Ubiquitous Interaction: Adapting to the "User in Context"

Berardina De Carolis and Sebastiano Pizzutilo
Department of Informatics,
University of Bari,
via Orabona 4 , 70126 Bari, Italy

Abstract: - *Context-awareness* is increasing of importance in achieving an effective communication of information and provision of services. Then, besides taking into account "classical" user-related features, other context-dependent factors (i.e. users' location, activity, emotional state and technical characteristics of the used device) have to be considered. In this paper, we present the Personalization component of a multiagent infrastructure that we have developed for supporting interaction between users and environments. This component establishes which information to present, how to organize it and how to set up the output layout according to "user in context" features.

Key-Words: - Human-Computer Interaction, User context adaptivity, Ubiquitous computing, Multiagent systems

1 Introduction

New technologies (wireless handheld devices, active everyday objects, and so on) are available for interacting with services of various type, like access to information, e-commerce, scheduling, smart guidance, and so on. The nature of these interaction technologies makes these services potentially accessible from everywhere, simultaneously with other user's activities to which they are often related. Personalizing interaction then becomes necessary for achieving an effective communication between service providers and their users.

Several factors have to be taken into account, which are related to the nature of the provided service, to the device/s employed in interaction to user-related "long-term" factors and to context-related "short-term" factors. This means that, besides adapting interaction to characteristics of users, other features related to the context have to be considered [1, 16]. For instance, the users' location, the activity in which they are involved, their emotional state and, finally, the technical characteristics of the device/s they are using.

Information about the location may be employed to contextualize generation of natural language texts, presentation of graphical maps and other. The users activity, their emotional state and the input/output capabilities of the device they are employing may influence the way the information is accessed and presented: for instance, browsing in a large information space, searching for some specific data or receiving fast and well focused 'hints'. In addition, users might be interested to add a social dimension to their activity, by being enabled to introduce and share comments on the information they receive.

We see the environment as a multiagent system populated by i) service-agents, that can provide different kinds of services, ii) D-Me agents, that represent users in the environment and negotiate services by acting on their behalf, and iii) users themselves.

We developed a FIPA (Foundation for Intelligent Physical Agents [9]) compliant infrastructure for supporting interaction between users and environments providing services of various types.

In this paper, we will focus on the personalization component, that establishes how to negotiate services and present the results according to the "user in context" characteristics and needs.

The paper is organized as follows: Section 2 outlines some architectural requirements and describes the organization of our multiagent infrastructure. Section 3 illustrates how this architecture may be used for supporting "personal" interaction in ubiquitous computing. Section 4 reports an example of its application and, finally, the Section 5 contains conclusions and future work.

2 D-Me Infrastructure

Interaction with environments of different types, that provide different kinds of services, presents problems related to the variability of the environment and of its users and to the fact that devices employed at present are not simple enough to provide a natural interface to every user. The idea of delegating to an agent, representing a particular user in a particular environment, the request of services to agents living in another environment, may be a solution to these problems. These agents may insure the provision of personalized presentation of results, that is appropriate to the "user in context". To this aim, different types of information need to be understood and communicated among the user and different agents in the environment: a) information about the *tasks* that the user needs to perform in that environment; b) information about the *user*, represented by its agent; c) information about the *environment*; d) information about the *results* that have to be presented to the user in a personalized way.

D-Me architecture tries to take these factors into account, through the following functions:

- Access, by external entities, to a not-fixed number of services of various kinds. This enables users of different types and with different needs to interact with the services exported by the environment, according to the proprietary protocols of the architecture;
- User modeling and execution of different services according to what has been inferred by this component;
- Personalization of the output, that is of the results of services, according to user and context related features.

Our main objective has been the development of an open infrastructure allowing the definition of a pattern for instantiating multiagent environments in which its entities are distributed on at least two platforms: one for the mobile agent representing the user and the other for the environment managing all the available services. To define the roles of every agent in this infrastructure, a typical case has been analyzed: "The entity representing the user perceives the presence of a smart environment, analyses the available services and requests some of them accordingly with its goals. Every invoked service adapts its behavior to the user, according to what has been inferred by the user modeling component, and gives back personalized results".

The entities explicitly involved in this case are: i) the user, ii) the agent that represent her/him in the environment (D-ME Agent), iii) a service provider and iv) the environment itself, as a manager of the available services. This can be seen as an abstraction for representing an open platform for supporting the development of multiagent environments.

