Brain Inspired Cognitive Systems
August 29 – September 1, 2004
University of Stirling, Scotland, UK

CONTEXT BASED MESSAGE SELECTION STRATEGIES IN A BIOLOGICALLY INSPIRED AMBIENT INTELLIGENCE SYSTEM

S. Piva,
D.I.B.E. - University of Genova via Opera Pia 11, 16145 Genova, Italy

L. Marchesotti(1), C. Bonamico(2), C.S. Regazzoni(1)
(1)D.I.B.E. - University of Genova via Opera Pia 11, 16145 Genova, Italy
(2)C.N.I.T. Genova Research Unit via Opera Pia 11, 16145 Genova, Italy

ABSTRACT

The presented work is concerned with the wide research field of the Ambient Intelligence systems. These systems, also called Smart Spaces (SS), are designed to provide users with augmented services that do not need intensive and invasive interaction efforts. To obtain this, SS main functionalities are *sensing, analyzing, deciding, acting / communicating*. In the scope of this wide research field, this paper particularly deals with the analysis task through the study of a theoretical biologically inspired brain model and with the deciding and communicating tasks through the description of a rule engine-based message customization system. Virtual Characters (*avatars*) displayed on the user's device represent the final front-end of the system. This choice is due to the high expressive capabilities guaranteed by this kind of communication means, able to enrich the message content with prosody, facial expressions, gestures and to make the communication more effective. Through examples and architecture diagrams, our solution for the problem of the context awareness-based parameterisation of the system reaction is explained.

INTRODUCTION

Nowadays research in the field of Information and Communication Technology (ICT) are going in the direction of integrating a wide range of complementary technologies to expand user's range of tendered services. One of the results of this attempt is the increasing publication of works related to the concept of Ambient Intelligence and to Smart Space systems.
Systems of this kind are designed to understand the state of a particular environment and appropriately react by instantiating a customized communication with each of the users who populate the monitored environment.
In this work we particularly face issues related to the decision and the communication with the user. Feedback for users has to be personalized with respect to the knowledge of user typology, profiling info and of the context surrounding him and influencing the system.
ISTAG gives in [1] a more formal definition of Ambient Intelligence (AmI) that points out how it should provide technologies to surround users with intelligent sensors and interfaces and to support human interactions. According to this, Brooks in [2] states that an Intelligent Environment has to make computation "ready-at-hand", putting computers out into the real world of people more than people into the virtual world. In these systems the central concept of *context awareness* represents the possibility for the system of biasing itself and its reactions to the environment. This means acquiring and processing knowledge of many static and dynamically changing parameters [3].
In this work we explore the idea of exploiting context awareness to condition the system's reactions. Data are collected, processed and a context label for interesting events or situations is produced through a context manager. Then this information is ready to control the system reaction through the appropriate physical actuators or a message content selection strategy (in this work we concentrated on the message customization). This is the way the system employs contextual data to obtain user-centered adaptivity; with this respect adaptivity becomes key of the problem: the user must be surrounded by a system that, knowing him, works to improve his capabilities, to extend the number of the services he has access to and to make these services appropriate for him.
The final goal, here, is to describe a context based message / action adaptation. The chosen communication interface is based on a Virtual Character (i.e. an *avatar* [4]) because the character can adapt not only message content but also the related emotive info by means of non-verbal modalities (prosody, facial expressions, gestures). In the last part of the paper the description concerns the definition of a rule engine and the concepts supporting the design of a set of rules, to demonstrate the effectiveness of this user centered approach.

CONTEXT AWARENESS

The central idea of the user centered adaptation is based on the context information exploitation. But what do we mean with *context*?

To give a more formal definition of the context-related issues we can say contextual information can be defined as *an ordered multilevel set of declarative information* concerning events occurring both within the sensing domain of a SS and within the communication and actions domains of the SS itself. Relations among events are also included (explicitly or implicitly) within contextual information.

An *event* can be defined as the occurrence of some fact that can be perceived by or be communicated to the SS; an event is characterized by attributes that basically answer questions about 'where' (position) and 'when' (time) the event occurred. Other attributes involve 'what' (core) consists the event of, 'who' (identity) is involved in the event, and possibly 'why' (reason) the event occurred.

