EC INFORMATION SOCIETY TECHNOLOGIES PROGRAMME

Cognitive Systems Integrated Project Summary
Area / Call: FP6-2002-IST-2 / IST-2002-2.3.2.4
Proposal Number: 004250

‘CoSy’: Cognitive Systems for Cognitive Assistants

http://www.cas.kth.se/cosy.html

Coordinator: **Henrik I. Christensen,**
http://www.nada.kth.se/~hic
Kungl Tekniska Högskolan (KTH), Sweden

Partners:

Jeremy Wyatt & Aaron Sloman
http://www.cs.bham.ac.uk/~jlw/
http://www.cs.bham.ac.uk/~axs/
School of Computer Science, University of Birmingham – UK

Hans Uszkoreit & Geert-Jan Kruijff
http://www.coli.uni-sb.de/~hansu/
http://www.coli.uni-sb.de/~gj/index.phtml
Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) – DE

Aleš Leonardis
http://lrv.fri.uni-lj.si/~alesl/
University of Ljubljana, Slovenia, SI

Bernhard Nebel & Wolfram Burgard
http://www.informatik.uni-freiburg.de/~nebel/
http://www.informatik.uni-freiburg.de/~burgard/
University of Freiburg - DE

J. Kevin O’Regan
http://nivea.psycho.univ-paris5.fr/
LPE-CNRS, University of Paris V - FR

Bernt Schiele
http://www.vision.ethz.ch/schiele/schiele.html
The Swiss Federal Institute of Technology (CH) (moving to TU Darmstadt, DE)

An additional partner with expertise in psychology will be recruited later.

ABSTRACT

Cognitive Systems for Cognitive Assistants

This document is a summary of a proposal produced in October 2003, inspired by the visionary FP6 objective

> *"To construct physically instantiated ... systems that can perceive, understand ... and interact with their environment, and evolve in order to achieve human-like performance in activities requiring context-(situation and task) specific knowledge"*

We assume that this is far beyond the current state of the art and will remain so for many years. However we have devised a set of intermediate targets based on that vision. Achieving these targets will provide a launch pad for further work towards the long term vision. In particular we aim to advance the *science* of cognitive systems through a multi-disciplinary investigation of *requirements*, *design options* and *trade-offs* for human-like, autonomous, integrated, physical (e.g. robot) systems, including requirements for architectures, for forms of representation, for perceptual mechanisms, for learning, planning, reasoning, motivation, action, and communication. The results of the investigation will provide the basis for a succession of increasingly ambitious working systems to test and demonstrate the ideas. Devising demanding but achievable test scenarios, including scenarios in which a machine not only *performs* some task but shows that it *understands* what it has done, and why, is one of the challenges to be addressed in the project. Preliminary scenarios have been proposed. Further scenarios, designs and implementations will be developed on the basis of (a) their potential contribution to the long term vision, (b) their achievability and (c) the possibility of practical applications, for instance in machines to help house-bound invalids who prefer not to impose too much on fellow humans. Tools will be developed to support this exploration. The work will use an 'open' framework facilitating collaboration with a variety of international projects with related objectives.

The problem

Despite impressive progress in many specific sub-topics in AI and Cognitive Science, the field as a whole moves slowly. Most systems able to perform complex tasks that humans and other animals can perform easily, for instance robot manipulators, or intelligent advisers, have to be very carefully crafted, normally their field of expertise is very narrow, and they are hard to extend. Whatever intelligence they have could be described as 'insect-like', with very little flexibility or self-understanding. Part of the reason for this is that over the last few decades research has become highly fragmented: with many individuals and research teams focusing their efforts on narrowly defined problems, for instance in vision, or learning, or language processing, or problem solving, or mobile robotics.

We propose to try to overcome these limitations by using ideas from several relevant disciplines to investigate an ambitious vision of a highly competent robot, combining many different capabilities in a coherent manner, for instance a non-trivial subset of the capabilities of a typical human child a few years old. This work will pursue two main types of objectives concerned with *theory* and *implementation*, and related *subsidiary* objectives

Theory objectives

We aim to produce a body of theory, at different levels of abstraction, regarding requirements, architectures, forms of representation, kinds of ontologies, types of reasoning, kinds of knowledge, and varieties of mechanisms relevant to embodied, integrated, multi-functional intelligent systems. The results should be useful both for enhancing scientific understanding of naturally occurring intelligent systems (e.g. humans) and for the design of artificial intelligent systems.

We expect such a theory to be built around the core idea of a self-modifying architecture combining many capabilities which develop over time, which are deployed concurrently and which interact with one another asynchrnously. The theory would cover both analysis of *requirements* for such an architecture and also *design options* with their trade-offs. Sub-theories would be concerned with different sorts of components of the architecture and the forms of representation and varieties of knowledge that they can use. Theory construction will build on empirical results in related disciplines (e.g. psychology and linguistics) and achievements of computer science, software engineering and AI (including robotics).

Since different sorts of designs are possible the theory will include an analysis of architectural options and trade-offs as well as design-options and trade-offs concerning components.

Implementation objectives

We expect to produce well-documented implementations of a succession of increasingly sophisticated working systems demonstrating applications of parts of the theory, e.g. in a robot capable of performing a diverse collection of tasks in a variety of challenging scenarios, including various combinations of visual and other forms of perception, learning, reasoning, communication and goal formation. Initially two main kinds of robot will be investigated both of which will learn from and interact with human teachers. One of them will be concerned with finding its way around a complex building, showing others where to go and answering questions about routes and locations. The other robot will be concerned with manipulation of structured objects on a table top. These two require somewhat different physical, perceptual, reasoning, planning and learning capabilities, though both can be combined with natural language capabilities and some social competence. A parallel strand of investigation ('how to build a philosopher') will investigate common requirements for self-understanding, for reasoning, for motivation, and for reflection on what is perceived and learnt.

A common feature of all these tasks is where to draw the nature/nurture boundary, between what is designed in initially and what arises through learning and development. Alternative options will be explored. Another common feature is the requirement to integrate different forms of represenation, needed for different sub-capabilities. Another is the requirement for the robot to have multiple ontologies which are used as appropriate for different activities. We shall avoid dogmatism on all these issues, exploring various alternatives and analysing trade-offs.

Subsidiary activities

The project will also produce a succession of workshops and summer schools, publications, and an 'open' web site containing code, development tools, theoretical papers, various kinds of re-usable libraries, demonstration packages, etc, including contributions from external collaborators, academic and industrial. We expect to have to share development of tools with other projects.

For more information see the Cosy Web site: `http://www.cas.kth.se/cosy.html`