Modeling Context Is Like Taking Pictures

Daniela Petrelli Elena Not Carlo Strapparava Oliviero Stock Massimo Zancanaro
Cognitive and Communication Technology Div.
ITC−irst
I−38050 Povo − Trento − Italy
+39 0461 314 444
{petrelli|not|strappa|stock|zancana}@irst.itc.it

PREFACE

The discussion reported in this paper derives from our experience in designing and implementing two location−aware adaptive systems: HyperAudio [4, 13] and HIPS [5]. Both systems are hand−held electronic museum guides that adapt their behavior to the one of the individual visitor[1]. The common idea behind is to create an augmented museum where the main interaction modality is the physical move: visitor's movements are traced by the system and interpreted as implicit input. While moving in the augmented museum the visitor contemporary explores the associated information space [8]. Data are organized as an adaptive hypertext [2]: each node has a marker set describing content and form and link are labeled. Markers and labels are analyzed and used at run time to compose presentations on the fly. A presentation has an audio message and a set of suggested links; both are adapted depending on the context. Adaptation is realized taking into account the physical space, the history of interaction and the user model (or in a broader sense the visit model). Particularly important in our scenario is the interpretation of the visitor's physical position and movements seen as implicit interactions with the system. For example, when approaching a new exhibit a description is automatically provided; after a long stay in front of an object, another description is proposed. Thus the interaction is implicit since there is no intentionality on the part of the visitor to communicate his/her position or interest. The selection of the content and the linguistic form of the description are context sensitive. Phrases containing direct reference to the space ("in front of you", "this is"), reference to already seen objects ("you saw previously", "you just saw"), or suggesting new exhibits ("located behind you", "on the opposite wall") are introduced, on the bases of the context, to exploit presentation effectiveness.

[1] HyperAudio is a ITC−irst project. It adopts an"all−on−board" hardware solution; the prototype was realized for a Natural Science museum. HIPS, instead, is a European founded project based on a client−server architecture; it supports a visit in an artistic domain and allows for marking with spots space or objects. Respect to the discussion in this paper, none of the differences between the two systems is relevant since when talking about HIPS here we mainly refer to the sub−system that was under ITC−irst responsability, the one discussed in the architecture chapter.

WHY USE CONTEXT: TO SMOOTH INTERACTION

The objective of context−aware applications should be the one of making the technology invisible. In his latest book *The Invisible Computer* [7], Donald Norman says that "We need to move to the third generation of PC, the generation where the machines disappear from sight, where we can once again concentrate upon our activities and goal in life." Context−aware applications are going towards this wish of unaware use of computers, but this goal would be reached only if the modeled context completely and reliably would represent the situation. A disappearing museum guide would be the one that you can forget to have, one that taken at the entrance and put in the pocket, would fully support a free, natural visit providing the most appropriate information at the right time and place. Then the only activity would be the enjoyment of the exhibition: the interaction would be with the (augmented) museum, not with the guide.

WHO MAY BENEFIT: BOTH SYSTEM AND USER

As between people, interactions between humans and computers requires mutual intelligibility, and shared understanding. The context of interaction as source of information brings an enlargement of the communication bandwidth between human and computer.

From a system point of view, this means to have much more data available to interpret correctly user actions and deciding the most appropriate reaction.

From a user's point of view, this means a smoother interaction since less explicit and conscious actions are required. As a consequence, the user can concentrate more on the ongoing activity, almost forgetting that a computer is there. In our scenario, for example, the physical movement is an (implicit) input: approaching an object triggers the system reaction. Of course all this works if the assumption that behavior in museum (i.e. visitor's movements) represents personal attitude toward the exhibition is true. That is the point: assumptions have to be the right ones and have to be used at the right time and in the most appropriate way.

WHAT IS CONTEXT: POSSIBLY ANYTHING

A full understanding of human activities requires the situation of users' actions in the context of particular, concrete circumstances. The term *situated actions* [15] well represents the fact that "every course of action depends in essential ways upon its material and social circumstances". Anthropologists define this broad view *holism*, "the idea that any and all aspects (of a culture) are related. Not all aspects (of a culture) are smoothly integrated, but they do connect to one another, sometimes in relations of conflict and contradiction." [6]. Our experience in modeling the context match this ethnographically−oriented view of human−computer interactions that subsumes integration and synergism but also conflict and contradiction among the factors that characterize the environment.