Events can be used to represent any information that can characterize the situation of an interacting user as well as of a SS part, i.e. an entity. An entity can be a person, a place, or an object that is relevant to the interaction between a user and an application. The user and the SS parts themselves are entities. The multilevel nature of contextual information is related to the possibility of detecting and representing events at multiple abstraction levels.

Context-awareness refers to the property of a SS to represent internally in terms of events the state of its users, of their surroundings and of SS parts, as well as to be provided with rules that make it possible its behaviour to be adapted accordingly. Therefore, in a context aware SS, contextual information can be either used by itself, i.e. it can have its own value for the interacting user, or it can be used to select which services or which parametric form of given services can be provided to the users, i.e. it can be used to optimise the adaptation and personalization process in such a way to maximize the service value.

Pascoe [5] introduced a set of four context-aware capabilities that applications can support:

- *Contextual sensing*: a system detects the context and simply presents it to the user, augmenting the user's sensory system.
- *Contextual adaptation*: a system uses the context to adapt its behaviour instead of providing a uniform interface in all situations.
- *Contextual resource discovery*: a system can locate and use resources which share part or all of its context.
- *Contextual augmentation*: a system augments the environment with additional information, associating digital data with the current context.

This approach results particularly useful in the design of a complex distributed systems characterized by a high-level analysis and interaction capabilities as the one described within the scope of this paper.

THE INSPIRING MODEL

In order to enhance the level of *context awareness* in the design of the SS internal structure, a biological model has been chosen [6]; following this approach, a primary distinction is made upon state of the monitored objects and system's inner state. In figure 1 Damasio's model is depicted, as it can be seen, a module a called *Proto-Self* is inserted as an interconnected and temporarily coherent collection of neural patterns which represent the state of the system. Proto-Self constitutes the basis for the Core-Self, which is generated for any object that provokes the core-consciousness mechanism [7] that is to say the continuous environment awareness due to the analysis of stimulating objects. Because of the permanent availability of provoking objects, it is continuously generated and thus appears continuous in time. The mechanism of Core-Self requires the presence of Proto-Self and generates as output inferences on the current situation, which are mapped on the long term memory (i.e.: A*utobiographical Memory*). The Damasio Model represents the kernel of the proposed architecture which can be described in terms of the user-centered closed loop outlined in figure. As it can be seen the final aim of this approach is to mimic the behaviour of humans during their normal activities. In particular four main steps have been inserted to summarize possible states of the system: Sense, Analyse, Decide, Act & Communicate.

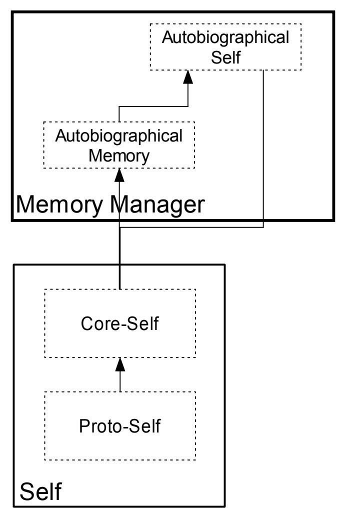


Figure 1:Interpretation of the Damasio's Brain Model.

The next step in the design of the architecture is to map the Damasio model in a logic functional schema according to the aforementioned closed-loop approach. The result can be seen in fig. 1 where a primary distinction is made between the External World (i.e. *Proto-Self*) and the Internal World (i.e. *Core-Self*). In this scenario, SS users are part of the External World, and have to be localized, and possibly identified, within the scope of the system. In this global context, two typologies of users can be present into the space: *registered* users, namely users equipped with multimodal communication devices already registered in the Smart Space (for instance, an intelligent laboratory) and enabled to get its services and *guest* users which are not equipped with or owns unknown devices.