The communication problem has been approached using FIPA Agent Communication Language (ACL) compliant messages. This approach, that combines standard protocols for communication with ontological bases for describing concepts, has the advantage of making communication independent of the inner representation languages or standards. The only constraint is that every entity that supports this communication must share with the other parties a portion of an ontology, whose concepts are the subject of the conversation.

3 Personalising Interaction

After this brief description of the infrastructure, we will show an application, by focusing on its *Personalization component*: this is realized through two specialized service agents: the *UMAgent* and the *Presentation Agent*.

As we mentioned in the Introduction, personalized interaction depends on several features, related to both the user and the context. By analyzing these features, we may distinguish between two families of factors that influence adaptivity of the presentation:

[1] *long-term* factors, that persist during time or evolve slowly. These are related to the user's background knowledge, cognitive capacity, experience, personality, job, sex, age, interests about topic families and so on. These factors are usually stored in a User Profile, that can be used to build an individual User Model;

[2] *short-term* factors, that are related to a particular interaction session or situation, depending on the context in which the user is and moves [10]; *context-awareness* regards (i) the environment in which the application operates; it is identified by the physical and social surroundings of the user, her emotional state, the location, the activities going on in the environment and other external properties; (ii) the activity or task the user is performing; (iii) the device employed: a PC, an handheld device, a mobile phone, a TV, a watch, and so on, or a combination of these devices.

These factors influence different aspects of the interaction and, in particular of information presentation.

Adaptation may generally be introduced in the following phases of the generation process:

a. when *planning the presentation content*, that is when deciding "what to say" to the user, at which level of detail and by using which media ;
b. when *rendering the presentation* at the surface level, that is when deciding "how to display information" to the user, by selecting an appropriate presentation.

Considering context in addition to long-term factors requires adding to the usual adaptation criteria other issues. For instance:

- **activity**: the task the user is performing influences the type of information and the way information is accessed. Providing a search or a browsing access modality or pushing relevant information for emergency are examples of adaptation to activity;
- **emotional state**: factors concerning the emotional state influence the presentation agent's response, the emotion to be expressed, its intensity and so on. This does not mean that the expression relies on an embodiment of the Agent: affective

expression can be conveyed through speech or text style, figures, colours etc. For instance: if a user requests information in conditions of emergency, the Presentation agent will have to avoid engendering panic, by using "ensuring" expressions or voice timbre.
- **device**: the display capacity affects the way that information is selected, structured and rendered. For instance, natural language texts may be more or less verbose, complex figures may be avoided or substituted with ad hoc parts or written/spoken comments. The possibility of using more than one display at the same time influences, in particular, the distribution of information items among the devices according to their capabilities. For instance, if a beamer is available in the environment and the user is interacting through a PDA, the presentation may be controlled via PDA and full-media complex results may be displayed on a wall.
- **environment**: texts, images and other multimedia may be selected according to the environment. For instance: a contextualization of the NLG step allows to generate sentences .

If these requirements are taken into account, the personalization component has to fulfil the following tasks:

a) modeling "user in context" features
b) establishing if there are other devices, in the environment, that can be used to improve the quality of the presentation (especially if the user interacts through a small device). This task can be accomplished by interacting with a service agent (Device Manager Agent) that acts as a smart gateway between the D-Me presentation component and the environment;
c) establishing the information to be presented and how to organize it according to the "user in context" features;
d) setting up the output layout according not only to these features but also to the used devices.

The resulting personalized presentation, as a combination of user modeling and presentation techniques, is a function of several variables related to the type of environment and services it offers to the user. Therefore, a unique personalization policy is not feasible: each environment should be able to adopt the user modeling and personalization strategies that are more suited to the nature of the provided services.

For this reason, each environment agent uses its user modeling approach and presentation strategy. When a D-Me agent enters in the environment, these two components are *attached* to D-Me that computes the results, presents them to its "owner" and, if authorized, gives back a feedback to the environment.

In this way, every type of environment can use an appropriate user modeling and presentation approach, by leaving to the D-Me agent the computation.

