Moving pictures? experimental evidence of cinematic influence on political attitudes. Social Science Quarterly 95(5), 1230–1244 (2014)

2. Bedford, L.: Storytelling: The real work of museums. Curator: the museum journal 44(1), 27–34 (2001)
3. Belinky, I., Lanir, J., Kuflik, T.: Using handheld devices and situated displays for collaborative planning of a museum visit. In: Proceedings of the 2012 International Symposium on Pervasive Displays. p. 19. ACM (2012)
4. Birchall, D., Faherty, A.: Big and slow: adventures in digital storytelling. In: Proc. of Museums and the web (2016)
5. Butler, L.D., Koopman, C., Zimbardo, P.G.: The psychological impact of viewing the film "jfk": Emotions, beliefs, and political behavioral intentions. Political psychology pp. 237–257 (1995)
6. Christopoulos, D., Mavridis, P., Andreadis, A., Karigiannis, J.N.: Using virtual environments to tell the story:" the battle of thermopylae". In: Games and Virtual Worlds for Serious Applications (VS-GAMES), 2011 Third International Conference on. pp. 84–91. IEEE (2011)
7. Damiano, R., Gena, C., Lombardo, V., Nunnari, F., Pizzo, A.: A stroll with carletto: adaptation in drama-based tours with virtual characters. User Modeling and User-Adapted Interaction 18(5), 417–453 (2008)
8. Davis, D.W., Davenport, C.: The political and social relevancy of malcolm x: The stability of african american political attitudes. The Journal of Politics 59(02), 550–564 (1997)
9. Deci, E.L., Koestner, R., Ryan, R.M.: A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. Psychological bulletin 125(6), 627 (1999)
10. Dim, E., Kuflik, T.: Automatic detection of social behavior of museum visitor pairs. ACM Transactions on Interactive Intelligent Systems (TiiS) 4(4), 17 (2015)
11. Froschauer, J., Arends, M., Goldfarb, D., Merkl, D.: Towards an online multiplayer serious game providing a joyful experience in learning art history. In: Games and Virtual Worlds for Serious Applications (VS-GAMES), 2011 Third International Conference on. pp. 160–163. IEEE (2011)
12. Froschauer, J., Merkl, D., Arends, M., Goldfarb, D.: Art history concepts at play with thiatro. Journal on Computing and Cultural Heritage (JOCCH) 6(2), 7 (2013)
13. Froschauer, J., Seidel, I., Gärtner, M., Berger, H., Merkl, D.: Design and evaluation of a serious game for immersive cultural training. In: Virtual Systems and Multimedia (VSMM), 2010 16th International Conference on. pp. 253–260. IEEE (2010)
14. Fullerton, T.: Game design workshop: a playcentric approach to creating innovative games. CRC press (2014)
15. Haven, K.F.: Story proof: The science behind the startling power of story. Greenwood Publishing Group (2007)
16. Hernandez, S.P., Bulitko, V., Spetch, M.: Keeping the player on an emotional trajectory in interactive storytelling. In: Eleventh Artificial Intelligence and Interactive Digital Entertainment Conference (2015)
17. Heussner, T., Finley, T.K., Brandes-Hepler, J., Ann, L.: The game narrative toolbox. OUP Oxford (2014)
18. Higgins, E.T.: Promotion and prevention: Regulatory focus as a motivational principle. Advances in experimental social psychology 30, 1–46 (1998)
19. Johnsson, E.: Telling Tales: A guide to developing effective storytelling programmes for museums. Museums Hub (2006)

20. Kanda, T., Hirano, T., Eaton, D., Ishiguro, H.: Interactive robots as social partners and peer tutors for children: A field trial. Human-computer interaction 19(1), 61–84 (2004)
21. Katifori, A., Perry, S., Vayanou, M., Pujol, L., Chrysanthi, A., Ioannidis, Y.: Cultivating mobile-mediated social interaction in the museum: Towards group-based digital storytelling experiences. In: Proc. of Museums and the Web (2016)
22. Koenitz, H.: Towards a specific theory of interactive digital narrative. Interactive Digital Narrative pp. 91–105 (2015)
23. Kuflik, T., Wecker, A.J., Lanir, J., Stock, O.: An integrative framework for extending the boundaries of the museum visit experience: linking the pre, during and post visit phases. Information Technology & Tourism 15(1), 17–47 (2015)
24. Loewenstein, G.: The psychology of curiosity: A review and reinterpretation. Psychological bulletin 116(1), 75 (1994)
25. Loewenstein, G., Adler, D., Behrens, D., Gillis, J.: Why pandora opened the box: Curiosity as a desire for missing information. Unpublished manuscript, Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA (1992)
26. Miceli, M., Castelfranchi, C.: Expectancy and emotion. OUP Oxford (2014)
27. Mortara, M., Catalano, C.E., Bellotti, F., Fiucci, G., Houry-Panchetti, M., Petridis, P.: Learning cultural heritage by serious games. Journal of Cultural Heritage 15(3), 318–325 (2014)
28. Rossi, S., Barile, F., Improta, D., Russo, L.: Towards a collaborative filtering framework for recommendation in museums: from preference elicitation to group visits. In: Proc. of the International workshop on data mining on IoT Systems (2016)
29. Rossi, S., Cervone, F.: Social utilities and personality traits for group recommendation: a pilot study. In: Proc. of the International conference on Agents and artificial intelligence. vol. 1, pp. 38–46 (2016)
30. Roussou, M., Pujol, L., Katifori, A., Chrysanthi, A., Perry, S., Vayanou, M.: The museum as digital storyteller: Collaborative participatory creation of interactive digital experiences. In: Proc. of Museums and the web (2015)
31. Schell, J.: The art of game design. CRC Press (2015)
32. Suits, B.: The Grasshopper: Games, Life and Utopia. Broadview Press (1990)

Designing Interactive Experiences to Explore Artwork Collections: a Multimedia Dialogue System Supporting Visits in Museum Exhibits

Antonio Origlia[1,2], Enrico Leone[1], Antonio Sorgente[2], Paolo Vanacore[2], Maria Parascandolo[1], Francesco Mele[2], and Francesco Cutugno[1,2]

[1] PRISCA-Lab, Federico II University, Naples, Italy,
{antonio.origlia, enrico.leone, maria.parascandolo, cutugno}@unina.it,
[2] Institute of Applied Sciences and Intelligent Systems - CNR, Naples, Italy,
{a.sorgente, p.vanacore, f.mele}@isasi.cnr.it

Abstract. Speech and natural language processing have a central role in the implementation of systems designed to make the museum more reactive to users' inputs and to improve the overall interaction quality. In this paper, we present the design and implementation of a dialogue system to provide multimedia presentations for museum visits. A corpus of speech recordings in Italian was collected with a mobile application to obtain a reference set of possible ways for the users to express their intentions. On the basis of this corpus, a set of recurring syntactic patterns associated to device requests was extracted to let the dialogue system separate device commands from information queries. Disambiguation strategies depending on the context are also applied in presence of partial syntactic patterns. Information queries are answered by automatically assembling portions of semantically annotated texts and are synchronized with relevant multimedia resources. A case study on the '800 exhibit at the Capodimonte museum in Naples is presented[3].

Keywords: dialogue systems, cultural heritage

1 Introduction

Italian has a very poor visibility in the area of spoken dialogue systems basic research. The EVALITA evaluation campaign held in 2009 [1] showed the state of art for telephonic systems was limited to the features offered by VoiceXML standard. Participants in that evaluation campaign were able to set up three different system initiative dialogue managers in the field of train services (ticketing, booking and timetable queries) but performances were found to be below the performance that is usually obtained on English. Also, recent dialogue systems for Italian equipped with semantic reasoning capabilities were presented in [2–4], but they only considers chat based interaction. In the passage from telephonic

[3] This work is supported by the Italian PAC project *Cultural Heritage Emotional Experience See-Through Eyewear* (CHEESE).

to mobile applications first and to generalized spoken language understanding systems, most of the Italian researchers participated into international projects and mainly worked on languages different from their own. This is the case, for example, of the recently published SPEAKY [5] development environment for robotic vocal interfaces. At the same time big companies that were investing in personal assistant mobile apps and similar products, extended their native solutions to our language following industrial procedures that did not give raise to knowledge to be spread in the scientific community. In this paper, we will describe the development of a dialogue system integrated with remote augmented reality interfaces in a cultural heritage setting. We will include a brief description of the problems that arise when dealing with delicate environments, like the '800 exhibit in the Capodimonte museum. These pose serious limitations to technological interventions that have an impact on the overall design process. We will describe how these were addressed and the how the system architecture was implemented.

2 Material

An important problem that arises when working with an environment that requires technology to be non-invasive, like a museum exhibit, is that it is difficult to involve end users in the early steps of system development. Exhibits may not be always open to the general public, estimated visitor attendance can vary due to external conditions and wifi connectivity is not always guaranteed. In our case study, the '800 exhibit at the Capodimonte museum in Naples, wifi connection presents a problem of its own as the exhibit is located inside the Bourbon Royal Palace, where walls are very thick and the possibility of intervention are limited. For this reason, in order to obtain a reference set of possible ways for the users to access the system's functions before this is deployed, we used a simple prototype application to collect speech utterances and design the dialogue system accordingly.

The application is implemented as an Android app running on a smartphone in uncontrolled environments. To avoid influencing participants in producing always the same utterances and obtain a higher expressive variability, we chose to present the scenarios using an iconographic approach. Each participant, at each step, was prompted with a set of icons representing a specific user request. An example of the *VolumeUp* scenario is presented in Figure 1. The participant can record the utterance, listen to it and submit it to the remote collection server when she is satisfied. Scenarios cover both device commands (volume control, taking pictures, recording videos...) and content-related queries. If a prompt was not clear, the user would tap the single icons to visualize a single word explaining the icon. This way, no suggestion about how to combine the icons to derive the scenario were given. The users were, in general, able to derive the meaning of the prompt. After a manual check, only 171 utterances have been discarded because the users provided inconsistent recordings.

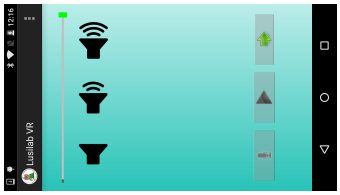


Fig. 1. A screenshot of the mobile app used to collect the reference corpus.