The duality between Internal/External World into the sensorial set up of the architecture is reflected into Observation (*Eso-*) Receptors which are specifically designed to sample the External World whereas State (*Endo-*) Receptors are able to measure internal variables constituting the internal state of the system.
Endo/Eso Receptors outputs are gathered in a module called Analysis and Data Fusion; an associative memory (based on a Self-Organizing Map [7]) is here used to map and coherently integrate information on the basis of the measured data and according to a certain behaviour policy of the intelligent environment that has to be specified during the SS design. The associative memory module is able to output a semantic representation of the event currently taking place in the environment of interest (i.e. arrival state, intense human work activity, etc.).
On the basis of sensors observations, fused sensors observations, detected events and situation assessed by data fusion block, a Decision Module takes the most appropriate decision. This module is implemented as a rule engine keeping into account past experiences, which are stored by the Memory Management module.
The decision taken results both into system's internal state changes and into interaction through system actuators and communication facilities. In the simplest cases, this module just selects one from a set of predefined reactions, while in more complex cases it is able to dynamically generate the multimode feedback that is then presented by the Virtual Character on the user's PDA. The complex process of delivering a personalized multimode feedback to mobile users is achieved through the close interaction of three modules.
In the first step, the Decision module produces an high level representation of the Virtual Character behavior, including both the message to be spoken to the user, in textual form, and XML tags describing the non-verbal behavior of the character, such as the facial expressions, or gestures to be performed.

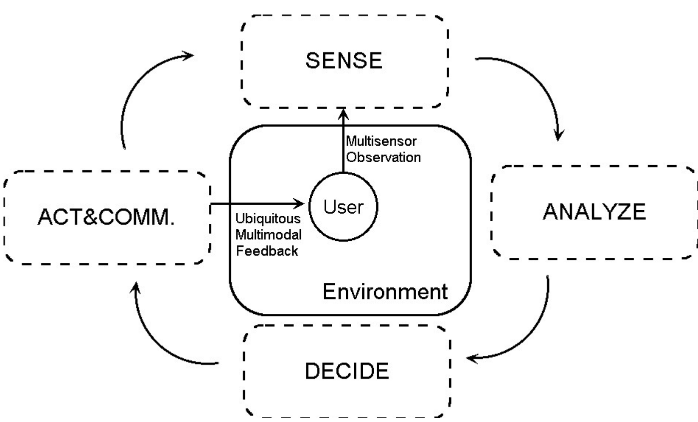


Figure 2: User centered smart space functionality approach

This high level representation is the input to a Multimodal Synthesis Engine, which is composed of several tools, each of which generates a specific communication modality. A Text-to-Speech engine synthesizes the character's voice, while a Text-to-Animation engine generates the character's facial movements, including both lip movements corresponding to speech, and expressions that reflect the system classification of an event (e.g., happiness, worry, surprise). The animation is encoded in a high-level parameter stream, which is sent together with the voice stream to the user's device, where an Animation Engine synchronously renders the message by animating and displaying a specific character model. Operatively, this conceptual model has been implemented in a system following the logic-functional architecture outlined in figure 2: different computational units (PC 1-2-3-4 and PDAs) have been used to support SS functionalities. A network infrastructure, based on wired LAN and an 802.11 WLAN, connects the different physical units and allow real-time data exchange. In particular tasks are allocated in 5 different computational units:

- **PC 1**: Image Processing and Resource Management algorithms implement the Sensing capabilities of Smart Space;
- **PC 2**: Information Conversion and Data Fusion methods complete the SS Sensing;
- **PC 3**: Rule-based engine & Databases;
- **PC 4**: Communication server that manages the generation of the multimodal output and the interaction with the mobile terminals;
- **PDA**: Personal Digital Assistant that maintains the communication with the SS infrastructure and it displays the audio/visual feedback to the user.