3.1 User modeling

User modeling (UM) is crucial for the personalization of services and, therefore, of the results to be presented. The need to let the user free to interact with smart services everywhere and continuously in time [1, 16] produces obvious changes in the way the user modeling component has to be designed and developed. This has to be thought not as a part of a standalone system, but as an independent component able to provide its services to other entities that require them. In this optics, new problems and challenges arise:

- **UM location**: the user can move physically in different smart environments; this requires the management of different strategies for locating information about the user, so that this is available in every moment and place in which the user requires a personalised service.
- **Security**: the user model is not always used by the same system, but "moves" with its "owner" among different environments. This requires the need of establishing privacy and security policies.
- **Consistency**: the user may interact with more than one environment at a time; then, it is necessary to develop a strategy for keeping individual user information consistent.

Possible solutions to these problems are represented by a centralised, a distributed or mobile approach [11]. All these approaches presents advantages and problems. In traditional client-server information systems, the most frequent design choice is to store the User Model on the server side, by enabling the user to access his/her model after having been recognised by the system. In the distributed solution, user information are stored in different servers, reducing in this way the computational load. This is obviously a complex strategy that presents problems of redundancy/incompleteness of user information and consistency. In the mobile approach, the user "brings" always with her the user model, for instance on an handled device; when interaction with an environment starts, his/her profile is passed to the environment user modelling component. This approach presents several advantages: the continuous availability of user information, wireless data communication, absence of information redundancy and easy management of consistency. However, we cannot assume that the user will have with him/herself an handled device and this type of device, nowadays, still presents hardware-related limits (capacity, computational speed, battery,…).

In our architecture, the user is represented by an agent that dialogues, communicates and negotiates with other agents in the environment. In this view, the User Model should preferably be stored on the client side and should communicate the needed personal information to the environment. This approach is motivated by efficacy and privacy.

Once the users access through their D-Me Agent some information service, D-Me asks to a Remote UM Agent the data that are relevant to personalise the presentation for that particular service. Data are then passed to the environment User Modelling Agent (UM Agent) that starts the modelling process, as not all information slots are relevant for all information services provided by the environment.

The Remote UM Agent component has been introduced to overcome the location problem: in fact, what we mean as ‘client side’ may be an external server, the user car, a wheelchair , an handheld device,

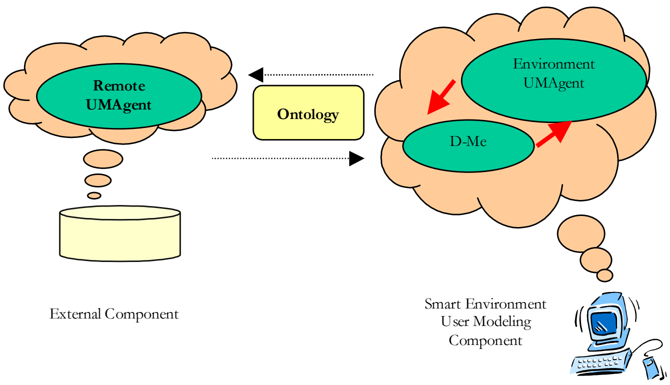

Fig.1 The UM architecture

etc.

Besides the long term factors in the user profile, when the interaction starts, the short term factors can be perceived by the D-Me Agent. This Agent will pass them to the modelling component, that elaborates them, together with the long term ones, during the whole session. When interaction ends, the environment sends back to the client the portion of the long-term model, updated according to what has been inferred (Fig.1).

The communication problems between the agents, interacting for accomplishing the user modelling task, is solved using the same approach of the underlying infrastructure: they talk sharing ontologies according to FIPA ACL. This allows to overcome problems of agents that use different representations of user profiles.

3.2 D-Me Presentation Agent

The Presentation Agent, that we implemented as a specialised service agent of our infrastructure, includes a *Mind*, that decides the presentation content, and a *Body* that displays the content according to the factors that influence the personalization [7]. In this way, each environment, according to the nature of the services it

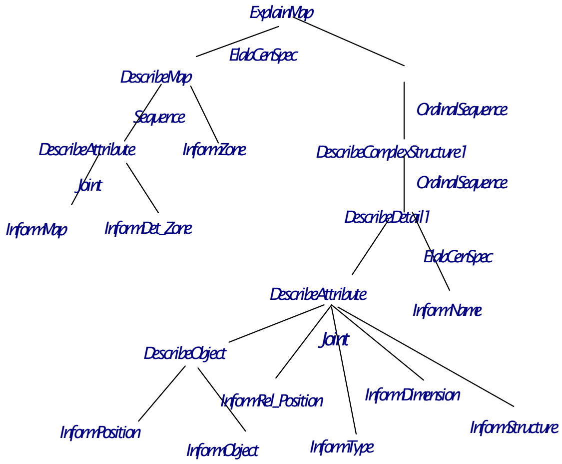

Fig.2 A plan describing an image

provides, can use its own strategy represented, for instance, as a plan.