The prompts were randomly presented to the users and were proposed five times each to encourage people to provide multiple ways to ask for the same service, increasing expressive variability. 22 gender-balanced participants with good technological competence were recruited and 17 scenarios were foreseen. A total number of 1870 recordings were collected this way. An expert linguist listened to the material and provided the correct transcription to obtain an estimate of the ASR errors. After a manual check, the correct transcription was found as 1-best in 70% of the cases. In 11% of the correct transcription was either presented a 2-best or 3-best. In the remaining 19% of the cases, the Google ASR engine was not able to provide the correct transcription on the first 3-best. This is mainly caused by the variable quality of the recordings and it represents a good approximation of the performance we can expect from the ASR. Also, it indicate that a good number of cases may be recovered by applying re-ranking techniques. A common error committed by the ASR engine was providing the transcription "Firma l'audioguida" (Sign the audioguide) as the 1-best while the correct "Ferma l'audioguida" (Stop the audioguide) was found as 2-best. In our context, of course, "Ferma l'audioguida" makes more sense than "Firma l'audioguida" and would be recovered, ideally bringing the system's performance around 81% correct transcription.

Concerning the question answering system, museum experts provided a collection of texts and media related the '800 exhibit. This material contains textual information describing 4 museum rooms and 7 artworks and it also contains 123 media objects linked to the relevant parts of the reference texts. This allows the question answering system to control the timing of potential accompanying media presentations when the answer to a question is assembled.

3 System architecture

In this section, we describe the client-server architecture used to deploy the CHEESE system. Most of the logic is located on the server side but some issues related to the client and how these were managed are worth mentioning.

Client side. Although the dialogue management is independent from the client interface, limits due to the chosen wearable device influence the configuration of the speech manager. In our case, we use the Epson Moverio BT-200 glasses for augmented reality, which are equipped with Android 4.0.4. This is an important

issue as, in this version of Android, no offline recognition support is offered by the system. For this reason, we developed an Android app that continuously listens to the microphone and streams audio towards the server. On server side, an audio acquisition thread collects the recorded input and applies Voice Activity Detection before connecting to Google Speech to obtain the transcription. To perform audio streaming and segmentation, adintool, which is part of the Julius ASR engine [6], is used. In order to connect the two parts of the system, the audio streaming procedure replicates, in Java, the C++ procedure used by adintool when operating in client mode. As future versions of wearable glasses will be equipped with higher versions of Android supporting offline recognition, the system can also be configured to manage strings instead of audio.

Server side. The server side of the dialogue system is centered on the Opendial framework [7], which provides a flexible environment to design dialogue systems using and XML-based language and can also be extended with customized plugins using Java. Opendial represents the dialogue state as a set of variables and it lets the user define a series of internal models triggered by variable updates that automatically produce reactions accordingly to the the observed state.

Although not mandatory, there are typically three main models in an Opendial application. The *Natural Language Understanding* (NLU) model analyzes the user input and it maps it on a finite set of possible user actions. The *Action Selection Model* (ASM) connects the user action to the correspondent machine action. The *Natural Language Generation* (NLG) model produces spoken content in accordance with the selected machine action. In the CHEESE framework, we have three separate NLU models to handle different moments of the interaction: the first separates commands to the device (volume control, taking pictures or videos, etc.) from user queries concerning cultural heritage items that are part of the considered exhibit. The second detects requests for device-related functions. The third one detects incomplete commands and summarizes the possible outcomes so that clarification strategies can be applied to recover the interaction.

As the system is focused on Italian, a set of plugins to include, in Opendial, the following software tools, needed to process this language: a plugin to receive the audio stream from the client and transcribe it using Google Speech; a plugin to obtain POS tags from the Treetagger[8] tool; a plugin to normalize the utterance substituting synonyms of target terms with the target term itself; a plugin to extract the dependency-based parse tree of the normalised utterance using the Turin University Linguistic Environment (TULE)[9]; a plugin to connect Opendial to the higher-level system handling user queries.

Concerning the last plugin, a communication protocol based on JavaScript Object Notation (JSON) has been adopted. The JSON string contains the multimedia response for the user and defines the synchronisation of synthesised text and media. Its structure is based on a simplified version of Synchronized Multimedia Integration Language (SMIL)[4]. The main modules used for the in-

[4] https://www.w3.org/TR/REC-smil/

terpretation of user requests are a parser to identify its grammatical structure, a set of semantic services implemented for the detection of semantic concepts in the text such as events, entities, locations, etc, and services to access semantic repository as MultiWordnet[10] and Wiktionary (https://www.wiktionary.org/).

4 Dialogue System

In this Section, we describe the dialogue management logic governing the system. The system is modular and it applies a pipeline process after receiving an input utterance to evaluate its content and plan a reaction. First of all, the input string is preprocessed to obtain a normalised utterance using the plugins described in the previous Section. The dialogue state is then updated considering the incoming utterance and the current position of the user, which is provided externally and is relevant to answer queries like "Chi l'autore di questo quadro" (Who is the author of this painting). The NLU model dedicated to the detection of WH-questions is the first to run. If this model detects a WH-question, the ASM gives control to the question answering system, otherwise the NLU command detection model is run. If a syntactic pattern associated with a device command is detected, the ASM activates the corresponding device action, otherwise the NLU model for recover strategies is activated. This model checks if incomplete syntactic patterns can be detected in the utterance and, if this is the case, the ASM instructs the NLG model to pose an appropriate question to the user to disambiguate the command. This module also attempts to resolve ambiguities based on the current context. If no partial syntactic pattern can be found, the system asks the user to confirm the automatic transcription and, if this is confirmed, aborts the interaction as it is not able to help.

4.1 Command/query separation

In our approach, we have focused on wh-question (or content questions), queries containing a question word (called wh-words) such as, for example, 'chi' (who) or 'quando' (when) [11]. For the identification of wh-questions, a set of lexico-syntactic patterns are defined. These are implemented as regular expressions detecting direct and indirect queries and identifying specific linguistic expressions. From the analysis of the reference corpus we observed that many users do not just make specific requests, but also use general queries such as "*dammi altre informazioni*" (give me more information), "*mi dici qualcosa sul quadro*" (can you tell me something about the picture).

4.2 Device commands management

Starting from the material collected with the procedure described in Section 2, an expert linguist analysed the output of the Treetagger and TULE tools applied to the normalised utterances to obtain the Opendial model dedicated to command recognition. Rules in this model attempt to match recurring syntactic patterns

related to specific user requests against the dependency parsing tree obtained from the received utterance. The rules consider the presence of a variable length syntactic dependence, in the tree, between a set of target subtree roots (usually verbs), covering the ones observed in the corpus, and, possibly, a set of target terms (usually nouns) in the tree. For both target roots and terms, it is possible to admit their MWN synsets. It is possible to summarize the structure of these rules as a tuple $< R, T, l, S >$, where R is a set of target subtree roots, T is a set of optional target terms, l is the maximum length of the dependency chain linking a member of R to a member of T and S is a subset of the union of R and T containing the terms for which synonyms are accepted. If T is empty, l is always 0 and the terms included in R must be the root of the entire tree for the rule to be matched. Multiple tuples can be associated with a single command to describe different syntactic patterns. For example, the *TakePicture* command is associated with the tuples

$$< \{scattare, \ldots, foto\}, \{\}, 0, \{scattare, \ldots, foto\} >$$
$$< \{fare\}, \{foto\}, 2, \{foto\} >$$

The first tuple handles the cases in which users use isolated words, like imperative verb forms, to control the device ("scatta!", "fotografa!", "foto!"), the second tuple handles the case of the most natural sentence "fai una fotografia" (Take a picture) where *fotografia* is a synonym of *foto*(photo) and also covers the frequent use of *Googlese*(talk using keywords) by the users as, in the corpus, utterances like "fai foto" were not uncommon. This is also the reason why dependency types are not included in the rules: we observed that, when users shift to *Googlese*, the structure of the dependency tree is preserved in most cases but the reported relationship types are erratic.

When no command pattern can be matched exactly, the system checks for partial matches in the preceding rules represented by the presence of target terms in the input utterance to recover the interaction. In this phase, the system also checks the active processes to reduce the set of possibilities when binary actions are considered. For example, if the target term "registrazione" is detected and the device is already recording a video, the only possible action related to the target term is *RecStop*, which stops the recording. The system asks for confirmation before performing the action as a safety measure. When context does not help to reduce the set of possible actions to one, as in the case of "Avvia!"(Start) when both video recording and the audioguide are not running, the system prompts the user to specify an action among the possible ones to proceed. When system prompts conclude an interaction turn, control is given directly to the system query handler to complete the interaction after processing the user utterance answering the question.

4.3 Question answering

In this section we introduce the approach used to understand the user's request and generate the response that satisfies it. First, we check if the sentence is a Wh-query type. Then, if the check has success, a set of rules are applied to detect the

topic and the information request about topic. Discovered the type of information request, queries are generated to research answers in a knowledge base. This latter is composed by stories that have been annotated through an event-based formalism. Finally, texts and media retrieved are composed to generate an unique multimedia response to be returned to the user.

For the extraction of the topic and the arguments of the request, we have defined a set of rules based on the relations contained in the dependency tree of the sentence and contextual information of user's position. The rules defined are used for discovering the topic (subject which is discussed) and about what the visitor asks. The recognition process identifies the components of the query using the same semantics adopted to annotate the story of an artwork. For this, analysing the dependency tree and the topic of the query, we discover the events from their components: the action (what), the location (where), the participants (who), and the interval of happening (when).

For example: "*Chi ha dipinto la Morte di Cesare*(Who painted Cesar's Death)" is related to the action "dipingere (paint)" and participant "la morte di Cesare". Also, the visitor want know the author ("Chi (Who)"). In this example, the rules used for detected the action and the participant are:

$$verb(S,V) \wedge \sim auxiliary(S,V) \rightarrow action(S,V)$$

$$verb(S,V) \wedge obj(S,V,O) \rightarrow participant(S,O)$$

where $verb(S,V)$ means that V is the verb of the sentence S, $auxiliary(S,V)$ means V is an auxiliary verb of S and $obj(S,V,O)$ means O is the object if V in S. In this example, the topic is explicit and corresponds to the object of verb. If the query has a passive form ("*da chi stata dipinta la Morte di Cesare*") the participant is the subject. There are other rules to discovery the location and/or time interval of the event. If the topic is not explicit, is detected from contextual information analysing the user's position.

