Generating Personalized Tourist Map Descriptions

B. De Carolis, G. Cozzolongo, S. Pizzutilo

Dipartimento di Informatica -Università di Bari
`http://www.di.uniba.it/intint`

Abstract. When visiting cities as tourists, most users intend to explore the area looking for interesting things to see or for information about places, events, and so on. An adaptive information system, in order to help the user choice, should provide contextual information presentation, information clustering and comparison presentation of objects of potential interest in the area where the user is located. To this aim, we developed a system able to generate personalized presentation of objects of interest, starting from an annotated city-map.

1 Introduction

User-tailored information presentation has been one of the main goals of the research on adaptive systems: features such as the user interests, background knowledge and preferences were considered to settle, at the same time, the information to be included in the message and its 'surface' realisation [1,2,3]. With the evolution of devices (PDA, mobile phones, car-computers, etc.), network connections (GSM, GPRS, UMTS, WLAN, Bluetooth, …) and localization technologies (GPS) for interacting with information services, users can access to these services potentially everywhere and anytime[4]. In this case, the main goal of an adaptive information system is deliver targeted information to the users *when* they need them, *where* they need them and in a form that is suited to their *situational interests* and to the technological context (*how* they need the information).

In general, achieving this objective requires the following system's capabilities:

- accessing the description of the domain data in order to select objects of interest and use their representation for generating related information presentation;
- accessing the description of the current context in order to understand the situation in which the user is (location, activity, device, etc.);
- modelling the situational interests of the user in order to use these data to personalize the selection and presentation of information [5];
- generating information presentation accordingly [6,7].

In this paper, we present a solution to the personalization of information presentation that combines the use of XML annotation for domain knowledge representation, ***Mobile User Profiles*** (MUP) for managing contextualized user preferences and interests, a media-independent content planner and a context-sensitive surface generator.

In order to show how the system works, we will use the tourist domain as an example. Indeed, as mobile phones and other portable devices are becoming more advanced, tourism is one obvious application area. Tourism has been a popular area for mobile information systems. In particular, the Lancaster GUIDE system [8], and other systems based on mobile devices [9,10] are examples of application in this field.

When people visit cities as tourists most users intend to explore the area and find interesting things to see or information about places, objects, events, and so on. According to [11] most of the times they do not make very detailed and specific plans "so that they can take advantage of changing circumstances" and, moreover, when choosing where to go and what to see they tend to "pick up an area with more than one potential facility". According to these findings, it would be useful to support the user choice with contextual information presentation, information clustering and comparison presentation of object of potential interest in the same area.

The paper is structured as follows: after a brief illustration of the system architecture, we focus on the description of the process of generating personalized description of places of interest using an annotated town-map. In particular, we describe the structure of the map annotation scheme, the role of the MUP and the generation steps necessary to produce a personalized map description. Finally, conclusions and future work are discussed in the last session.

2 System Architecture

Let's consider the following situation: "a user is traveling for business purposes, she is in the center of a town and requires information about a place using a personal mobile device. She wants to know what is going on in that area."

In this case, the user is "immersed" in the environment and she is presumed to look for "context-sensitive" information. One of the most common ways for tourists for requesting information about places of interests in a particular town is to use a map. Then integrating information provision with a graphical map of the place is one of the most used metaphors supporting this type of interaction. However, if this map is only a graphical representation of the town, it cannot be "explained" to the user by an automatic system. In order to generate targeted information about places of interest, the map has to be annotated so as to define a correspondence between graphical objects and metadata understandable by the system that has to generate the presentation of information. With this aim, we developed a system that, starting from an XML representation of

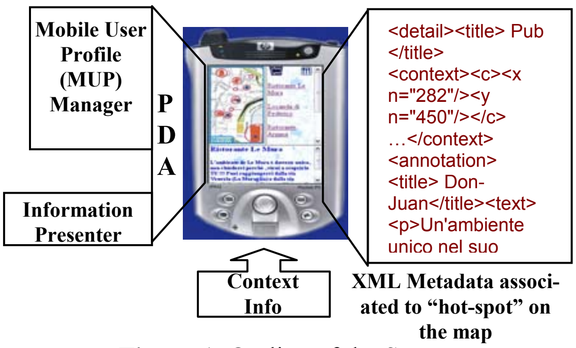

Figure 1: Outline of the System

domain knowledge, decides which information to provide and how to present it either after an explicit user request or proactively in presence of interesting objects, events and so on.