The context we should consider in human−computer interaction includes both material and social circumstances. Aspects like the place of use (in an office, at home, in a museum, in the open air), the device (a workstation, a hand−held device) or the available infrastructure (networks, GPS, infrared) are relevant, but equally important are social aspects (being alone or not, who the others are, if this brings pressure) and personal traits (attitudes, preferences, interests). To represent this holistic view, not only facts (i.e. the physical place), but also system assumptions (i.e. the user interest model) have to be included in the context.

Facts are not always tangible, as is a physical space. The history of interaction until now is a precious source of facts. For example for an electronic museum guide it is fundamental to know if the visitor is back to a place already visited or if an exhibit related to the current one was already seen or not before. Given these two facts the system avoids repetitions and explicitly refers to other objects, recalling what was already seen or suggesting where to go next.

Significant facts can be collected directly from the user: the key point is to select the few, relevant ones. HyperAudio gives an interesting example [14]. Families and classes were up a two−thirds of the total of natural science museum visitors. Both type of groups have internal dynamics that make the museum visit completely different from individual or adult−group visit. Moreover who is at his/her first visit behaves differently from a frequent visitor. Knowing in advance these attributes, collected through a very short questionnaire, allows HyperAudio to adapt from the very beginning by selecting different contents (anecdotal vs. scientific) or setting a different system behavior (automatically starting new presentation vs. signaling that new information is available) or regulating presentation length and detail depth.

Besides facts, system assumptions reveal relevant context factors. System assumptions includes planning choices and dynamic modeling of the ongoing interaction. Planning choices locate the system's behavior in relation to contextual conditions. For example, the precise positioning of the visitor in relation to an object allows to identify two different visitor's status: who is close to the object is assumed to be looking at it; while who is facing the exhibit from a distant position can be distracted. Thus in the first we suppose that the object is already in the visitor's attention focus, while for the second the objective is getting visitor's attention. Given these context features, at run time, the system prefers "this is" when the object is in focus, while "in front of you" is preferred to attract attention.

A second type of system assumption is dynamic user modeling: it represents what the system supposes is going on. In HIPS, for example, the number and length of stops, as well as the sequence of moves are used by the Visiting Style Model to represent visitor's attitude towards the exhibition [10]; presentations and visitor's reaction to them are used instead by the User Interest Model to make hypothesis on what type of information or exhibit is the most preferred one [11]. User Models are critical points since the system's "guess" is based on the correct interpretation of the user's action and they strongly influence system reaction.

Finally, for linguistic applications, like those discussed, the discourse context has to be considered. In semplified terms, the discours context contains both the communicative context (the world around speaker and hearer) and the textual context (the text that surrounds a certain utterance), also called co−text [12]. Let clarify the concept with two examples. A reference to an object in the physical space like "*this is* the great fresco La Maesta'" or "*in front of you there is*..." concerns the communicative context since a correct interpratation is given only merging the utterance with the environment. An example co−textualization occurs when a pronoun substitutes the subject like in "*It* is one of the absolute masterpiece of the european gothic painting.". The comprehension of this phrase requires to know the previous utterance. In our systems the discourse context includes all the features discussed until now plus some co−text features. The effect of a correct use of the discourse context is a higher fluency of presentations.

Physical environment, hardware and devices, explicit and implicit interaction, user models, social environment, personal traits: context includes potentially everything. Moreover time is a fundamental variable since the context is by definition unstable and changing. A given configuration of context is valid only at a given time.

Other scientific communities are investigating the context concept, mainly its computational side [1]. Our commitment should be the one of investigating how the context affects human−computer interaction. In this perspective, unlike other authors [3] who states that context is local to entity (i.e. person, place or objects), we think that only the person who is interacting with that specific application at a certain time is relevant. In this perspective our definition is: *a context feature is any information that can be used to characterize and interpret the situation in which a user interacts with an application at a certain time.*

As all the definition of context, also this one risks being too broad for being on any usefulness. Reducing the number of aspects the application has to consider seems to be a must.

HOW: AN IMPLEMENTED ARCHITECTURE

HyperAudio and HIPS are both systems able to consider the context of interaction and change their behavior depending on it. Their architecture is fairly complex [9], but, as most augmented environment systems, the core is a loop triggered by (implicit or explicit) user input (fig. 1). Besides the modules here described, there is the localization sub−system that, using different technologies, traces the visitor's movements in the environment.