To assemble the multimedia response, we first have to obtain a textual answer. To do this, we adopt a modified version of the system presented in [12]. All the histories and media relative to an exhibit are archived in a semantic repository annotated with an event-based formalism. In this work, this formalism is the Cultural Story Web Language (CSWL)[13], which represents. Starting from the results of query interpretation, we reformulate the user request as a CSWL query, and apply query expansion using semantic lexical databases (MultiWordnet and Wiktionary). The list of answers (events) obtained from query results are ranked and the best answer is expanded through extra events correlated with it, if necessary. Then, the corresponding text associated to the selected events, and the relative media recovered in the repository, are assembled as a presentation. The process takes in account the semantic annotation associated to the elements (texts and media) and synchronizes them so that media items are visualised coherently with the relevant time instants in which a synthetic voice is talking about the content it represents[14].

5 Conclusions

We presented the design and implementation of a spoken dialogue system for Italian aiming to assist the visitors of a museum exhibit. We reported the difficulties encountered in designing an interactive system using natural language understanding caused by the particular environment of museums located in historical places and how we addressed them. We also described the material collected to obtain a first working prototype of the CHEESE system. The system architecture is flexible, modular and can easily be adapted to future updates due to upcoming technologies. Also, spoken dialogue systems for Italian are not common and our contribution explores an area of recent interest for interactive environments, cultural heritage, which may find relevant applications in Italy. Future work will consist in collecting on-site feedback to evaluate the system and take full advantage of the probabilistic environment offered by Opendial to fine-tune the system.

References

1. Baggia, P., Cutugno, F., Danieli, M., Pieraccini, R., Quarteroni, S., Riccardi, G., Roberti, P.: The multi-site 2009 evalita spoken dialog system evaluation. In: Proc. of EVALITA. (2009)
2. Sorgente, A., Brancati, N., Giannone, C., Zanzotto, F.M., Mele, F., Basili, R.: Chatting to personalize and plan cultural itineraries. In: UMAP Workshops. (2013)
3. Stock, O.: Language-based interfaces and their application for cultural tourism. AI Magazine **22**(1) (2001) 85
4. Stock, O., Zancanaro, M.: PEACH-Intelligent interfaces for museum visits. Springer Science & Business Media (2007)
5. Bastianelli, E., Nardi, D., Aiello, L.C., Giacomelli, F., Manes, N.: Speaky for robots: the development of vocal interfaces for robotic applications. Applied Intelligence **44**(1) (2015) 43–66
6. Lee, A., Kawahara, T.: Recent development of open-source speech recognition engine julius. In: Proc. of APSIPA ASC. (2009) 131–137
7. Lison, P.: A hybrid approach to dialogue management based on probabilistic rules. Computer Speech & Language **34**(1) (2015) 232 – 255
8. Schmid, H.: Improvements in part-of-speech tagging with an application to german. In: Proc. of the ACL SIGDAT-Workshop. (1995) 47–50
9. Lesmo, L.: The rule-based parser of the nlp group of the university of torino. Intelligenza artificiale **2**(4) (2007) 46–47
10. Pianta, E., Bentivogli, L., Girardi, C.: Developing an aligned multilingual database. In: Proc. 1st Intl Conference on Global WordNet. (2002)
11. Rossano, F.: Questioning and responding in italian. Journal of Pragmatics **42**(10) (2010) 2756 – 2771
12. Mele, F., Sorgente, A.: Semantic mashups of multimedia cultural stories. Intelligenza Artificiale **6**(1) (2012) 19–40
13. Mele, F., Sorgente, A.: CSWL - un formalismo per rappresentare storie culturali nel web. Technical Report 180/15, Inst. of Cybernetics "E. Caianiello", CNR (2015)
14. Sorgente, A., Vanacore, P., Origlia, A., Leone, E., Cutugno, F., Mele, F.: Multimedia responses in natural language dialogues. In: Proceedings of AVI*CH 2016. Volume 1621 of CEUR Workshop Proceedings., CEUR-WS.org (2016) 15–18

Multi Attributes Approach for Tourist Trips Design

Ilaria Baffo[1], Pasquale Carotenuto[2], Antonella Petrillo[3], Fabio De Felice[4]

[1] Industrial Engineering School (DEIM) - University of Tuscia, Italy
ilaria.baffo@unitus.it
[2]Istituto per le Applicazioni del Calcolo (IAC) - CNR, Italy
p.carotenuto@iac.cnr.it
[3]Department of Engineering - University of Naples "Parthenope", Italy
antonella.petrillo@uniparthenope.it
[4]Department of Civil and Mechanical Engineering - University of Cassino and Southern Lazio, Italy
defelice@unicas.it

Abstract. The authors propose a Multi Attributes approach to meet the demand of personalized tourist tours into cultural cities. Respecting to others works present into the literature, in this paper the decisional process includes two phases and a high number of variables that don't increase the complexity of the problem. A real application in an Italian city, Florence, is presented to demonstrate the great potential of this system into real context. The first phase of optimization is solved applying an innovative Genetic Algorithm, the second one a Multi Criteria Method, Analytic Hierarchy Process (AHP). The combination of these two approach gives flexibility to the system with respect to number of variables and allow to return a good solution for tourist in few second of computational time.

1 Introduction

The increasing use of mobile devices in everyday life has favored the creation of intelligent applications also in field such as tourism, culture and hobbies. One of the most interesting applications provides opportunity to help tourists in choosing the best route given a maximum time of visit. During the last years several researchers have faced this problem with the aim to reduce uncertainty and increase the personalization of the tourist paths. In the scientific literature, the problem is better known as Tourist Trip Design Problem (TTDP). The authors in this work propose a Multi Attributes Decision Support System (MDSS) to meet the demand of personalized tourist tours into cultural cities. As described in the following many researchers have successfully addressed this problem, but, only few of these have proposed real applications able to consider several attribute to decisional process. The Multi Attributes Decision Support System proposed, facing the TDPP in two steps. Firstly, a Genetic Algorithm (GA) solves in optimal way the Orienteering Problem (OP) on real instances coming from Italian historical cities, Florence. The validation of algorithm is guaranteed creating an A Mathematical Programming Language (AMPL) model able to return the optimal solution for small instances of the same problem. Finally, the optimal solutions are compared on basis of different attributes having the features to be added

up. The last step has been conducted with the use of Multi Criteria Analysis Methods. Thank to this approach, the obtained solutions can be evaluated on basis of several *attributes* like: time of decision with respect to time of travel, duration of travel, levels of freedom into the decision, degree of additionality of preferences, number of decision-makers, numbers of cultural heritages to visit, etc. Real cases studies are given with the aim to offer development ideas for implementation of smart applications for tourist trips' design. The proposed methods can reach important results where the variables of the problem are numerous and the optimization process too long and complex. The rest of paper is organized as follows: a review of literature is given in section 2, the mathematical model and its resolution with genetic algorithm is explained in section 3. Section 4 presents AHP approach and its integration with optimization model is discussed. Application and results are analyzed in Section 5. Finally, conclusion and future researches are proposed in section 6.

2 Literature Review

As well explained in [5,10] the described route-planning problem has been considered as an application of the Orienteering Problem or its variants. In the OP, several locations with an associated score have to be visited only once in order to obtain a total trip score. The objective is to obtain a total trip score as high as possible without violating a time restriction. There have been works on exact methods that have yielded solutions to smaller sized problems. Due to the computational limitations of the exact algorithms, several heuristic were explored to many researchers in the past. The first heuristics were proposed by Tsiligirides (1984) in [9] and are known as the S-algorithm and the D-algorithm. The S-algorithm uses the Monte Carlo method to construct routes using probabilities correlated to the ratio between node score and node distance from the current node. The D-algorithm is built based upon the vehicle scheduling method proposed by Wren and Holiday (1972) [12]. Others heuristics have been explored in the years to solve this kind of problems, in [8] a Tabu Search approach has been used for Multi Constrained Team Orienteering Problem and its application in touristic trip planning. In [7,11] the authors propose to solve the Multi–Constraint Team Orienteering Problem with Multiple Time Windows with known heuristic like Iterated Local Search (ILS) and Greedy randomized adaptive search procedure (GRASP). Iterated Local Search is proposed also in [4] to solve real-time Time Dependent Team Orienteering Problem with Time Windows (TDTOPTW). During the last years, many authors have chosen to implement evolutionary algorithm trying to reduce the computational time and to improve the quality of proposed solution to tourist. In [2] the authors propose an evolutionary algorithm to solve a multi-objective problem for personalized tours in street networks. In this case, the trip is long multiple–day and a Team Orienteering Problem with Time Windows (TOPTW) is solved with a heuristic approach whose computational times are not clearly declared. The aim of the authors is to give in few second an optimal solution of OP and secondly introduce more elements that can be analyzed with multi criteria analysis. In this way, the complexity of the problem is lower and the computational time too, then the approach is more usable for real time digital applications.

3 Problem Description

In this work the authors face the single day Tourist Trip Design Problem (TTDP) modelling an Orienteering Problem (OP). In the OP, a set of *N* location, also called Point of Interest (PoI*s*), corresponding with nodes *i* is given, each with a score S_i. The starting point and the end point are fixed. The time t_{ij} needed to travel from vertex *i* to *j* is known for all vertices. The problem became more complex because not all the vertices can be visited since a given threshold T_{max} limits the available time. The goal of the OP is to determine a path, limited by T_{max} that visits some of the vertices, in order to maximize the total collected score. Other works propose to optimize the problem maximizing the number of PoI*s* visited, the problems coincide if we suppose each point's score equal to 1. The scores are assumed to be entirely additive and each vertex can be visited at most once. In this algorithm the maximum time available to visit the PoI*s* is considered as constrain of the problem. At the same time, all solutions are defined by a set of PoI*s* visited, a time of visit for each point and a time of walking from one point to another point. The OP can be formulated as follows: $S_i \geq 0$ is the score associated to node *i*, c_{ij} is the cost associated to path between node *i* and node *j*. Usually *n* nodes are considered in the Euclidean plane. Since the distance and travel time between nodes are determined by the geographical measure, distance is used as the representative of path's cost. Generally, the mathematical model of the OP is formulated as follows:

$$MAX \sum_{i=1}^{n} \sum_{j=1}^{n} S_i \, x_{ij} \tag{1}$$

Subject to

$$\sum_{j=2}^{n} x_{1j} = \sum_{i=1}^{n-1} x_{in} = 1 \tag{2}$$

$$\sum_{i=2}^{n-1} x_{ik} = \sum_{j=2}^{n} x_{kj} \leq 1 \qquad k=2,\ldots,n-1 \tag{3}$$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} \leq T_{max} \qquad \forall i, j=1,\ldots,n \tag{4}$$

$$2 \leq u_i \leq n \qquad \forall\, i = 1,\ldots,n \tag{5}$$

$$u_i - u_j + 1 \leq (n-1)(1-x_{ij}) \qquad \forall\, i,j = 1,\ldots,n \tag{6}$$

$$x_{ij} \in \{0,1\} \qquad \forall\, i,j = 1,\ldots,n \tag{7}$$

Where (1) represents the problem's Objective Function to maximize the total collected score S. Constraint (2) guarantee that the path starts in vertex 1 and ends in vertex N. Constraint (3) ensures the connectivity of the path and guarantee that every vertex is visited at most once. Constraint (4) ensures the limited time budget T.