Like in most NLG systems [13], the Presentation agent applies the following strategy: XML-annotated filtered results are passed to its *Mind* that, using an approach based on ontology sharing like the one explained previously, understands the meaning of the data to be presented and uses it as a factual knowledge base for the selected presentation strategy. The result is a presentation plan that is then passed to the Body component, which decides how to render it at the surface level according to the current situation. For instance, the presentation agent could be embodied in a character and converse in a natural way to the user during the interaction or it could be just a voice comment to an hypermedia presentation.

In order to show an instance of this strategy, we provide an example of application in the context of tourist services provision.

Let us suppose that the user is accessing this type of information from a PC at home using the web. In this case, he/she is presumed to need general and detailed information about the place he/she is going to visit, accommodation suggestions, restaurants and so on. The main information presented and the follow-up suggestions may be adapted to his/her needs: for instance, budget limits or food preferences. Instead, in a mobile access to the same information service, the user is “immersed” in the environment and is presumed to look for “context-sensitive” information about a monument or a cheap restaurant that is close to where he/she is. In this case, showing a map of the place would help the user in locating exactly where the described object lies. As display of images on a small device like a cellular phone or a PDA is still a problem, contextualizing the information to be displayed might help considerably in improving the information presentation [2].

In the prototype that we implemented in this context, the Presentation agent uses the following strategy (see [5, 6] for more details):

Pass a] In order to support contextually the user by generating situated descriptions, *Mind* receives, as input, an XML representation of a map of the visited place and a set of results filtered according the "user in context" features.

To annotate maps, we use a language that we defined in another project for marking-up radiological images and storing them as a XML structures. Starting from annotated images, a NLG module was able to generate the corresponding user-tailored textual descriptions, that were employed for tutoring puposes[6].

This language allows to characterise objects of particular interest with their shape, texture, edges etc and the way they are aggregated into larger regions. In our XML-based image mark-up language, object aggregations are called *overlays*, objects are called *details* and their attributes are called *annotations.* Overlays, details and annotations depend, of course, on the particular domain to which the image refers.

Pass b] By associating this type of metadata with every image, we may reason on its content and generate user and context adapted descriptions using XSL-plans representing the structure of these presentations.

For instance, the plan in Fig. 2, when instantiated with the data related to an annotated map, allows to describe an item (i.e. a restaurant).

Pass c] The instantiated plan is then passed to the *Body* component that, according to strategies for producing the final output [5, 6, 13], decides:

- *how to aggregate plan nodes* into information units;
- *how to display an information unit* according to the device.

  For instance: a full-media web page includes a title, a body and a footer as typical tokens while, in a WML page, the title appears only at the beginning of the navigation session, the body tends to be composed of homogenous information items and the footer depends on the programmability of the softkeys, that is related to the telephone type. Linguistic transformations of information items can be more or less verbose and employ terms with which the user is familiar;
- *which media to select* for conveying every information token: some media are preferable to convey particular information tokens, according to the user characteristics and to the device employed. For instance: text is preferred over images on small devices; to this aim, textual descriptions may be generated from the metadata associated with the image;
- *which type and level of guidance* to provide during interaction. These features may be adapted to the user level of experience (i.e. guided-mode with strong orientation support for non-expert users, strong orientation support for instructional purposes, less orientation for information, etc: for more details, see [5]). The device influences, as well, this choice: the poor usability of some device (small display, limited bandwidth, small memory capacity) suggests structuring menus so as to propose a guided navigation structure with step-by-step interaction, feedback provided only in negative cases, appropriate use of softkeys, etc. [3, 15].

4 An Example

Considering a tourist service example, let's see how the Personalization component behaves in different situations as described in Table 1. These scenarios differ in: i) the type of user background and interests, ii) information access goal (*browsing* vs. *searching*) and iii) the device employed.

As far as the last feature is concerned, we will take two cases that, with the present technological offer, may be considered as representatives of the 'tabs' and 'pads' envisaged by Weiser [16]: a PC and a PDA. We are well aware of the fact that this is a rough image of the diversification of devices that will be available in a future society of 'Ubiquitous Computing': however, we claim that, even in diversified situations, the method we propose will still keep a validity.