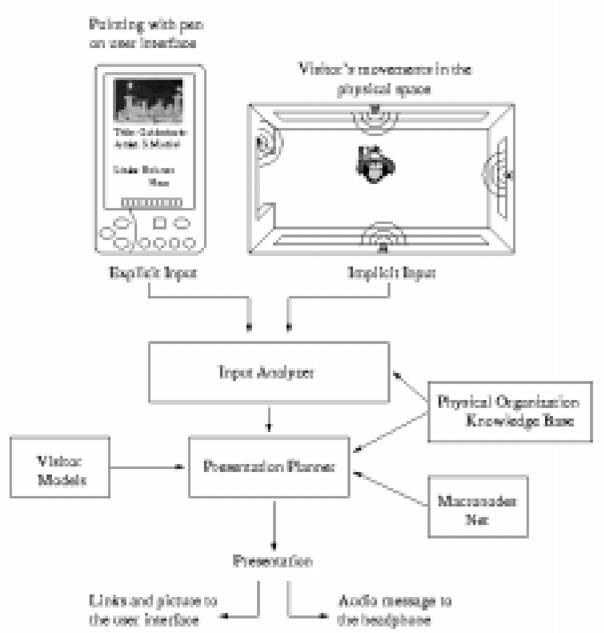

Fig. 1 The (simplified) implemented architecture.

The *Input Analyzer* is in charge of interpreting the visitor's interactions and deciding if a reaction is appropriate. In general, it triggers the Presentation Planner when the visitor approaches a new exhibit, when the visitor stays for long in the same position (and no presentation is running), or when a link is explicitly selected.

The *Visitor's Models* box hiddens many moduls. They contain the facts the system knows about visitors (i.e. where they have been, what they have listened at) and the suppositions the system makes on the visitor status. These models can be very different from one another. HyperAudio, for example, has a static profile (set by the questionnaire discussed above) that describes the potential visit. HIPS, instead, has mechanisms to infer characteristics of the visitor such as attitudes, interest and knowledge [10, 11].

The *Interaction History* records both the places visited and the information received. The order is particularly inportant since it affects the discource context.

The *Physical Organization Knowledge Base* contains a description of the space in terms of rooms, areas, and exhibits. Given a certain position and orientation of the visitor it provides the reference to the exhibit(s) the visitor is in front of and his/her distance from it (them).

The *Macronodes Net* is the data structure. It is a kind of adaptive hypertext where the nodes are marked and link labeld in order to allow the system to select the most appropriate information for the current situation at run time. Each node is about a single concept. A concept could correspond to an exhibit in the physical space, but could also be "abstract" (e.g. the author, the blazon). The correspondence exhibit−concept is what allows to anchor the information space (the Macronodes Net) to the physical space (the Physical Organization Knowledge Base).

The *Presentation Planner* is the module where the context is used to plan the system reaction. Here the context and the co−text are fused in the discourse context. On this bases, the Presentation Planner decides what to present to the visitor's; it composes a presentation structure that is sent to the hand−held device for delivery.

As the Input Analyzer, the Presentation Planner is a rule based module. Rules are composed by condition and action; when the condition matches the context the rule is fired and the system performs the action. The condition part encompasses tests on context variables. There is no limit in the number or type of context features tested. The context can be imagined as a set of variables, continually updated. It includes the visitor's position and the exhibit the visitor is looking at, the history of interaction (places visited and information received), the discourse context (mainly if a concept was given and when), and the status of the visitor's models (visiting style, knowledge and interest in HIPS, visit profile in HyperAudio). The behavior of the system is determined by rules organized in clusters. Each cluster is applied to accomplish a certain action. There is a cluster to decide what is the exhibit in focus; one for selecting the information that will compose the audio message and finally another set for selecting the suggested links. An additional cluster is used by the Input Analyzer to decide if the user input is meaningful (i.e. if a reaction of the system is appropriate).