As outlined in Figure1, the system runs on a PDA and uses two other components: the Mobile User Profile (MUP) Manager and the Information Presenter. These components, given a metadata representation of a map, cooperates in establishing which information to present and the structure of the presentation according to the "user in context" features.

In this paper we will not discuss about information filtering, context detection and proactivity issues, but we will focus on the process of generating adaptive information presentation while interacting with the city-map. Let's see in more details which are the methods employed to implement the system.

2.1. Understanding the Map

Understanding a map means extracting and describing objects of particular interest with their descriptive features. Data annotation is a typical solution to achieve this objective. Since we do not use automatic image features extraction techniques, the description of the map components, their attributes and the relationships among them, is achieved using metadata.

In this case, the map image is annotated in a modality-independent way using a markup language and encapsulates tourist information in a XML structure. To build these metadata, we use a tool in Java (Inote [12]) that is available on line and provides a way of annotating images in a user-friendly way. Inote allows to attach textual annotations to a image and to store them in a XML file. Then, Inote's mark-up language is very general and may be applied to every kind of image. For instance, we have been using it for describing radiological images in another project [13].
With Inote it is possible to identify:

- a region of interest, a part of the image, called "<overlay>";
- each overlay may contain some objects of interest denoted as "<detail>" and
- each <detail> may have attributes;
- each attribute is denoted as "<annotation>", and may be given a name;
- a <text> may be associated with every annotation of every detail, in order to add the description of that attribute.

To tailor it to map description, we defined a parser able to interpret the tags according to the following ad hoc semantics (illustrated in Figure 2):

A map region has some "General Properties" that identify it: the name of the town, the described area, its coordinates, and so on. In this wide region it is possible to identify some areas of interest, these are denoted as overlays. The main information content of each overlay then consists in a list of details that correspond to the category of places of interest (eating places, art, nature, and so on); each place of interest is described by a set of attributes (type, position, etc.) denoted as "annotation" whose value is described by the "text" tag.

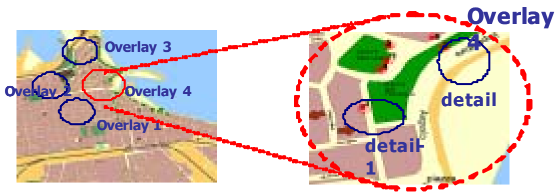

Figure 2: Illustration of the Map Annotation Scheme

The following is an example of structure generated by Inote following this scheme:

```
<overlay><title>bari-zone1</title>
        <detail><title>eating<title>
                <annotation><title>type</title>
                <text>fast-food</text> </annotation>
                <annotation><title>name</title>
                <text>Bar Città Vecchia (da Cenzino)
                 </text> </annotation>
                <annotation><title>coordinates</title>
                <text>41°06'14.800"N 16°45'57.013"E </text>
                </annotation>
                <annotation><title>view</title> <text>historical center</text>
                </annotation>
                <annotation><title>wheelchair accessibility</title>
                <text>yes </text>
                </annotation>
   …</detail></overlay>
```

2.2 Mobile User Profiles

The illustrated interaction scenario depicts a situation in which the user is interacting with the information system with a mobile device. Mobile personalization can be defined as the process of modeling contextual user-information which is then used to deliver appropriate content and services tailored to the user's needs. As far as user modelling is concerned, a mobile approach, in which the user "brings" always with her/himself the user model on an personal device, seems to be very promising in this interaction scenario [14]. It presents several advantages: the information about the user are always available, updated, and can be accessed in a wireless and quite transparent way, avoiding problems related to consistency of the model, since there is always one single profile per user.

Based on this idea, our user modeling component uses profiles that allows to:

- express context-dependent interests and preferences (i.e. "I like eating Chinese food when I'm abroad");
- allows to share its content with environments that can use it for personalization purposes following the semantic web vision [15].