Constraints (5) and (6) are necessary to prevent sub-tours. Constraints (7) requires that the variables are binary. The decisional variable x_{ij} is equal to 1 if the node *j* to *i* are connected, 0 else. This formulation guarantees as result a tour able to: Visit as many PoI*s* as possible; Visit PoI*s* at most once; Visit PoI*s* that maximize the Total Score (Objective Function); Visit PoI*s* connected among them; and Visit PoI*s* respecting the limitation time. The set of solutions coming from the genetic algorithm represent the inputs for multi criteria approach with AHP.

4 Multi criteria approach: Analytic Hierarchy Process.

The AHP breaks down a decision-making problem into several levels in such a way that they form a hierarchy with unidirectional hierarchical relationships between levels [3]. The AHP for decision making uses objective mathematics to process the inescapably subjective and personal preferences of an individual or a group in making a decision. With the AHP, one constructs hierarchies or feedback networks, then makes judgments or performs measurements on pairs of elements with respect to a controlling element to derive ratio scales that are then synthesized throughout the structure to select the best alternative. The top level of the hierarchy is the main goal of the decision problem. The lower levels are the tangible and/or intangible criteria that contribute to the goal. The bottom level is formed by the alternatives to evaluate in terms of the criteria. The modeling process requires a *pairwise comparison*s of the elements in each level using a scale of 1-9, as suggested by Saaty [6]. The result of the comparison is the so-called dominance coefficient a_{ij} that represents the relative importance of the component on row (*i*) over the component on column (*j*), i.e., $a_{ij}=w_i/w_j$. The pairwise comparisons can be represented in the form of a matrix. After all pairwise comparison is completed, the priority weight vector (*w*) is computed as the unique solution of $Aw=\lambda_{max}w$, where λ_{max} is the largest eigenvalue of matrix A. Finally, consistency index CI is estimated. CI could then be calculated by: $CI=(\lambda_{max}-n)/n-1$. In general, if *CI* is less than 0.10, satisfaction of judgments may be derived. For this application the AHP takes as input the results of optimization process and evaluates these under several criteria better explained in the next paragraph. As first activity, the mathematical model is supposed to model the capacitated orienteering problem, as second phase we solve the formulation implemented a meta-heuristic known as Genetic Algorithm (GA). Then, we applied a multi-criteria analysis (AHP model) to algorithm's results to order the founded solutions respect to the following objectives: Time L-path, Visited places, Ticket cost.

5 Problem Resolution

The authors propose an optimization approach that draw inspiration on Genetic Algorithm (GA) as developed by Holland (1975) and then described by Goldberg (1989). GA starts initializing a First Population composed by a pre-determinate number of individuals. Each individual represents a solution for the faced problem and it is characterized by several elements such a chromosome and a fitness value.

The individual's chromosome is composed by several gene that can assume binary or integer value, the combination of these genes usually represents the problem's solution thanks to a correspondence between the gene's value and the problem's variables value. The individual's fitness value corresponds to solution objective function's value, and it's basic into determinate the individual's probability to survive at evolution process, so that bad individuals are destined to not have a long life into population. The difference between the old and the new population in the evolutionary process is guaranteed by the presence of two important operations called Mutation and Crossover. After several generations, the best solution converges, and it hopefully represents the optimum or suboptimum solution to the problem at hand. GA are successfully applied to different contexts in order to solve a very large number of problems, for this purpose various and original genetic operators are developed, that revisit the original concept of mutation and crossover. In this application the authors adopt a circular representation of chromosome and several heuristic techniques to improve the evolutionary process through mutation and crossover operations. In our GA the chromosome is composed of all nodes that the tourist want visit, without the start node and the end node that can be added at the end of optimization process. This coding of chromosome with respect to problem's solution representing by a single tour, allows to have only admissible solutions better or worst on basis of Total score reached by visiting more or more better PoI*s*. The Fig. 1 shows like each chromosome represents a sequence of visiting PoI*s*, i.e. a tour, i.e. a solution for the OP. The Closed loop Structure of Chromosome code allows a wider and faster research of solutions.

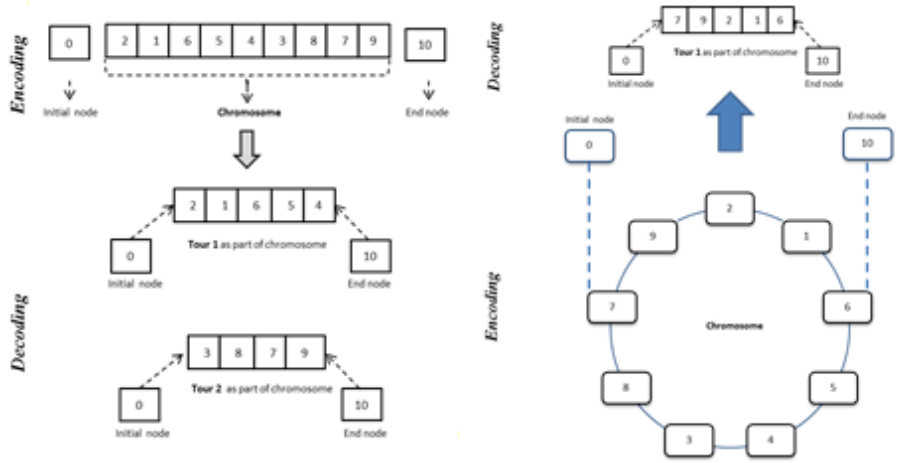


Fig. 1: Encoding and Decoding process

As previously described the GA structure is based on constitution of a population at each iteration. This population step by step evolves and the algorithm found better solutions for the faced problem. Following a synthetic description of the implemented GA pseudocode is reported.

```
Parameters Setting
Initialization
   Generate a feasible solution randomly as individual
```

```
   Save them into the population Pop (i)
   Loop until the population's size S is reached
   Calculate Fitness to define the Best Individual BI(i) of P(i)
   While the number of iterations isn't reached do
Set iteration i = 0
   Selection: Roulette function for selecting two individual as parent
   Crossover (Single, PMX)
   Evaluate Fitness
   Update the Best individual when offspring is better than actual Best
   Individual BI(i)
   Mutation (Smart Swap)
   Evaluate Fitness
   Update the best individual when offspring is better than actual Best
   Individual BI(i)
Build
   Build new Population Pop (i+1) as collection of individuals generated
   with Crossover and Mutations Operations
Next i
End - Return the best individual containing the best tour
```

More details are given into a previous publication of authors in [1]. The solutions given by optimization process constitute inputs for Multi Criteria Analysis Approach. Finally, the Multi Attributes Approach for Tourist Trips Design Problem (MATTDP) gives as output a solution that takes into account a lot of variables, some of these are considered into optimization process and others into Multi Criteria Analysis Process.

5.1 Real Application

In this paragraph, we propose the results obtained for the single-day Tourist Trip Design Problem (TTDP) in the city of Florence. We consider 24 PoI*s*, all included into urban areas of the city. We fix the PoI*s* of departures and of arrival and the maximum time that the tourist has to visit the city, variable from 390 minutes to 930 minutes. We calculated the distance matrix and set the algorithm parameters (numbers of individuals, population and Best Solution Containers's (BSC) size, Frequency rate of genetic operators) on the basis of preliminary tests that have given the best configuration of algorithm. For each PoI, we suppose that the tourist assigns a different score on the basis of your preference. The Total time of tourist path is calculated as summarize between the time to reach the PoI and the time to visit it that it is supposed to be equal to 30 min for each one. The Table 2 presents the results coming from the optimization process; the authors found 5 solutions/tours (called 1, 2, 3, 4, 5) for 5 different time each one starting from PoI 1 and ending to PoI 24. Each PoI is represented by a number from 1 to 24 and each Tour can be considered as a sequence of visited PoIs. As showed in Table 1 for each solution a Real time of tour, a total time of tour, a total score of tour and the number of visited PoI*s* are given. T_{max} real is expressed in minutes like Time for visit the PoI*s* and the relative Total Time. T_{max} is expressed in a conventional measure used into the algorithm, Total Score and Visited PoIs have no units of measure. For summarizing, each solution coming from optimization process is characterized by a number or a sequence of visited PoI*s*, and a Total time needed for completing the tour as sum of time to reach all PoI*s* of tour and time to visit them.

Table 1: Data of solutions

Solutions	1	2	3	4	5
Tmax	5	8	10	15	20
Real Tmax	60	96	120	180	240
Time of visit	330	450	540	660	690
Total time	390	546	660	840	930
Total score	75	110	145	170	185
Visited PoIs	11	15	18	22	23

The outputs of optimization process as before described are given as inputs o**f** AHP model built as in Fig. 2. In the present paper AHP Absolute model is applied as in [**3**].

Table 2: Solutions of optimization

Solutions	Tour
1	1 10 7 18 3 2 5 6 4 23 24
2	1 20 18 12 2 3 10 4 6 5 8 17 22 23 24
3	1 18 12 14 9 3 5 2 6 4 8 17 22 21 23 16 11 24
4	1 14 19 18 12 2 3 10 17 21 22 7 5 6 9 8 4 23 16 11 24
5	1 2 9 6 5 3 7 12 10 18 20 19 22 4 17 8 21 23 16 11 14 24

AHP Absolute model is based on paired comparisons among the elements of a set with respect to a common attribute. Experts team developed pairwise comparison matrices to determine the weights of criteria. Consistency index has been estimated (CI 0.018). Results show that the most important criteria is "C1. Time L_{path}" with a score of 56%, followed by "C3. Ticket cost" with a score of 32%, and finally is the criteria "C2. N° Visited places" with a score of 12%. In AHP Absolute model criteria are further subdivided into a level for intensities. The scores of these intensities are each weighted by the priority of its criterion and summed to derive a total ratio scale score for the alternative. Each criterion has ratings listed under it. Tab. 3 shows the final ranking of the AHP Model.

