Dialog Manager Concept

Communication for quality of life.

Project arconym:	ALIAS
Project name:	Adaptable Ambient Living Assistant
Strategic Objective:	ICT based solutions for Advancement of Social Interaction of Elderly People
Project number:	AAL-2009-2-049
Project Duration:	July, 1st 2010 – June, 30th 2013 (36 months)
Co-ordinator:	Technische Universität München
Partners:	Technische Universität Ilmenau Metralabs GmbH Cognesys GmbH EURECOM Guger Technologies Synergiums S.A. pme Familenservice YOUSE GbR

D3.1

Version:	1.05
Date:	2010-10-26
Author:	B. Schönebeck (COG)
Dissemination status:	PU

This project is co-funded by the Ambient Assisted Living (AAL) Joint programme, by the German BMBF, the French ANR, the Austrian BMVIT and the Luxembourgish Luxinnovation and FNR:
http://www.aal-europe.eu

D 3.1	Executive Summary
This deliverable presents the basic structure of the *dialog manager* concept, central in the WP3. The first task was to specify the different software and hardware modules of the partners involved in the development of the dialog manager. After that, a common picture of the entire concept was sketched so that all partners can identify their modules. This is required to develop the interfaces and connections between the modules during the next project months.	

Dissemination Level of this deliverable (*Source: Alias Technical Annex p20 & 22*)	
PU	Public
Nature of this deliverable (*Source: Alias Technical Annex p20 & 22*)	
R	Report

Due date of deliverable	30.09.2010
Actual submission date	26.10.2010

Authorisation			
No.	**Action**	**Company/Name**	**Date**
1	Prepared	COG / B. Schönebeck	20/10/2010
2	Approved	TUM-MMK / J. Geiger	21/10/2010
3	Updated	EURECOM / R. Troncy	15/11/2010
4	Updated	COG / S. Lapp-Emden	10/12/2010

1. Introduction

The WP3 "Dialogue Manager includes the following tasks:

- User identification (via speech or face recognition)
- Knowledge representation
- Development of the dialogue system
- Development and integration of a game collection
- Web 2.0 wrapper
- Linking verbal/conceptual representation through sensor inputs (robot) and symbolic and/or alphanumeric inputs
- Adaptive and proactive behavior of the robot platform
- Linking CES (CES: Cognitive Ergonomic System) Core System (natural language understanding, cognesys) and Speech recognition (VODIS-server, Synergiums)
- Physiological monitoring / integration into the dialogue manager.

The aim of the deliverable D3.1 "*Dialog manager concept*" is to establish a general conceptual basis to connect all different knowledge bases and subsystems, which will be developed and implemented by the project partners who participate in the dialog manager.

2. Developments on the concept

In preparation of this deliverable, the first task was to find out which information from the modules will be necessary to define the robot behaviour.

The adaptive, proactive and polite robot behavior requires at least three sources of information:

Input from the human (human wishes, intentions, commands).

1. Input from robot sensor systems (mainly spatial orientation).
2. Knowledge about human needs / Knowledge about available net-based services and helpful tools to assist the user.

The first point - human input - will be given via voice (speech recognition), keyboard, mouse/touchpad and brain interface. The second point - robot spatial orientation - will be based on the sensor system input, but has to consider the human intentions too. The third point will be based on user inclusion, but also depending on programs, games and tools, accessible via the web.

3. The challenge: communication between different knowledge bases

The speech interface will be realized by the partner Synergiums. The brain interface will be delivered by the partner Guger technologies. Both systems are highly complex modules that require their own specific knowledge bases.

The movement of the robot and its spatial orientation require sensor system input, connected to special knowledge bases that will be developed by the partners TUM and Metralabs.

Net-based services, provided by the partner Eurecom, will be developed based on user needs requirements identified during user studies performed in the WP1. Furthermore, specific multimedia knowledge bases will be developed to enable event-centric interfaces for sharing multimedia content.

The day to day vocabulary and also the necessary world and expert knowledge, based on needs of different user groups, will be implemented in additional knowledge bases. The dialog manager will integrate all of those different systems, knowledge bases and modules.

4. The solution: "understanding and deciding" by the CES core system

Information from the three sources mentioned above (human Input, robot Input, human needs and knowledge concerning net-based services) will be collected by a so-called Cognitive Ergonomic System (CES).

Operating mode of CES: The concept-based knowledge base of CES is capable to integrate various types of information. First, the incoming information is translated into concepts. Using inferences and other cognitive operations on a conceptual level, an integrative representation of the information will be constructed in a second step. Based on this resulting "situational model", CES will be able to understand the situational context of robot actions, desired by the user. Finally, the appropriate robot action will be selected from the pool of all possible actions. This decision will be made by CES using human-like cognitive operation resulting in human like and intuitively predictive robot behavior.

The overall framework is presented in Fig. 1.