Then, as far as **representation** is concerned, beside considering static long term user features (age, sex, job, general interests, and so on), it is necessary to handle information about more dynamic "user in context" features. Instead of defining a new a ontology and language for describing mobile user profiles, since this is not the main aim of our research, we decided to adopt UbisWorld [5] language as user model ontology of our user modeling component. In this way we have a unified language able to integrate user features and data with situational statements and privacy settings that better suited our need of supporting situated interaction. This language allows representing all concepts related to the user by mean of the UserOL ontology, to annotate these concepts with situational statements that may be transferred to an environment only if the owner user allows this according to privacy settings. An example of a situational statement is the following:

```
<Statement id="14">
<content><subject><UbisWorld:Nadja /></subject>
<predicate><UserOL:eating /></predicate>
<predicate-range><UserOL:restaurant,fast-food,pizzeria/>
</predicate-range><object>fast-food <object>
</content>
<restriction><location>tourist info<location></restriction>
<meta>
<owner><UbisWorld:Nadja /></owner>
<privacy><UbisWorld:friends /></privacy>
<purpose><UbisWorld:information /></purpose>
<retention><UbisWorld:short /></retention>
<explanation confidence="high" creator="Nadja" evidence="
Interface input " method="acquire_pref" />
</meta>
</Statement>
```

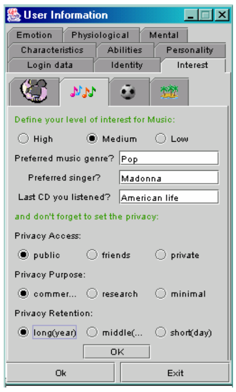

Figure 3. MUP interface.

User preferences, interests, etc. are collected in two ways:

- using a graphical interface (Figure 3) in which the user can explicitly insert her preferences and related privacy settings regarding particular domains,
- deriving other information (i.e. temporary interests) from user actions or from other knowledge bases (i.e. user schedules, agenda, etc. [16]).

User feedback and actions in the digital and real world may reproduce changes in the user model. The MUP manager observes the user actions: when new information about the user can be inferred, it updates or adds a new slot in the MUP and sets the "confidence" attribute of that slot with an appropriate value that is calculated by the weighted average of all the user actions having an impact on that slot. The confidence attribute may be set to low, medium and high.

2.3 Generating Context-Sensitive Information

The Architecture of the Information Presenter is based on the model of Natural Language Generation (NLG) systems [17]. Given a set of goals to be achieved in the selected domain (tourist information in this case), the Agent plans what to communicate to the user and decide how to render it according to the context. In this case, situational user preferences play an important role in order to adapt the description of object to the situation. As it has been already proven in previous research on language generation (e.g.,[7,18]), user-related information could be used to constrain generator's decisions and to improve the effectiveness and tailoring of the generated text. Such an information is useful at any stage of the generation process: i) for selecting relevant knowledge; ii) for organizing information presentation (the organisation strategies or plans can have preconditions dependent on user information); and iii) for the surface realisation (use of words which depends on the context).

3 Selecting Relevant Knowledge

Let's consider the following example: suppose the user is travelling for business reasons and, during lunch break, she is visiting the centre of the town. While she is there, information about places of interest close to where she is will be emphasized on the interactive map running on her personal device.

In this case, the Information Presenter will ask to MUP manager to select the situational statements regarding "time_of_day = lunch time" when "reason_of_travel=business purposes" and when the user "location=town-centre". In the set of selected statements, the one with the highest confidence value will be chosen.

Referring to the previously mentioned example, in the described context, the MUP Manager will infer that the user prefers to eat something fast but in a place with a nice view on the town center. Then, according to this preference, the Information Presenter will select, in the XML description of the map, all places (<details>) of category "eating" being "fast-foods" with coordinates that show that the place is relatively close to the user position (within 500 mt). Moreover, the system will check for other features matching the presumed user preferences (i.e. view="historical center"). Then a new xml structure containing the selected places will be generated to be used for the presentation. Selected items are then ordered on the bases of number of matched user features. As the user moves, the map is updated as well as the context information.