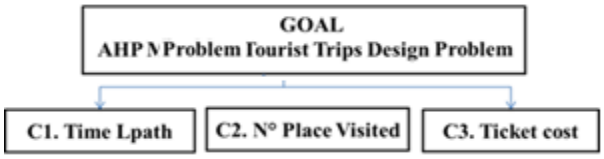


Fig. 2. AHP Absolute Model

	Priorities	C1 Time Lpath 0.5584	C2 N* Place Visited 0.1219	C3 Ticket cost 0.3196
5	0.3750	0.5 km	11	10-20 €
8	0.2423	1 km	15	20-40 €
10	0.1599	1.5 km	18	40-60 €
15	0.1174	2 km	22	60-80 €
20	0.1052	2.5 km	23	80-90 €

Tab. 3. Final Ranking

6 Conclusions and Future Researches

In this work the authors present a Multi Attributes Approach for solving single-day Tourist Trips Design Problem (MATTDP). The proposed System is applied to a real case into the Italian city of Florence. Thanks to two phases of this approach it is possible to include several variables into the decisional process without increase the complexity of the problem. For future researches, the authors want to implement the system in a big instance of the problem considering also time windows constraints for visiting PoIs. The next step could be to implement a mobile application for smartphone and tablet in order to test the real usability of this system in real context.

References

1. Baffo, I., Carotenuto, P., Storchi, G., A genetic algorithm to design touristic routes in a bike sharing system, 14th international conference on modeling and applied simulation, At Bergeggi (Liguria) – Italy, (2015).
2. De Falco, I., Scafuri, U., & Tarantino, E. A multiobjective evolutionary algorithm for personalized tours in street networks. In European Conference on the Applications of Evolutionary Computation. Springer International Publishing. (2015) 115-127.
3. De Felice F., Petrillo A., Absolute measurement with analytic hierarchy process: A case study for Italian racecourse. International Journal of Applied Decision Sciences. Volume 6, Issue 3, 2013, (2013) 209-227.
4. Garcia, A., Arbelaitz, O., Linaza, M. T., Vansteenwegen, P., & Souffriau, W. Personalized tourist route generation. In International Conference on Web Engineering. Springer Berlin Heidelberg. (2010) 486-497.
5. Gavalas, D., Konstantopoulos, C., Mastakas, K., & Pantziou, G. A survey on algorithmic approaches for solving tourist trip design problems. Journal of Heuristics, 20(3), (2014) 291-328.
6. Saaty, T.L., (1977. A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology 15, 234–281.
7. Souffriau, W., Vansteenwegen, P., Vanden Berghe, G., & Van Oudheusden, D. The multiconstraint team orienteering problem with multiple time windows. Transportation Science, 47(1), (2013), 53-63.
8. Sylejmani, K., Dorn, J., & Musliu, N. A tabu search approach for multi constrained team orienteering problem and its application in touristic trip planning. In Hybrid Intelligent Systems (HIS), 12th International Conference on IEEE. (2012) 300-305.
9. Tsiligirides, T., Heuristic Methods Applied to Orienteering. Journal of Operational Research Society, vol. 35, no. 9, (1984) 797-809.
10. Vansteenwegen, P., Souffriau, W., & Van Oudheusden, D. The orienteering problem: A survey. European Journal of Operational Research, 209(1), (2011) 1-10.
11. Vansteenwegen, P., Souffriau, W., Berghe, G. V., & Van Oudheusden, D. Metaheuristics for tourist trip planning. In Metaheuristics in the Service Industry. Springer Berlin Heidelberg. (2009) 15-31.
12. Wren A., Holiday A., Computer Scheduling of Vehicles from One or More Depots to a Number of Delivery Points. Operations Research Quarterly, vol. 23, (1972) 333-344.

Integrating Archaeological Datasets: the ARIADNE Portal

Paola Ronzino[1], Achille Felicetti[1], and Sara Di Giorgio[2]

[1] PIN, Polo Prato, Italy
{paola.ronzino, achille.felicetti}@pin.unifi.it
[2]Istituto Centrale per il Catalogo Unico delle biblioteche italiane, Roma, Italia
{sara.digiorgio@beniculturali.it}

Abstract. One of the emerging needs of the archaeological community is represented by the importance of availing of systems that allow to tackle new research questions, by querying diverse available resources. Usually, archaeological digital data is stored in non-standardised individual databases with a limited possibility of integration and a high level of fragmentation. The EU-funded project ARIADNE, has developed an e-infrastructure which enables the integration of archaeological datasets from various different institutions, integrating resource discovery metadata using controlled vocabularies, thesauri, gazetteers and ontology (CIDOC CRM). This paper presents the ARIADNE infrastructure, describing the activities undertaken by the project to achieve interoperability of archaeological resources at the dataset and item level. Moreover, the architecture of the ARIADNE Infrastructure and the Portal, with the different ways to search and access the resources are described.

1 Introduction

In the recent years we have assisted to an increasing awareness of the importance of creating networks of data that allow integrated access to documentation and to digital archives of archaeological resources. An important condition for the development of such networked accesses lays in the definition of standards and guidelines that establish a degree of compatibility between the datasets that make these networks up.
Usually, data is stored in non-standardised individual databases with a limited possibility of integration and a high level of fragmentation of data. This is mostly due to the different needs of the various research communities who store and structure their data according to the standards that apply to their specific research domain. However, when the different communities agree to share their data with the wider community and for a broader purpose, the related problem of data interoperability arises. This is the challenge that ARIADNE is facing [1, 2]. The EU-funded project has developed an e-infrastructure that enables the integration of archaeological datasets from various different institutions. ARIADNE's main objective is to provide researchers with an integrated access and to guarantee the semantic interoperability of archaeological datasets distributed throughout Europe. The main expectation of the project is that researchers will make use of these resources and benefit from them through the use of technologies and services made available by the infrastructure itself, and to challenge

new scientific questions. In the following sections we will introduce the consortium and the content made available into the ARIADNE infrastructure, touching on the user requirements that influenced the structure of the portal. Moreover, the paper describes the overall architecture of the infrastructure and the core services, which make ARIADNE a powerful system for the archaeological research community for sharing, discovering, accessing and reusing available data.

2 The ARIADNE Research Infrastructure

The ARIADNE infrastructure is supported by a consortium of archaeological institutes and data archives providing content, and by technology developers.
The consortium consists of 23 partners in 16 countries, and a number of associate partners that ensure an almost complete coverage of the European territory.
Most of the archaeological institutes involved in ARIADNE started to integrate archaeological datasets under a common portal, responding to the need of digital preservation, open access and networked access. Part of the network (to cite a few) are the UK's Archaeological Data Service (ADS) [3], which currently provides access to over 36,000 unpublished fieldwork reports and over 1000 data digital archives, the Data archiving and Networked Services (DANS) [4] providing access to over 21,000 reports and 4,000 excavation archives of Dutch archaeology, the Swedish National Data Service (SND) [5], based at the University of Gothenburg, the CulturaItalia Portal of the Italian Ministry of Cultural Heritage (MIBACT-ICCU) [6].
Content provided by partners was necessarily created and documented in different ways, using different languages and encoded by means of different metadata schemas. This obviously makes data integration a complex process.

2.1 Gathering user requirements

Before starting the ARIADNE infrastructure design, the project carried out various surveys and interviews to the archaeological research community, to find out the existing and emerging needs, so that the infrastructure would build on their basis. The research consisted of an extensive literature review, numerous interviews involving members of the ARIADNE partners and other stakeholders, two online questionnaire surveys with participation of more than 600 archaeological researchers and repository managers [7]. Subsequently, a survey of existing data portal involved 23 ARIADNE archaeological researchers and data managers, to get further insight for the development of the ARIADNE portal services [8]. The result of the two international online surveys made clear that archaeological researchers lack appropriate data repositories and services that allow to find and access relevant data. The 95% of the responses expressed that the most important need is to have a comprehensive overview of the available datasets. Most researchers asked for a data portal that offers an overview of the resources available online, with the possibility to search across resources scattered in different places, using new mechanism to discover and access data. For space reason we limit the description of analysis result to the top-level needs. The detailed

analysis of the user requirements can be found at [7]. The recommendation and the user needs collected through the activities presented above, influenced and guided the technology partners in the design and implementation of the infrastructure architecture, of the portal interface and of the services.

2.3 The ARIADNE Catalogue

The first step towards data interoperability consisted in a preliminary analysis of the available archives, in order to identify formats, standards, and services in use by the content providers [9]. Key elements, common to all archives, were identified and encoded using existing international standards and terminological tools, and referring to the "what, where, when" paradigm. The descriptions of content have been encoded using the ARIADNE Catalogue Data Model (ACDM) [9], developed by ARIADNE with the aim to produce a detailed representation of the archaeological information of the legacy archives made available by the consortium.
The ACDM was built on the DCAT vocabulary [10] and was extended with classes and properties needed to better describe the ARIADNE archaeological resources.
The main classes, which reflect the ARIADNE assets, are: *DataResources* (including the resources that are the containers of the data, like databases and collections), *LanguageResources* (such as vocabularies and metadata schemas) and *Service* (owned by the ARIADNE partners and offered to the project for integration).
The ARIADNE Catalogue aggregates metadata describing datasets, metadata schemas, vocabularies, which were harvested either manually or trough the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). The metadata and object repository aggregator (MORe), was customized to meet the ARIADNE content needs and used to aggregate metadata from multiple sources and in multiple formats. The Catalogue, and the detailed information it contains, represents the core of the entire integration process. The Catalogue lays at the base of the portal, and provides the necessary support for retrieval, analysis and resource discovery facilities.

2.3 Datasets integration

The highest level of integration started at the conceptual level, by identifying key elements, common to each archive. The "what, where, when" paradigm was used to identify objects, places and time periods, which are fundamental elements in the archaeological domain.

2.3.1 The WHAT

The "what" represents the subjects of the various datasets. These are described using terms derived from the Art and Architecture Thesaurus (AAT) [11] of the Getty Research Institute, which was adopted as the spine for the whole framework of terms in ARIADNE. Each of the terminological resources used by ARIADNE content providers was mapped to the AAT concepts to demonstrate the semantic and conceptual similarity between the different archives. The mapping activities were facilitated by a

mapping tool developed to establish correspondences between concepts coming from different vocabularies [12].

2.3.2 The WHERE

The "where" represents the spatial entities. Most of the archaeological archives had already standardized spatial information. To enable browsing the archives, ARIADNE recommended content providers to provide geographic information in the WGS84 format. When the only information available was the name of a place, the spatial coordinates were retrieved through the GeoNames gazetteer.
Geographic information about historic names were retrieved from Pleiades [13], through a collaboration between ARIADNE and the Pelagios project.