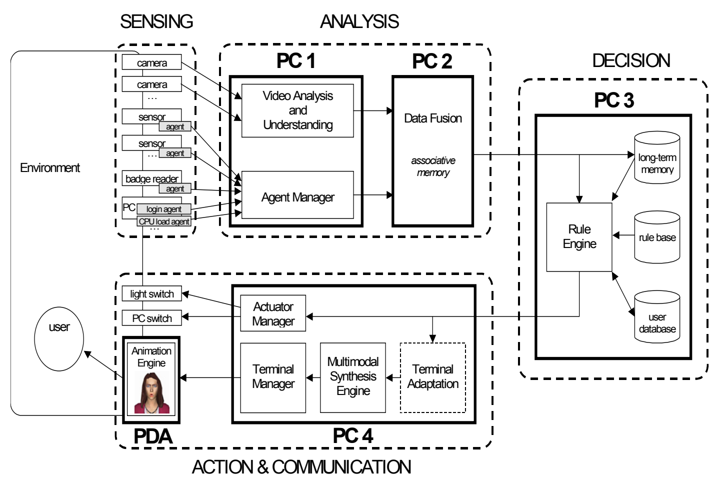


Figure 3: Smart space logical architecture and tasks

DECISION MODULE

In the process of conversion between multilevel contextual information and effective SS actions / communications, a central role is covered by the *decision module(DM).*
The importance of taking the right decision is crucial because its effects are directly reflected to users. Data available for decision are collected at different semantic level through the sensing modules and are then processed by the Proto- and Core-Self according to their nature and source. The ensemble of internal and external state information is encoded by the context manager (the module devoted to the contextual data extraction and processing – outside of the scope of this paper) and the decision manager modules in a hierarchical representation, where a *super-state* (e.g. ARRIVAL, meaning

new moving object entering the scene) gives the highest-level summarization of the context, while lower-level attributes describe the position and behavior of individual people, objects and resources (e.g., single PCs, parts of the system).
The reference hierarchy is the following:

1. *Event level:* super-states
2. *Behaviour level:* trajectories of users
3. *Object level:* identity of users
4. *Feature level:* number and position of users, system components status
5. *Signal level:* raw, non processed data

In such a structure, lower levels contribute to the estimation of the higher ones, a super state (an *event*) can be evaluated by considering Objects and related Features (i.e., the object's position). Events and Behaviour levels define the context in which objects (SS's users) act, whereas the remaining levels contribute to the definition of more specific and personalized description for objects.

Decisiom Module internal structure

In this work, the proposed approach is that of designing and developing the Smart Space decision module depicted in figure 3 basing on a rule-processing engine as represented in figure 4.

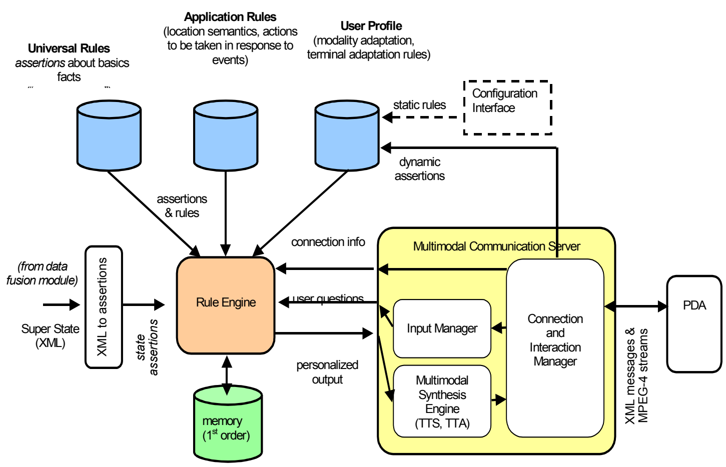


Figure 4: Architecture of the rule-based decision module for the reference Smart Space.

The rule engine receives updates of the system state and of the estimated ambience context from the video analysis modules and the agent network (*Agent Manager* - see figure 3 – not described in this work) that monitors resource activities. This hierarchical state description is provided in XML format. At the same time the rule engine interacts with the Multimodal Communication module that maintains the network connections with user terminals in order to receive updates about user preferences.
These two flows of input data are processed by means of a set of rules that define message selection and personalization strategies. The rule engine then produces an XML description of the message to be produced by the Multimodal Synthesis Engine, including a specification of the modalities to be generated, and the Virtual Character model to be employed.
The Multimodal Communication task finally manages the delivery of the personalized message to the user.