3.1 Organizing the Information Presentation

There are several computational approaches to planning "what to say" when presenting information. Important milestones in this research field were the introduction of *text schemata* [19] and *Rhetorical Structure Theory* (RST), as formulated by Mann and Thompson [20]. Meanwhile, RST has been operationalized by the application of a traditional top-down planner [21], and has been further refined by the introduction of intentional operators [22]. Planning, however, is an heavy computational task. Considering the need of dealing with real-time interaction on a *small* device, our approach is based on the idea of using a library of non-instantiated plan-recipes expressed in an XML-based markup language: DPML (Discourse Plan Markup Language [23]). DPML is a markup language for specifying the structure of a discourse plan based on RST: a discourse plan is identified by its name; its main components are the nodes, each identified by a name. Attributes of nodes describe the communicative goal and the rhetorical elements: role of the node in the RR associated with its father (nucleus or satellite) and RR name.

The XML-based annotation of the discourse plan is motivated by two reasons: i) in this way, a library of standard explanation plan may be built, that can be instantiated when needed and can be used by different applications, in several contexts; ii) XML can be easily transformed through XSLT in another language, for instance HTML, text or another scripting language driving for instance a TTS, favoring in this way the adaptation to different context and devices.

Once a communicative goal has been selected, explicitly as a consequence of a user request or implicitly triggered by the context, the Information Presenter selects the plan in this library that best suits the current situation. The generic plan is, then, instantiated by filling the slots of its leaves with data in the XML-domain-file that is associated with the map to describe. In this prototype we consider the following types of communicative goals:

- Describe(Ag, U, x) where x is a single object to be described;
- Describe(Ag, U, list_of(y_i)) where list_of(y_i) represent a set of objects of interest of the same type (i.e. restaurants) to be described;
- DescribeArea(Ag, U, list_of(z_i)) where list_of(z_i) represent a list of objects of interest belonging to different categories.

Considering the previous example, the Presentation Agent will select the plan correspondent to the Describe(Ag, U, list_of(y_i)) goal for listing the eating facilities matching the user preferences and then it will instantiate it with the selected data (fast foods close to where the user is, with a nice view and open at the current time). A small portion of the XML-Instantiated-Plan that was generated for describing some eating facilities in the area is shown in Figure 4.

```
<d-plan name="describe_set_of_objects">
<node name="n1" goal="Describe(where_to_eat, area1)" role="root" RR="Elab">
    <node name="n2" goal="Inform(existence(fast_foods))" role="nucleus" RR="null"/>
    <node name="n3" goal="Describe(fast_foods, area1)" role="sat" RR="ElabGenSpec">
            <node name="n4" goal="Inform(number(fast_foods, 3))" role="nucleus" RR="null"/>
            <node name="n5" goal="Describe(list(fast_foods))" role="sat" RR="OrdinalSequence">
            <node name="n5.1" goal=" Describe(fast_foods, “La Locanda di Federico”)" role="nucleus"
            RR="ElabObjAttr">
                  <node name="n5.1.1" goal="Inform(name, “fast_foods”)" role="nucleus" RR="null"/>
            <node name="n5.1.2" goal="Describe(Specific Features, image)" role="nucleus"
            RR="OrdinalSequence">
                  <node name="n5.1.2.1" goal="Inform(type, “osteria tipica barese”)" role="nucleus"
            RR="null"/>
                  <node name="n5.1.2.2" goal="Inform(rel_pos, “100 meter North”)" role="nucleus"
            RR="null"/>
                  <node name="n5.1.2.3" goal="Inform(timetable, “12.00-24.00”)" role="nucleus" RR="null"/>
                  <node name="n5.1.2.4" goal="Inform(telephone, “0805240202”)" role="nucleus"
            RR="null"/>
                   <node name="n5.1.2.5" goal="Inform(description, “a osteria where it is possible to eat
            good typical bari food….”)"  role="nucleus" RR="null"/>
            </node>
    </node>…
    </node>
…</node></d-plan>
```

Figure 4. An example of XML-Instantiated-Plan.