Case 1a:	***Case 1b***:
Goal(U, Know-About (restaurants))	Goal(U, Know-About (restaurants))
Task(U, browse)	Task(U, search)
Device(U, PC)	Device(U, PDA)
Like(U, local food)	Like(U, local food)
Interested(U, modern art)	Is-Interested(U, historical-monuments)
	Is-Located(U, (x,y))

Table1. Two different interaction situations.

Let us suppose that the user is looking for a restaurant downtown. The system will answer to the query by presenting the list of restaurants available, with links to their description and to the description of that zone. If a PC is employed, the items will be ordered according to the user preferences and the description of the location (directions and places of interests) will be complete and exhaustive, by mentioning all objects, places and monuments of interest in that location. When access is through a mobile device, the user location is taken as a further adaptation parameter; besides ordering the list of restaurants according to her presumed preferences, the description of their location will be contextualized: the map will be focused on the place of interest in relation to the user position and objects, monuments and places of interests will be outlined in the map.

5 Conclusions

This work is our first step towards adapting the presentation of information to the 'user in context'. Even if the examples we illustrated are still displayed on a PC, a WAP-phone and on a PDA, we claim that the approach adopted is general enough to be device and environment independent. It allows to "plug-in" the different personalization components of the D-Me Agent (UM+ Presentation Agent) that are typical of that environment and to represent the Presentation Agent's 'Mind' decisions into different 'Bodies', by providing the right amount of contextualized information in an appropriate presentation form .

Acknowledgments
We wish to thank Fiorella de Rosis for her useful comments.

References:

[1] Abowd G.D and Mynatt E. "Charting past, present, and future research in Ubiquitous Computing". ACM Transaction on Computer-Human Interaction, Vol 7 n. 1, March 2000.

[2] Ardissono L, A. Goy, G. Petrone, M. Segnan, "A software architecture for dynamically generated adaptive Web stores", Proc. 17th International Joint Conference on Artificial Intelligence (IJCAI 01), Seattle, August 2001, 1109-1114

[3] Buyukkokten O., Molina H.G, Paepcke A., Winograd T. Power Browser: Efficient Web Browsing for PDAs. The 9th International WWW Conference (WWW9)., 1999.

[4] Dale R Oberlander J. Milosavljevic M. and Knott A. Integrating Natural Language Generation and Hypertext to Produce Dynamic Documents Interacting with Computers, 10, 1998.

[5] De Carolis B., de Rosis F., Berry D C and Michas I. Evaluating plan-based hypermedia generation. 7th European Workshop on Natural Language Generation. Toulouse, 1999.

[6] De Carolis B, de Rosis F, Pizzutilo S. Generating user-adapted hypermedia from discourse plans.AI*IA 97:Advances in Artificial Intelligence, LNAI M. Lenzerini (eds),Springer, 334-345, 1997.

[7] A. S. Rao and M. P. Georgeff", BDI-agents: from theory to practice". Proceedings of the First Intl. Conference on Multiagent Systems, San Francisco,1995.

[8] Hovy E. H. Automated Discourse Generation using Discourse Structure Relations. Artificial Intelligence, 63, (1993), 341-385.

[9] http://www.fipa.org.

[10] Jörding, T. Temporary User Modeling for Adaptive Product Presentation on the Web. Proceedings of UM99, Banff, Canada, pp. 333-334.

[11] Kobsa A., Generic User Modeling Systems. UMUAI vol. II nos.1-2 pp.49-63. Kluwer Academic Publisher. 2001.

[12] M. Perkowitz and O. Etzioni. Towards adaptive web sites: Conceptual framework and case study.Artificial Intelligence, 2000.

[13] E.Reiter and R. Dale. Building Natural Language Generation Systems. Cambridge University Press. 2000.

[14] Rich, E.: User Modeling via Stereotypes, Cognitive Science, 3, pp. 329-354, 1979.

[15] Schmidt A., Schröder H and Frick O. WAP - Designing for Small User Interfaces. ACM CHI 2000 Extented Abstracts, Conference on Human Factors in Computing Systems. April 2000. pp 187-8. ISBN 1-58113-216-6.

[16] Weiser M. The Computer for the 21st Century. Scientific American, september 1991.