Message content selection strategies

The fundamental objective of message customization is to maximize the message usefulness, in terms of its effectiveness in directing user behavior in a constructive way. This effectiveness can be defined through a cost functional. This cost maximizes results depending on the situation or the application.
Usefulness may be defined in terms of a number of objectives, such as maximizing ambient safety and security, or minimizing waste of time for users. With respect to this goal, the use of the multimodal feedback provided by a Virtual Character in addition to conventional modalities (text, diagrams) is expected to provide a significant advantage both because users tend to perceive an embodied character as more trustful than a non-embodied system [8], and because it allows for exploiting the natural interaction capabilities and intuitive reaction to natural modalities (e.g., speech and facial expression) that we all develop through interaction with other people.
In fact, several studies have outlined how in human-to human communication, the informative content provided by non-verbal modalities is often more relevant than the phrase being pronounced.
For instance, in a dangerous situation, the user is likely to perceive more immediately the danger from the facial expression of a Virtual Assistant than from a long textual message [9].
In general, the goal of increasing the effectiveness of the feedback provided to the user can be subdivided in three main subgoals:

- adaptation to the user profile;
- adaptation to the situation/events detected by the system;
- scalability with respect to the user's terminal capabilities.

In this work we are particularly interested in the two first tasks while not dealing with the issue of user terminal adaptation.
The system must be able to select suitable messages and to adapt the message delivery modalities depending on its internal state an on its estimate of the current context of the Smart Space.
As an example, table 1 summarizes the main events a testbed we are building based on a university campus scenario must be able to detect and for which a message must be sent to an user. The table also presents the level of abstraction of the attributes that influence the message selection/generation.

Decision impact assessment

After the system decides for the correct action or the appropriate message or action and this action has been performed, the last step is to have a feedback from the user to assess the result of the decision. In the reference scenario this feedback has to be collected mainly in a passive way through analysis sensors not to involve the user. For example this means that if the system suggests to the user to follow a predefined path to reach a desired location the network of video cameras is used to detect if he/she does not correctly follow the path. Besides this state estimate, other information about the situation might be directly provided by users who can select or type in questions on their mobile terminals. For instance, the user might explicitly ask the system for help, or for guidance in reaching a room that is part of the controlled environment.

Event	Super-state level	Behavior level	Object level	Feature level	Other
A generic user enters a room	X				
A specific user enters a room			X		
User approaches PC				X	
User login				X	
User follows a given path		X			
User leaves given path		X			
Danger condition in a room	X				
User asks about availability of PCs in a room				X	user question on PDA
User asks how to reach a room					user question on PDA

Table 1

Adaptation strategies

The message forwarded to the user is adapted to the properties of the person that is the target of communications, and of the context into which he/she moves. Object properties are used together with Feature level properties, such as his/her position, identity, or behaviour, and contextual properties define the environment in which he is immersed.

To accomplish its adaptation duties the system must be able to distinguish among different users classes in the monitored environment. Each class can have different levels of access to the systems granted.

A different avatar, thanks to the customization of communications, is sent to the recognized user. The user profile becomes a fundamental feature to influence the message content because, for example in the cited hypothesis of the university campus, the student user can be easily warned that the lesson he is going to attend will be held in a different room with respect to previous communications. Furthermore the database can be widened with access permissions to the various campus areas and with all the possible information the system supervisor thinks to be useful to improve the service quality.

The second main objective for message personalization is that of following context-aware strategies, which take into account the "state" of the user, including, ideally, his current activity and goals, and the state of the surrounding ambience.

Taking into account the context attributes that the analysis and data fusion module being developed is able to estimate, we identified 7 main attributes of the context tree according to which the message can be personalized:

- time;
- location;
- situation of a room;
- resource status;
- other users' activities;
- user behaviour (cooperative/non-cooperative);
- number of times the same situation is repeatedly detected.

In order to allow for a more effective and reusable definition of adaptation rules, we followed the approach of trying to express the adaptation rules at the highest possible level of abstraction, possibly by means of the introduction of multi-level rules that process events at increasingly higher level of abstraction.

Concerning the first point, the system must be able to react differently according to the current *time*. For instance during work hours, guidance information could be provided to all users. After-hours, users could only be able to get guidance to emergency exits and to their own office/laboratory. A very important part of the context is the *user's location*, both at high level (building/room he is located) and at low-level (proximity to a resource, door, and so on).

A first general criteria is that of giving higher priority to personalization in term of lower-level location features, which means that "local" rules related to a specific place override default rules declared at the building level. A second criteria concerns directing users to the nearest possible resource corresponding to their needs.

The most important criteria related to the monitoring of *room situation* is the distinction between normal conditions and danger/emergency condition. For example in an emergency condition, all messages not relevant to the user safety will be hidden not to distract the user's attention.