2.3.3 The WHEN

The "when" represents temporal entities. When the dates were expressed in numeric format, the temporal integration was easy to manage. When, instead, the periods were indicated as names, for example Bronze Age, this caused a lot of ambiguities because these were not referred to absolute dates. Beside converting each period in absolute time spans, the collaboration with the PeriodO project [14], allowed to manage collections of periods as intersections of documented events on specific geographical areas providing unique identifiers for each of those periods as Linked Open Data.

2.4 Item-level Integration

The possibility to answer a research question by using relevant information from several available heterogeneous sources, is one of the emerging needs of the archaeological community. To address the complexity of archaeological data integration, ARIADNE developed a global, extensible schema as a formal ontology to allow for integration without loss of meaning. The CIDOC CRM ontology [15] was chosen as the backbone of the ARIADNE Reference Model, which includes a suite of extensions developed to address the complexity of archaeological data integration.
CIDOC CRM (ISO21127) is a formal ontology created to facilitate the integration, mediation and interchange of heterogeneous cultural heritage information. It contains the basic relationships needed to describe what happened in the past, as for example people and things meeting in space-time, parts and wholes, use, influence and reference.

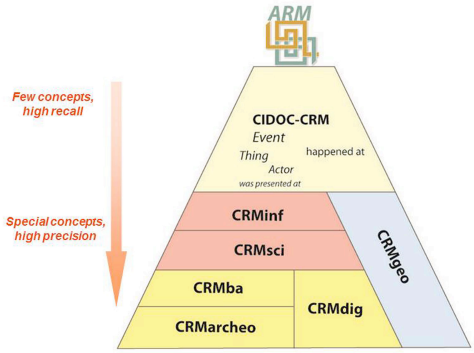


Fig.1 The ARIADNE Reference Model

The ARIADNE RM, presented in Figure 1 includes the following extensions:

- CRMinf [16]: is a formal ontology intended to be used as a global schema for integrating metadata about argumentation and inference making in descriptive and empirical sciences.
- CRMsci [17]: is a formal ontology intended to be used as a global schema for integrating metadata about scientific observation, measurements and processed data in descriptive and empirical sciences.
- CRMgeo [18]: is a spatiotemporal model that provides a link between the standards of the geospatial and the Cultural Heritage community in particular between GeoSPARQL and CIDOC CRM.
- CRMdig [19]: is an ontology to encode metadata about the steps and methods of production (provenance) of digitization products and digital representations such as 2D, 3D or animated models created by various technologies.
- CRMba [20]: the Buildings Archaeology is an ontology developed for investigating historic and prehistoric buildings, the relations between building components, functional spaces, topological relations and construction phases through time and space.
- CRMarchaeo [21]: the Excavation Model is an ontology to encode metadata about the archaeological excavation process.

With the definition of the ARIADNE RM, and the creation of an integrated knowledge base, the aggregation of several existing archaeological databases were transformed by mapping their individual schemas to the ARIADNE RM. The mapping process was supported by the X3ML Mapping Framework, a tool that ensures the integrity of the initial data and preserving the original meaning [22].
An advanced level of the interoperability was achieved with the integration of individual records of legacy archaeological archives. This item-level integration experiment had the aim to reach the deepest integration of archaeological data. Preparatory activities towards this goal included mappings with specific tools which allowed individual partners to track complex correspondences between the entities contained in their databases and the conceptual classes provided by the CIDOC CRM and its extension. Conceptual mappings for each partner's archives enabled the creation of

semantic representations for individual items in RDF, to form a complex graph of relationships to be queried, integrated with semantic technologies and published in Linked Open Data format. ARIADNE chose as a use case the numismatics field, to demonstrate the item-level integration process of archaeological datasets. Five datasets were selected. Four of them were mapped to the ARIADNE RM and transformed to RDF using the X3ML framework, while the fifth was already in CIDOC CRM RDF form, and therefore, compatible with the ARIADNE RM. As a common thesaurus for the aggregated knowledge base, the nomisma.org, and the AAT thesauri were adopted. The mapping and transformation workflow is presented in Figure 2. The main goal of the integration of the diverse coin datasets was to create a system enabling users to specify queries that will be evaluated on the common aggregated repository and will be able to combine results coming from the different datasets. The ARIADNE portal provides a main access point to integrated repository and an intuitive user interface will guide the user to formulate the query, browse the results and refine the search with facet view.

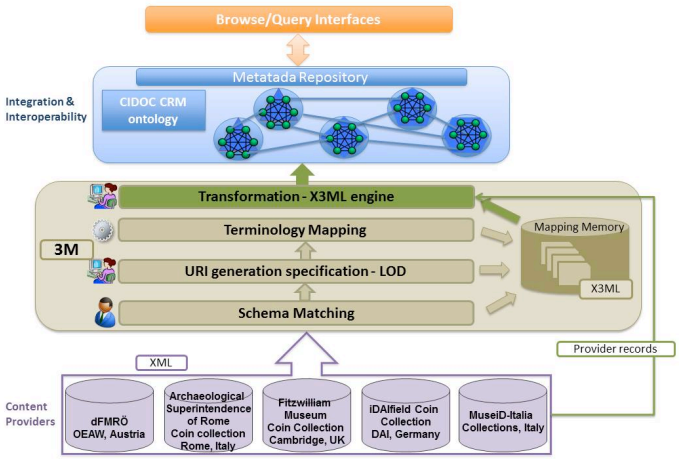


Fig.2 Mapping and transformation workflow

3 The Integrated Infrastructure and the ARIADNE Portal

The ARIADNE Portal represents the highest level of the infrastructure. It is the access point of the whole infrastructure, where users can browse, query and analyse available data and use the services to activate all the features provided by the system.

The integration platform designed and implemented by ARIADNE appears as a complex modular system, with advanced interfaces and features and an architecture able to interact with distributed archives, in a transparent way. The system is able to query and extract integrated information concerning legacy archives, to present them to users by means of advanced services and tools to visualize, analyse and possibly use them as part of subsequent queries. The search and browse operations are driven by the Catalogue, which, in addition to detailed descriptions, contains data related to

digital provenance. Catalogue information is used to address queries to the appropriate archives, which contain the information the user is interested in. A complex network of services will provide users with advanced features for using data in new ways, such as advanced visualisation, definition of use cases and scenarios potentially different from the ones in which the same data were created. Advanced interfaces for querying the item-level semantic network are also provided, so as to obtain relevant information about objects, places, events, people and types according to semantic criteria and to retrieved and display them in a user-friendly and meaningful way.

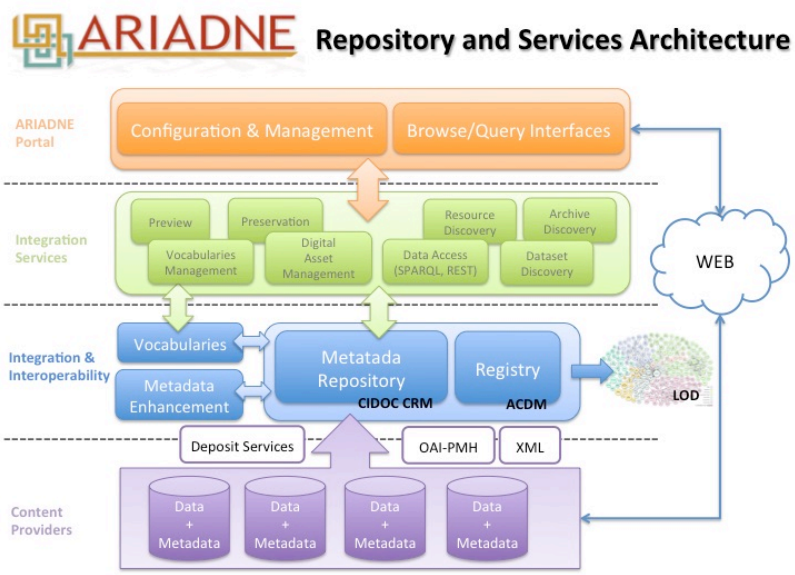


Figure 3: The ARIADNE Architecture

4 Conclusions

The ARIADNE portal was developed to create a unique global access point to provides researchers, repository managers, and the wider interested public, open access to integrated archaeological information. ARIADNE, in its scope, represents a substantial innovation for archaeology, as it provides a common platform where dispersed data resources can be homogenously described, discovered and accessed. It is also a crucial step towards the ambitious goal of offering archaeologists integrated data, tools and computing resources for web-based research and data reuse, which will pave the floor to the creation of new knowledge. The main ambition of ARIADNE is to achieve a wider engagement of the archaeological research community in sharing and reusing data through the ARIADNE portal.

Acknowledgments

The present work has been supported by the ARIADNE project, funded by the European Commission (grant 313193) under the FP7 INFRA-2012-1.1.3 call. The authors opinions do not necessarily reflect those of the European Commission.

References

1. Niccolucci, F., Richards, J.D.: ARIADNE Advanced Research Infranstructure for Archaeological Dataset Networking in Europe. International Journal of Humanities and Art, 7(1-2), pp.70–88, (2013).
2. ARIADNE 2016: www.ariadne-infrastructure.eu
3. ADS: www.archaeologydataservice.ac.uk
4. DANS: www.dans.knaw.nl/en
5. SND: www.snd.gu.se/en
6. CulturaItalia: www.culturaitalia.it
7. Selhofer, H., Geser G.: First Report on Users Needs. ARIADNE Deliverable D2.1 (2014). Available at http://www.ariadne-infrastructure.eu/Resources/D2.1-First-report-on-users-needs.
8. Selhofer, H., Geser G.: Second Report on Users Needs. ARIADNE Deliverable D2.2 (2015). Available at http://www.ariadneinfrastructure.eu/content/view/full/1188
9. Papatheodorou, C., et ali: Initial report on standards and on the project registry. ARIADNE Deliverable 3.1 (2013).
10. Goedertier, S.: DCAT application profile for data portals in Europe (2013). https://joinup.ec.europa.eu/asset/dcatn application profile/
11. AAT: http://www.getty.edu/research/tools/vocabularies/aat/
12. Binding, C., Tudhope, D.: Improving Interoperability using Vocabulary Linked Data. International Journal on Digital Libraries 17, 1 (2016), 5–21.
13. Pleiades: https://pleiades.stoa.org
14. PeriodO: http://perio.do
15. CIDOC CRM. Current Official Version of the CIDOC Conceptual Reference Model (2015). Available at http://www.cidoc-crm.org/docs/cidoc crm version 6.2.pdf
16. CRMinf: the Argumentation Model, version 0.7 (2015). Available at http://www.ics.forth.gr/isl/CRMext/CRMinf/docs/CRMinf-0.7.pdf
17. CRMsci: the Scientific Observation Model, version 1.2.3 (2016). Available at http://www.ics.forth.gr/isl/CRMext/CRMsci/docs/CRMsci1.2.3.pdf
18. Doerr, M., Hiebel, G.: CRMgeo: Linking the CIDOC CRM to GeoSPARQL through a Spatiotemporal Refinement. TECHNICAL REPORT: ICS-FORTH/TR-435, (2013
19. Doerr, M., Theodoridou, M.: CRMdig: A Generic Digital Provenance Model forScientific Observation. In USENIX workshop on the Theory and Practice of Provenance (TaPP). Heraklion, Crete (2011).
20. Ronzino, P.: CIDOC CRMba A CRM Extension for buildings archaeology information modelling. (Unpublished doctoral thesis). The Cyprus Institute, Nicosia, Cyprus (2015)
21. CRMarchaeo: the Excavation Model, version 1.4 (2016) Available at http://www.ics.forth.gr/isl/CRMext/CRMarchaeo/docs/CRMarchaeo_v1.4.pdf
22. Minadakis, N., et ali: X3ML Framework: An Effective Suite for Supporting Data Mappings. Proceedings Workshop EMF-CRM2015, Poznań, Poland, September 17 (2015) CEUR-WS.org, online CEUR-WS.org/Vol-1656/paper1.pdf