This plan first presents general information about the existences of open fast foods, then it lists them, describing in details their main features.

3.2 Rendering the Map Objects Description

Adaptation of layout (visible/audible) should support alternative forms of how to present the content, navigational links, or the presentation as a whole.

The appropriate transformation technology, especially when considering standard initiatives, is obviously XSL transformation (XSLT) in combination with DOM (Document Object Model) programming. XSLT is an effective way to produce output in form of HTML, or any other target language. Rule-based stylesheets form the essence of the XSLT language and build an optimal basis for the introduced adaptation mechanism.

The surface generation task of our system is then very simple: starting from the instantiated plan apply the appropriate template. This process is mainly driven by the type of the communicative goal and by the RRs between portions of the plan. The plan is explored in a depth-first way; for each node, a linguistic marker is placed between the text spans that derive from its children, according to the RR that links them.

For instance, the description: "There are 3 fast foods in this town area", in Figure 5, is obtained from a template for the Describe(Ag, U, list_of(yi)) where the Ordinal Sequence RR relates the description of the single objects in the list. We defined the templates' structure after an analysis of a corpus of town-map websites. At present, we generate the descriptions in HTML; however, our approach is general enough to produce descriptions in different formats and, therefore, for different interaction modalities [24].

Figure 5. List of eating places

In the example in Figure 5, the Information Presenter will display to the user a web page structured as follows: i) on the left side the portion of the map of the town area where the user is located and the graphical indications (icons denoting different categories of objects) about places of interests is displayed; ii)on the right side a description of those objects is provided; iii) on the bottom part, when the user selects one of the objects in the list, a detailed description of the selected object will be displayed. The user may access the same information directly clicking on the icons on the map.

Looking in more detail at the proposed information could be considered as a positive feedback in building the usage models. However, while this is important in the case of non-mobile information systems, when the user is moving in a real space, this is not enough. In this case, the digital action should be reinforced by the action in the real world: going to that place. We are still working on this issue since it is important to consider contextual events that may discourage the user to eat in that place (i.e. the restaurant is full). At the moment, for dealing with this kind of feedback, we ask directly to the user.

4 Conclusions and Future Work

In this paper, we described the prototype of a system able to generate context-sensitive description of objects of interest present in a map. Even if we selected the mobile tourism as a application domain to test our approach, the system architecture and employed methods are general enough to be applied to other domains. Moreover, the use of XML content modeling and domain-independent generation methods, allows the system to deal with the adaptation of the presentation modality. In this way, the provided information can be easily adapted to different devices and to the needs of user with disabilities. The system has been implemented in Java and XML related technologies. We tested on a iPAQ h5550 without GPS. We simulated the user location with an interface for managing context features.

In this phase of our work we are concerned more with the study of the feasibility of the proposed approach and employed methods than in evaluating the effectiveness of the generated description. At this stage we performed only an evaluation of the generated text against the descriptions present on Bari tourist guide and the results show a good level of similarity. However, this does not show any evidence that contextual information provision is more effective than non-contextual one. This will be the aim of our future user studies. After this study, in case there is an evidence that contextual information provision is effective, we will concentrate on the generation of comparative descriptions of places of interests in the same area.

Acknoledgements

Research described in this paper is an extension of the work we performed in the scope of the ARIANNA project. We wish to thank those who cooperated in implementing the prototype described in this paper: in particular, Gloria De Salve and Marco Vergara. In particular, we thanks Fiorella de Rosis for her useful comments on this work.

References

1. L. Ardissono, A. Goy, G. Petrone, M. Segnan and P. Torasso. Ubiquitous user assistance in a tourist information server. Lecture Notes in Computer Science n. 2347, 2nd Int. Conference on Adaptive Hypermedia and Adaptive Web Based Systems (AH2002), Malaga, pp. 14-23, Springer Verlag 2002.
2. Brusilovsky P.(1996).Methods and Techniques of Adaptive Hypermedia.UMUAI,6.87:129.