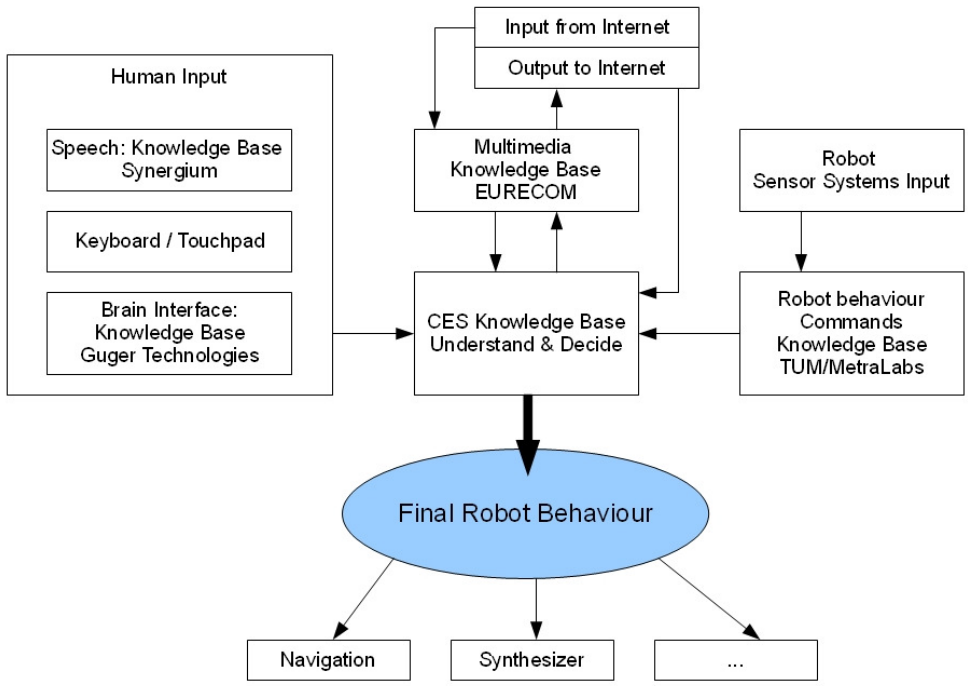

Figure 1: Overall architecture of the Dialog Manager

5. Summary and conclusion

The dialog manager must be able to integrate all information from at least three sources: human input, robot input and knowledge about available services. These sources of information come in different manner requiring to build separated knowledge bases. A fast and easy communication between these various knowledge sources is therefore essential. This will be achieved using the unique conceptual knowledge base and the cognesys core system as a central “translation” unit. Linking the various knowledge bases is challenging and requires a good "teamwork" of all involved modules. Finally, we expect the CES to be able to “understand” the user intention and to decide which robot action is appropriate in a given context.

The continuing task will be to develop interfaces, connections and technical commands with which the modules and the CES System will communicate.

6. Further processing

This deliverable describes the general concept of the dialog manager on which the further work will be based and proceeded. It is a high level concept at this stage and therefore it does not go into details. The details like descriptions of every module, the interfaces with the dialog manager and a detailed technical draft will be implemented in the next work steps. The progress of these tasks will be integrated in the next deliverables:

D3.4 Documented prototype of dialogue manager (software) – M8
D3.6 First dialogue system with integrated physiological monitoring – M12
D3.7 Dialogue system updated to user´s needs – M23
D3.8 Final dialogue system with integrated physiological monitoring - M32

Annex:

Example of a situation of robot behaviour and navigation

PREMISE 1: Robot microphone is on

PREMISE 2: Keyword "Robby" turns on speech recognition unit / or speech recognition is always on

1. Speech input from user: "**Robby, is there any new post from my grandson? I would like to see again the photos taken last month when Max gathered the whole family.**"
2. Speaker recognition module: identify the user.
3. Vodis filters the recognition hypothesis coming from the Speech recognition module according to the correct context (for instance eliminating '*Is there any post from TOM?* or *Is there in the bus from my son?*).
4. If several grandsons send an email to the current user, the Vodis server will say : "*You have received 2 emails, one from Max and one from Tom. Which one are you going to choose?*" The graphical interface appears with the two names of Max and Tom and with or without the photos if available.
5. The user pronounces a name or selects it on the graphical interface (existing touch screen, existing gesture etc.).
6. Vodis server activates the CES service.
7. CES interprets: *User request check emails from max.grandson@yahoo.de and prepares a slideshow*"
8. CES sends a request to the Multimedia Knowledge Base in order to guess what is the event the user is interested in. The module gathers all media taken at this event, either published on the web or send to the user's email or shared with the user's social network, and prepare a photo album slide show.
9. CES displays email and slide show requested
10. CES actualizes situation model as premise for step (6): check distance from user to robot
11. CES send request to robot spatial orientation system: distance to user has to be approx. 0,5 meter
12. Final robot behavior: move to the user and present e-mail and photos from user´s grandson
13. User is happy!