Syncretic Text Composition in Artificial Museum Guides

Antonio Sorgente, Antonio Calabrese, Gianluca Coda, Paolo Vanacore, and Francesco Mele

Institute of Applied Sciences and Intelligent Systems "Eduardo Caianiello" of the National Research Council
Via Campi Flegrei 34, 80078 Pozzuoli (Naples) Italy
{a.sorgente, a.calabrese, g.coda, p.vanacore, f.mele}@isasi.cnr.it

Abstract. In this paper, we present our ongoing research about the composition of syncretic text for artificial museum guides. During a museum visit, the visitors receive information about the cultural assets and responses to their questions. The aim is to reuse existing texts(for example those already published on the web) to compose responses for visitors that take into account the time at their disposal, and are balanced with respect to possible insights. Finally, system responses will result from a composition process that coherently synchronises media elements with a synthetic voice related to selected text.

Keywords: syncretic text, multimedia composition, cultural heritage

1 Introduction

Nowadays, the diffusion of new technologies (such as mobile and wearable devices) has allowed the practitioners and organisations operating in the area of Cultural Heritage to propose new approaches for the fruition of cultural assets. This approaches allow us to access to museum collections in multiple ways, both in and off site. Also, the amount of information related to the domain of Cultural Heritage built by experts, and published on the web, is growing day by day. In this scenario, an important aspect is related the extraction of information that must be coherent with the query submitted to the system.

The goal of a generation system is to produce text in response to a given stimulus. It must be able to choose what to include in such text and how to organise this information so that it can be easily understood, and increase the knowledge of the user. The most common information available on cultural heritage is related to the story of the asset or what it depicts. The best way to represent stories consists of using natural language. Our aim in this work is to propose an approach that allows us to dynamically generate information that can be close to the user request, re-using textual information provided by experts and/or already published on the web, integrating these lastly with media resources (photos and video) to generate a unique multimedia response.

In this paper, we present an ongoing research about the production of syncretic text for artificial museum guides. The texts have to take into account the time available to visitors, and to be balanced with respect to the possible insights. The latter means that the response provides an explanation of equal length for each topic involved in the dialogue. The main characteristic of this methodology is to propose an approach based on the thematic structure of the text, selecting appropriate contents related to a cultural item and then aggregating them with media resources (photo and video).

The construction of the thematic structure is based on CSWL formalism[1]. Also, we want to use the thematic progression as pattern for the selection and composition of the text to be proposed to the user. The application of the thematic progression permits us to improve the cohesion and coherence of the composed text provided to visitors.

This activity has been developed within the SIMArt project[1]. The aim of the project is to design interactive multimedia systems for the use of the cultural heritage based on the augmented knowledge paradigm[2].

We will briefly introduce the concept of thematic organisation of text and how we define it using CSWL annotation. We will present the approach adopted to compose text and how to synchronise it with media resources. Finally, some conclusions and future work will be presented.

2 A dialog model based on Theme-Rheme structure

In the construction of text, the speakers/writers construct their messages gradually introducing concepts in such a way that the message is clear, like a touristic guide. A way to achieve this aim is to organise the text through a thematic structure[3]. This structure is based on two elements: *theme* and *rheme*. The *theme* (called also topic) is related to 'what' the text is talking about, and *rheme* (called also comment or focus) is related to 'what' is said about the topic. This structure is known as thematic organisation of text. For the automatic text composition, our model is based on such structure.

As reported in [4], the theme typically contains information which has been previously mentioned or refers to the context of discussion, for example in a museum it can be a cultural asset. It is followed by the rheme that is the part of text that explains the theme introducing new information. An example of theme and rheme in a sentence is the following: *(The Basilica of Saint Clara)*$_{theme}$ *(in Naples was built between 1310 and 1340)*$_{rheme}$. In the sentence, the goal is to talk about the Saint Clara church (theme) and to say something about the story of its construction (rheme).

In order to build the response text in a dialog, some principles, based on the thematic organisation (called thematic progression[5]) have been defined. The thematic progression defines how the theme and rheme are introduced in order to have cohesion and coherence in the text. It can be seen as the skeleton of the plot.

[1] SIMArt (Interactive Multimedia Systems for the use of Art objects) is a project of National Research Council (CNR)

The main types of thematic progression, called patterns, are: linear progression, constant (or parallel) progression, split-theme, and split-rheme progression[5, 6].

In the Constant Progression (CP) the same theme appears in sequence in a series of sentences, in some cases using words having equivalent meaning (see Fig. 1.a). With Linear Progression (LP), the rheme of a sentence will be the theme of the next sentence (see Fig. 1.b). Instead, in Derived-Theme Progression (DTP), the theme of a sentence is linked to the theme of the next sentences. This means there exists a relation (hypertheme) between the themes. An example is the meronomic relation as shown in Fig. 1.c. Finally, the Split-Rheme Progression (SRP), that can be considered a general case of LP, is applied when a sentence introduces more rhemes. In this case, each rheme becomes the theme of a sentence (see Fig. 1.d). Analysing some cultural texts, it has been necessary to introduce another pattern that we called: Inverse Progression (IP). With respect to this pattern, the theme of the sentence becomes a rheme of the new sentence (see Fig. 1.e).

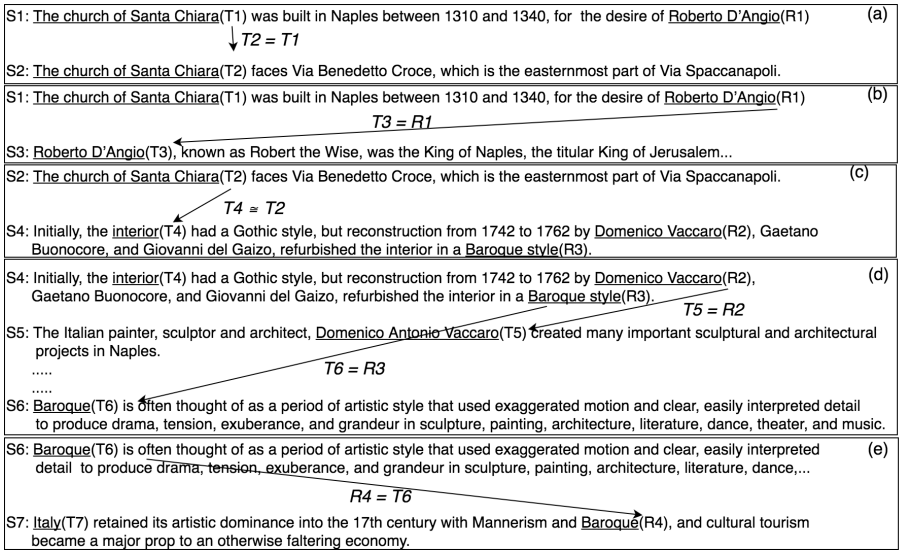


Fig. 1. Examples of Thematic Progression Pattern: (a) Constant; (b) Linear; (c) Derived-Theme; (d) Split-Rheme; (e) Inverse.

2.1 Correspondence between event components and theme-rheme elements

One of the key ideas of this work is the use of sentences annotated by the CSWL formalism to represent the thematic structure. Generally, a theme corresponds to

the first noun phrase of a sentence that is participant, circumstance, or process, while the last part of sentence contains the rhemes[7].

Our starting point is the annotation in CSWL[1]. CSWL is an event-based formalism and defines three types of entities changing over time: *simple events*, *complex events* and *fluents*. **Simple Events** are represented by four components: *When* - the time interval in which the event happens; *What* - the action happening in the event; *Where* - the location where an event takes place; and *Who* - the participants in the event. In CSWL, stories are represented through complex events. A **Complex Event** is constituted by a set of events, causal and temporal relationships between them, and all properties holding true over the time in which the story unfolds. Through the **Fluents** one can represent properties, mental events, spatial relations, and meronomic relations, which are all entities that can change over the time.

Using event-based annotation, for each sentence S we have one or more events annotated, and for each event e_i there are four components:

$$\begin{aligned} Event(S) &= e_1, e_2, ...e_n \\ Components(e_i) &=< what_i, who_i, where_i, when_i > \end{aligned}$$

In our approach, we select as theme the first event component that appears in the sentence and the remaining components as rhemes. If we consider the following sentence with related CSWL annotation:

S_1: La basilica di Santa Chiara$_{po1}$ in Napoli$_{loc1}$ fu costruita$_{act1}$ tra 1310 e 1340$_{int1}$.
- Event(S_1) = e_1
- Compoments(e_1)= $< act_1, po_1, loc_1, int_1 >$
  with $act1 : Action, po1 : PhysicalObject, loc1 : Location, int1 : Interval$

In accordance with this representation, the theme is po_1, that is "la basilica di Santa Chiara" and the rhemes are act_1, loc_1, int_1. Each thematic schema defined starting from the CSWL annotation, can be enriched with other semantic relations such as: meronomic, hyperonomic, synonymous. Then, using this representation, we implement the thematic progression presented in the previous section. Through this approach we choose the next sentences according to both thematic progression and semantic relations.