For what concerns the awareness of *available resources*, the system should try to maximize first the user productivity, then the resource usage efficiency. A key criterion is obviously that the one of minimizing interference with *other user's activities*, particularly in the choice of presentation modalities. This means avoiding voice-based messages or talking-avatar messages when the user is in a room where several people are working.

The attribute of *user behaviour* is used to trigger functions such as the localization technique: a cooperative user provided with the appropriate device can support positioning task by the use of local dedicated software and hardware: a GPS receiver, for example.

Repetition of the same event, then, can be used to increase the weight of an action. This concept can be connected to the following last criterion: providing a *proportional feedback* to users, especially in cases when the system must outline an incorrect or forbidden behaviour on behalf of the user. The feedback must be, on one hand, proportional to the severity of the user's misbehaviour with respect to his own safety and to the overall system security. On the other hand, the intensity of the message needs to be progressively scaled according to the user's reaction to previous suggestions. This can be effectively

achieved using non-verbal modalities, and particularly facial expressions of the avatar: the first time the user contradicts the avatar, he can just show some surprise for the user not following, for instance, the suggested path. The following times the Virtual Character can show an increasing degree of angriness.
In addition to these general criteria, adaptation rules can be specified for a peculiar ambience or application. As an example, three instances of message adaptation referred to the reference University Campus scenario are shown in table 2.

Message description	Context Attributes	Cases	Personalized message	Presentation modalities
1.a Welcome to the building	Superstate Time of the day	*ARRIVAL morning*	Welcome to the Biophysical and Electronics Engineering Department. Press the help button on your PDA if you need assistance	smiling and talking avatar
		ARRIVAL night	Welcome to the department. Please make sure that you have the after-hours access permit with you at all times	serious talking avatar
	Superstate	*Danger detected*	You must not enter the building as an emergency situation has been detected	text message
2.a Number of PC available in the ISIP lab	User's location User's activity Resources	*User standing in front of the lab*	There are 2 PCs available, please log on workstation 3 on your left	Text-message and map of the lab on a public display at the entrance of the lab
		User walking in a corridor	There are two PCs available, do you want me to reserve one for you?	voice message on the user's PDA
5.a How to reach the ISIP lab	Number of times question is asked	*0*	Path to the lab	2D map
		1	I'm sorry I was not clear, please follow the corridor and turn right at the end	talking avatar with sorry expression
		2		

Table 2

Rules

The rule engine used here is the *Java Expert System Shell* (JESS) [10]. JESS uses an algorithm called RETE in order to apply rules to data in a very efficient way. A rule-based program can have hundreds or even thousands of *rules*, and JESS will efficiently/continually apply them to data in the form of a *knowledge base*.

The rule is divided in a left and right (Sx and Dx) part respectively corresponding to decisional data and action/message.

If <context-aware condition>
and <personalized condition> **and** <personalized condition>......... } Sx part
Then <action>
or <message> } Dx part

[Abstract scheme for a rule]

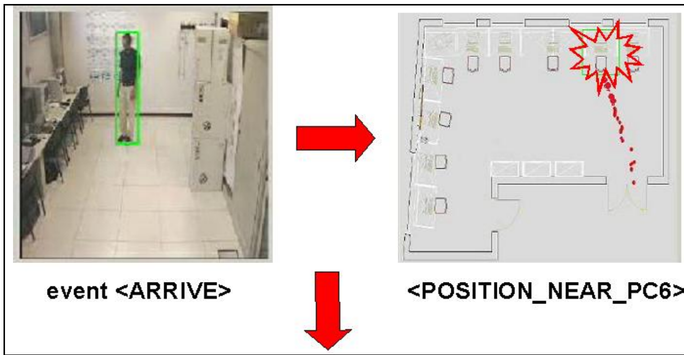
Specifically, if the result is a specific message for the user, a text-production function is activated, which, starting from a generic template for that message, derives a textual representation of the personalized message for the specific user to which the message is addressed.
The message personalization is achieved by expressing the message in the Virtual Human Markup Language[11]. VHML is an XML-based language, which allows for fully characterizing both the verbal and non-verbal behaviour of a Virtual Character. In this way, not only the words spoken by the Virtual Character, but also his expressions and gestures can be varied depending on the personalized conditions detected in the left part of the rule. The intensity and duration of the character expressions can also be varied by means of VHML tags: for instance, in the case of the system disappointment for the user performing a dangerous action, the intensity of the worried expression can be increased each time the message is repeated.