WHERE AND WHEN USE IT: TRY AND TEST

Deciding the strategies the system has to apply to build presentations and when to use them are the hardest points. The fact that human actions are affected by computer behavior and that computer behavior is not random, but planned in advance, make the design step the key point. From a technical point of view, a rule−based system offers the flexibility required to be context−sensitive in the broadest sense since many context features can be considered and evaluated at the same time. Moreover this architecture is very flexible and easily supports a prototyping approach: system behavior can be changed changing the rules, not the system code. Thus the problem of defining system behavior is equivalent to defining the rules and the priority among them. Unfortunately, rephrasing Nardi from above, any and all aspects of the context are related, but not all aspects of the context are smoothly integrated; they do connect to one another, but sometimes in relations of conflict and contradiction. To clarify this point let consider the following example. When approaching a new exhibit the system reacts sending a presentation, but this is not always the right choice. Consider the situation in which the visitor is listening to a presentation and moves towards a new exhibit; should the system stop the current presentation or not? should it start a new presentation or not? The most appropriate action to take may depend on other factors than visitor's movements, for example the type of presentation currently running: if it concerns the object previously on focus it has to be interrupted (or better it has to be shortened) since references to the space are no longer valid, but if the content of the presentation is general, then a full delivery is an acceptable behavior (the new presentation could be queued).

CONCLUSIONS

There is a user, there is the application, there is the world around them. The context is the world around them. Implementing a context−aware application means reduce it in both dimensions of width and depth. Modeling context is like taking pictures: there is the risk to leave out of the frame something important (reduction of the width) or that something is not as relevant as it seemed (reduction of depth). A careful design and a prototyping phase is needed to select what is actually important and avoid that "the (relevant) context is what you don't use but you should"[2].

ACKNOWLEDGMENTS

We thank all the people who participated in the HyperAudio project and the HIPS partners for the valuable discussions and work.

REFERENCES

1. Bouquet, P., Serafini, L., Brezillon, P., Benerecetti, M. and Castellani, F. (eds.) Modelling and Using Context. *Proceedings of the 2nd International and Interdisciplinary Conference*, *CONTEXT99*, Lecture Notes in Artificial Intelligence, no. 1688, Springer.

2. Brusilovsky, P. Methods and Techniques of Adaptive Hypermedia. *User Modeling and User Adapted Interaction*, 1996, 6, 2−3.

3. Dey, A.K. and Abowd, G.D. Towards a Better Understanding of Context and Context−Awareness. GVU Technical Report GIT−GVU−99−22, College of Computing, Georgia Institute of Technology. (ftp://ftp.cc.gatech.edu/pub/gvu/tr/1999/99−22.pdf)

4. HyperAudio home page:

   http://ecate.itc.it:1024/projects/hyperaudio/

5. HIPS home page:

   http://marconi.ltt.dii.unisi.it/progetti/HIPS/

6. Nardi, B.A. The Use of Ethnographic Methods in Design and Evaluation. *Handbook of Human−Computer Interaction*, second ed. 1997, 361−366.

7. Norman, D.A. *The Invisible Computer*. 1998, MIT Press.

8. Not, E. Petrelli, D., Stock, O., Strapparava, C., and Zancanaro, M. Augmented Space: Bringing the Physical Dimension into Play. in *Proceedings of the Flexible Hypertext Workshop*, held in conjunction with the 8th International Hypertext Conference − Hypertext97, 1997.

   (http://www.mri.mq.edu.au/~mariam/flexht/)

9. Not, E., Petrelli, D., Sarini, M., Stock, O., Strapparava, C., and Zancanaro, M. Hypernavigation in the Physical Space: Adapting Presentations to the User and to the Situational Context. in *The New Review of Hypermedia and Multimedia*, 4, 1998, 33−45.

10. Marti, P., Rizzo, A., Petroni, L., Tozzi, G. and Diligenti, M. Adapting the Museum: A Non−intrusive User Modeling Approach. in *Proceedings of UM99* 311−313.

11. Sprecht, M. and Opperman, R. User Modeling and Adaptivity in Nomadic Information Systems. In *Proceedings of i3 Annual Conference*, 1999, 65−68.

12. Petofi, J.S. *Semantica, pragmatica, teoria del testo.* In Conti M.E La linguistica testuale, Feltrinelli Economica, 1977, 195−223

13. Petrelli, D., Not, E., Sarini, M., Stock, O., Strapparava, C., Zancanaro, M. HyperAudio = Location−awareness + Adaptivity. in *Extended Abstract CHI99* (Pittsburgh, May 1999).

14. Petrelli, D., De Angeli, A. and Convertino, G. A User−Centered Approach to User Modeling, in *Proceedings of UM99* (Banff − Canada, June 1999), Springer−Verlag, 255−264.

15. Suchman, L.A. *Plans and Situated Actions.* Cambridge University Press, 1987.

[2] From a free−speech of Pedro Domingos at the workshop on "Learning and context" at ICML96, International Conference on Machine Learning, Bari, Italy.