3 Response as regular expansion

The construction of natural and self-explanatory responses for the visitor needs to be built with respect to some criteria. In a museum, a key constraint of the visit is time. Each visitor has a limited amount of time to spend for museum visits. So, the systems to generate responses have to take in to account such value. Also, during a presentation, the text need to be clear, so it is necessary to present and explain all new terms introduced for the first time. Of course, in such process, the text does not explain obvious things. The choice whether to explain a concept, or not, can be made in accordance to the visit context and/or user

profile (background). For example, if we consider the sentence *The Basilica of Saint Clara in Naples was built between 1310 and 1340*, and the visitor is located in Naples, it's pointless to explain something about Naples, vice versa, it can be useful if the visitor is listening to the story in another country. Finally, the response built for the user has to be balanced, ensuring that the text does not present insights too large with respect to specific subjects and thus, deflecting attention from the main topic.

3.1 Dialogue responses using thematic progression

In this section we present, through some examples, how we can expand a sentence using a thematic propagation that takes into account the semantic annotation of sentences. An important step is to define the procedures for browsing the thematic structure of sentences, to research the text that composes the response.

Listing 1.1. Pseudocode 1

```
1  def spQuery_Expansion(DialogueState,Sx, Tr)
2      Tr = Tr - time(Sx) # response time
3      Expansion = Sx     # expanded response
4      DS = DialogueState # sentences already uttered
5
6      if Tr>0:
7          for each S in (Text -(DS U Expansion)):
8              if  (thematic_progression(S,SX) and
9                   semantic_relation(S, Sx) and time(S) <= Tr+Dt):
10                     Expansion = Expansion + S
11                     Tr = Tr - time(S)
12     if Tr>0:
13         for each S in (Text - (DS U Expansion))
14             if  (linear_expansion(S,Sx) and time(S)<= Tr+Dt):
15                     Expansion = Expansion + S
16                     Tr = Tr - time(S)
17     if Tr>0:
18         for each S in (Text- (DS U Expansion))
19             if  (constant_progression(S,Sx) and Time(S)<= Tr+Dt):
20                     Expansion = Expansion + S
21                     Tr = Tr - time(S)
22     return Expansion
```

Generally, the main theme is related to a cultural asset, but in an interactive system based on dialogue, the starting theme for the search depends on the user query (Q). As a first step, the system identifies the event, and the corresponding sentence (Sx)[8], that answers the user query Q. So in the expansion phase we will have as the starting point the sentence Sx. The listing 1.1 shows the pseudocode of algorithm for expansion of sentence Sx with respect to a specific query Q (e.g. *"When did Vaccaro renovated the Basilica?"* or *"When was it bombed?"*.

As a first step (lines 6..10), the algorithm finds all sentences strongly connected with Sx. We assume that two sentences S_1 and S_2 are strongly connected

if there is a thematic progression and a semantic relation between them. For semantic relations, we have used causal relations and hyperonomy relations. In the first case, this means that there is a cause and effect relation between two sentences, and in the composition process it is more useful to choose both sentences for the explanation. While in the second case, there exists a specialisation of some sentence component, which is an insight.

After this selection, the algorithm, for the next text, searches (lines 11..16) the sentences S which have a Linear Progression ($Theme(S) \in Rheme(Sx)$), which are sentences that have as theme one of the rhemes belonging to Sx.

Listing 1.2. Peudocode 2

```
1  def genQuery_Expansion(DialogueState,Sx, Tr)
2      Tr = Tr - time(Sx) # response time
3      Expansion = Sx     # expanded response
4      DS = DialogueState # contains the sentences already uttered
5
6      if Tr>0:
7          for each S in (Text- (DS U Expansion)):
8              if (constant_progression(S,Sx) and time(S)<= Tr+Dt):
9                      Expansion = Expansion + S
10                     Tr = Tr - time(S)
11     if Tr>0:
12         for each S in (Text- (DS U Expansion)):
13             if (inverse_progression(S,Sx) and time(S)<= Tr+Dt):
14                     Expansion = Expansion + S
15                     Tr = Tr - time(S)
16     if Tr>0:
17         for each S in (Text- (DS U Expansion)):
18             if (derived_progression(S,Sx) and time(S)<= Tr+DX):
19                     Expansion = Expansion + S
20                     Tr = Tr - time(S)
21     if Tr>0:
22         for each S in (Text-(DS U Expansion)) and St in Expansion:
23             if (linear_progression(S,St) and time(S)<=Tr+DX):
24                     Expansion = Expansion + S
25                     Tr = Tr - time(S)
26     return Expansion
```

If the user's query contains a generic request (e.g. *"can you tell me something about the basilica?"* or *"can you give me some information about the style"*, this means that he/she does not ask specific information about the topic expressed in the query. The listing 1.2 shows the pseudo-code of the algorithm for the expansion of sentences with respect to a generic query Q. As in the previous case, the starting point is the first sentence Sx that the system provides as a response. In this case, the request is generic, and it asks information about a topic. So, the algorithm finds all sentences that have a Constant progression with respect to Sx (lines 6..10) and if there is more narration time, it searches (lines 11..15) sentences that have inverse progression ($Theme(Sx) \in Rhemes(S)$) with

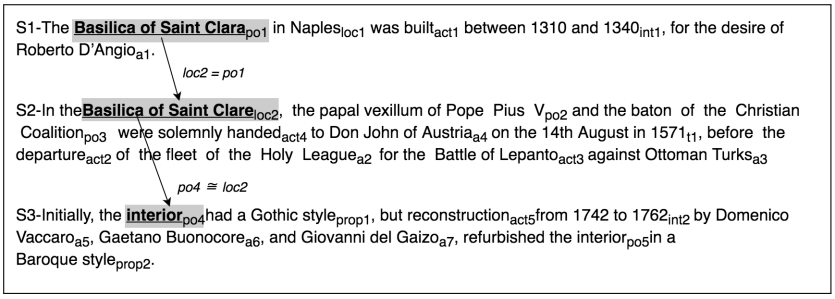


Fig. 2. Answer of the query *"Can you tell me about the Basilica?"*

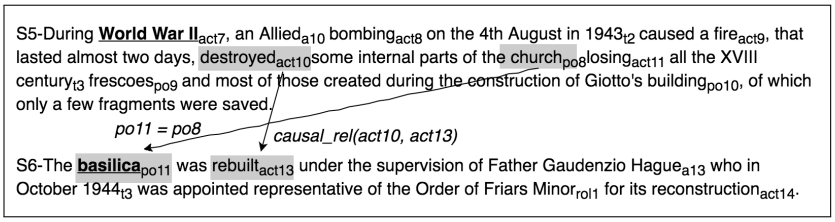


Fig. 3. Answer of the query *"When was it bombed?"*

respect to the theme of the query, or try to find sentences with derived theme progression (lines 16..20). If there is more available time for the narration, the algorithm finds some deepening of the sentences that were already selected in the previous steps (lines 21..24). Using this approach, the composed text will present a generic description of the topic required and, in accordance with the available time for the response, some deeper insights.

As experimentation of the adopted approach, we have considered a cultural text about the Basilica of Santa Chiara in Naples. Asking the query *"can you tell me about the basilica?"* we can obtain the answer shown in Fig. 2. In this figure, the first sentence is selected as the response by the system. To build an expansion, the algorithm in listing 1.2 is applied. It adds two sentences to the response: the first with constant progression and the second with derived theme progression. In the latter, the system can detect the theme through a meronomic relation. In fact, the interior (po_4) is *part of* of the basilica (loc_2).

If we consider the query *"When was it bombed?"* we can obtain an answer as shown in Fig. 3. In this case, starting from the sentence that contains the response, for the expansion the algorithm presented in listing 1.1 is used. We can observe that between the two sentences exist a linear progression and a causal relation, so they are strongly connected.

3.2 Syncretic approach for multimedia responses

To build the composed text for responding to the user question, we create a multimedia response temporally synchronising text and media according to semantic

annotations. This approach is called **syncretic text**[9], namely a text composed of heterogeneous languages within a unitary communications model[2], having features of cohesion and coherence, respect to a same enunciation instance. For these goals, the system selects, and carries out a ranking, using available multimedia objects, that can be associated to the composed text. The selection is based on annotated entities using the semantic of CSWL[8] formalism. Then, multimedia objects selected are synchronised with synthesised text, so that media items are coherently visualised with the time intervals in which a synthetic voice talks about the content represented in the media.

4 Conclusions and Future Work

In this work, we presented an ongoing research activity about the composition of balanced texts that uses a thematic progression structure built through an event based formalism. What we have presented here represents just a first application that composes texts using information coming from a single document, but we believe that this approach can be adopted for building texts integrating more documents. Future work will consist of analysing more texts to validate the patterns of thematic progression and discover new ones. We also believe that presented algorithms can be improved taking into account some characteristics of the user profile. In addition, because the approach is based on CSWL annotation, to reduce the time in such phase, we are working on an assisted tool that helps the users in the annotation.

References

1. Mele, F., Sorgente, A.: CSWL - Cultural Stories Web Language. Technical Report 180/15, Institute of Cybernetics "E. Caianiello" of CNR (06 March 2015)
2. Sorgente, A., Calabrese, A., Coda, G., Vanacore, P., Mele, F.: Building multimedial dialogues annotating heterogeneous resources. In Bordoni, L., Mele, F., Sorgente, A., eds.: Artificial Intelligence for Cultural Heritage. Cambridge Scholars Publishing (2016) 49–82
3. Ebrahimi, S.F., Ebrahimi, S.: Information development in efl students composition writing. Advances in Asian Social Science **1**(2) (2012) 212–217
4. Halliday, M.A.K., Matthiessen, C.M.: Halliday's introduction to functional grammar. Routledge (2013)
5. Danes, F.: Functional sentence perspective and the organization of the text. In Papers on Functional Sentence Perspective (1974) 106–128
6. Soepriatmadji, L.: Thematic progression in thesis abstacts written by english students of fbib unisbank semarang. Jurnal Ilmiah Dinamika Bahasa dan Budaya **3**(1) (2009) 28–40
7. Jing, W.: Theme and thematic progression in english writing teaching. Journal of Education and Practice **6**(21) (2015) 178–187
8. Sorgente, A., Vanacore, P., Origlia, A., Leone, E., Cutugno, F., Mele, F.: Multimedia responses in natural language dialogues. In: Proceedings of AVI*CH 2016. Volume 1621 of CEUR Workshop Proceedings., CEUR-WS.org (2016) 15–18
9. Cosenza, G.: Semiotica dei nuovi media. Laterza (2010)