RESULTS

The general form of the rule becomes a practical application in the scope of the system realized to the described concepts. In figure 5 an example of the results of the study is depicted: the upper left image represents the detection of an ARRIVE event perceived through video eso-sensors and recognized through the context manager module. On the right side of the latter, the picture of the interested laboratory map evidences the trace of the detected subject (tracking phase) and the perception of the approaching to a particular PC (PC6 in the example). What follows is the application of the described rule and the consequence is a dissuasion message: the arrived student, Paul, is going to use the machine already assigned to another student. The avatar appears on Paul's device and tells Paul to move to another PC.

CONCLUSIONS

A theoretic approach for funding the artificial analysis and decision core for an ambient intelligence system on a neurobiologically inspired human brain model is described in this paper. The further step of exploiting the acquired environment and user contextual knowledge is particularly explored with the definition and the exemplification of a set of concepts, methodologies and rules. The aim is to demonstrate the concrete possibility to make a Smart Space employ a certain degree of context awareness to manage a message selection policy. The choice for the communication interface towards the user relies on the use of animated virtual characters, able to deliver emotional and prosodic add-ons to the message content. This demonstrates to make the communication more comfortable for the human user.



(1) If <ARRIVE> and <POSITION_NEAR_PC6> and <STUDENT>
Then <MESSAGE(leave)>

(2) Paul, <angry intensity ="50%" > you are trying to use a Pc reserved for other students</angry>. Please move to another PC <happy duration="500ms">

Figure 5: Results' example, application of a rule

ACKNOWLEDGMENTS

This work was performed under co-financing of the MIUR within the project FIRB-VICOM and the project Per2.

REFERENCES

[1] ISTAG Scenarios for Ambient Intelligence in 2010 http://www.cordis.lu/istag.htm

[2] Brooks, R. A. with contributions from M. Coen, D. Dang, J. DeBonet, J. Kramer, T. Lozano-Perez, J. Mellor, P. Pook, C. Stauffer, L. Stein, M. Torrance and M. Wessler, "The Intelligent Room Project". Proceedings of the Second International Cognitive Technology Conference (CT'97), Aizu, Japan, August 1997.

[3] Thad Eugene Starner ,"Wearable Computing and Contextual Awareness" Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning,in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the MIT June 1999.

[4] J. Ostermann, "E-Cogent: An Electronic Convincing aGENT", in MPEG-4 Facial Animation: The Standard, Implementation and Applications, I. S. Pandzic and R. Forchheimer, Eds., 2002.

[5] Pascoe, "Adding generic contextual capabilities to wearable computers", in Proceedings of 2nd International Symposium on Wearable Computers, October 1998, pp. 92-99

[6] Antonio R. Damasio, "The Feeling of What Happens-Body, Emotion and the Making of Consciuousness", Harvest Books; Sept. 2000.

[7] L. Marchesotti, S. Piva, C.S. Regazzoni, "Structured Context Analysis Techniques in a Biologically Inspired Ambient Intelligence System", Proceedings of the IEEE on System Man and Cybernetics, Special Issue on Ambient Intelligence (2004 to appear).

[8] C. Pelachaud, V. Carofiglio, B. De Carolis, F. de Rosis, I. Poggi, Embodied Contextual Agent in Information Delivering Application, In Proc. of AAMAS 2002, Bologna, July 2002

[9] Timon C. Du, "Implementing Association rule tecnhniques in data allocation of Distribuited Database" Industrial Engineering, Chung Yuan Christian University, Taiwan

[10] The Java Expert System Shell. http://herzberg.ca.sandia.gov/jess/

[11] A. Marriott, S. Beard, J. Stallo, and Q. Huynh, "VHML - Directing a Talking Head", Proc. of The Sixth International Computer Science Conference. Active Media Technology, Hong Kong, 2001 VHML, 2001. http://www.vhml.org