3. Wilkinson R., Lu S., Paradis F., Paris C., Wan S., and Wu M. Generating Personal Travel Guides from Discourse Plans. P. Brusilovsky, O. Stock, C. Strapparava (Eds.): Adaptive Hypermedia and Adaptive Web-Based Systems International Conference, AH 2000, Trento, Italy, August 2000. Proceedings LNCS 1892, p. 392 ff.
4. Weiser M. The Computer for the 21st Century. Scientific American, september 1991.
5. UbisWorld: http://www.u2m.org
6. De Carolis, B., de Rosis, F., Pizzutilo, S.: Generating User-Adapted Hypermedia from Discourse Plans. Fifth Congress of the Italian Association of Artificial Intelligence (AI*IA 97), Roma , (1997).
7. Paris, C. User modelling in Text Generation. Pinter Publishers, London and New York. 1993.
8. Cheverst K., Davies N., Mitchell K., Friday A. and Efstratiou, Developing Context-Aware Electronic Tourist Guide: Some Issues and Experiences, Proceedings of CHI'2000, Netherlands, (April 2000), pp. 17-24.
9. Gregory D. Abowd, Christopher G. Atkeson, Jason I. Hong, Sue Long, Rob Kooper, Mike Pinkerton: Cyberguide: A mobile context-aware tour guide. Wireless Networks 3(5): 421-433 (1997).
10. Pan, Bing and Daniel R. Fesenmaier (2000). "A Typology of Tourism Related Web Sites: Its Theoretical Background and Implications." In Fesenmaier, Daniel R., Stefan Klein and Dimitrios Buhalis (Eds.),Information and Communication Technologies in Tourism 2000 (pp. 381-396). Springer-Verlag.
11. B. Brown, M. Chalmers (2003) Tourism and mobile technology, In: Kari Kuutti, Eija Helena Karsten (eds.) Proeceedings of the Eighth European Conference on Computer Supported Cooperative Work, Helsinki, Finland, 14-18 September 2003., KluwerAcademic Press.
12. Inote: Image Annotation Tool. http://jefferson.village.edu/iath/inote.html.
13. De Salve, G., De Carolis, B. de Rosis, F. Andreoli, C., De Cicco, M.L. Image Descriptions from annotated knowledge sources. IMPACTS in NLG, Dagsthul, July 25-28, 2000.
14. Kobsa A., Generic User Modeling Systems. UMUAI vol. II nos.1-2 pp.49-63. Kluwer Academic Publisher. 2001.
15. A. Sinner, T. Kleemann, A. von Hessling: Semantic User Profiles and their Applications in a Mobile Environment. In Artificial Intelligence in Mobile Systems 2004.
16. Cavalluzzi A., De Carolis B., Pizzutilo S., Cozzolongo G.: Interacting with embodied agents in public environments. AVI 2004: 240-243.
17. Cozzolongo, G., De Carolis, B., Pizzutilo, S.. Supporting Personalized Interaction in Public Spaces. In Proceedings of the Artificial Intelligence in Mobile Systems 2004. Baus J., Kray, C., Porzel, R. (Eds.). Nottingham, UK, 2004.
18. Reiter E. and Dale R. Building Natural Language Generation Systems. Cambridge University Press. 2000.
19. McKeown, K. Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text. Cambridge University Press, Cambridge, England. 1985.
20. Mann W.C., Matthiessen C.M.I.M., Thompson S. (1989). Rhetorical Structure Theory and Text Analysis.ISI Research Report- 89- 242.
21. Hovy, E., (1988), Generating Natural Language under Pragmatic Constraints, Hillsdale, NJ: Lawrence Erlbaum Associates.
22. Moore J. and Paris C.. ``Planning Text for Advisory Dialogues: Capturing Intentional and Rhetorical Information." Computational Linguistics Vol 19, No 4, pp651-694, 1993.
23. De Carolis B., Pelachaud C., Poggi I. and Steedman M. APML, a Mark-up Language for Believable Behavior Generation. In H Prendinger and M Ishizuka (Eds): "Life-like Characters. Tools, Affective Functions and Applications". Springer, in press.
24. Palle Klante, Jens Krösche, Susanne Boll: AccesSights - A Multimodal Location-Aware

Mobile Tourist Information System. ICCHP 2004: 287-294.