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
Centre for Autonomous Systems
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/
Faculty of Computer & Information Science
University of Ljubljana, Slovenia, SI

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

J. Kevin O'Regan
http://nivea.psycho.univ-paris5.fr/
Laboratory of Experimental Psychology CNRS,
University of Paris V - FR

Bernt Schiele
http://www.vision.ethz.ch/schiele/schiele.html
Dept of Computer Science, TU Darmstadt, DE

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

Contents

Cognitive Systems for Cognitive Assistants

1 Abstract

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 robot 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 (which may not be obvious in advance) and (c) the possibility of practical applications. 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.

2 Objective of project

2.1 The problem

Despite impressive progress in many specific sub-topics in AI and Cognitive Science, work on building integrated cognitive systems 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' insofar as they have capabilities that they do not understand, they do not know why they do things one way rather than another, they cannot explain what they are doing, they cannot improve their performance by taking advice from a human, and they cannot give advice or help to someone else doing similar tasks. 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 very 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 distant 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. The scientific importance of this objective is that such a robot would require generic capabilities providing a platform for many different sorts of subsequent development, since such a child can develop in any human culture and benefit from many forms of education. However, we do not underestimate the profound difficulties of this challenge.

The project will make use of results in the various component disciplines of AI and cognitive science, for instance, new results on perception, learning, reasoning, language processing, memory, plan execution, and studies of motivation and emotion. It should also provide new substantive contributions to those disciplines in the form of new theories and working models, and also new research questions.

The goal of producing a robot with many of the capabilities of a human child cannot be achieved in the time-frame of this project: it is an enormous long term challenge. However, by analysing in great detail the many requirements for moving in that direction, we can derive sets of successively less challenging sub-goals that should provide significant steps towards the distant goal. Some of these sub-goals are achievable in the time-frame of the project.

The project has two main types of objectives concerned with *theory* and *implementation*, and related subsidiary objectives

2.2 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 comprising different sorts of capabilities which develop over time. 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. Key ideas for the architecture will be inspired by biological considerations.

Requirements for perceptual and motor systems that operate concurrently with, and in close coordination with, processes in all the different architectural layers will be analysed, as will varieties of learning (discussed below).

Different varieties of communication and social interaction will be related to the different architectural layers: for instance, (a) dancing, fighting and moving heavy objects require coupled *reactive* systems; (b) linked collaborative actions spanning spatial and temporal gaps, e.g. in building houses and bridges, require *deliberative* capabilities; (c) the ability to empathise, exhort, persuade, may require extensions of self-understanding in a *meta-management* (reflective) system to support other-understanding. (All of these influences can go both ways: e.g. meeting requirements for social developments may enhance individual capabilities.) The theory will also have to account for affective and motivational mechanisms that allow an individual to exist as an autonomous agent instead of always having to be told exactly what to do or how to deal with conflicts and choices.

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.

2.3 Implementation objectives

We expect to produce well-documented implementations of working systems demonstrating applications of parts of the theory, e.g. in a robot capable of performing a collection of diverse tasks in a variety of challenging scenarios, including various combinations of visual and other forms of perception, learning, reasoning, communication and goal formation.

Distinctive features of such a robot will include integration of sub-functions (e.g. vision and other senses can be combined in making sense of a scene, vision can be used to disambiguate a sentence by looking at what the sentence refers to, and learning processes can enhance different kinds of capabilities, including linguistic, visual, reasoning, planning, and motor skills), and also self-understanding. Central to all of this will be an understanding of *affordances*.

Nature vs. Nurture: How much should be programmed into such a robot and how much will have to be learnt by interacting with the environment, including teachers and other agents? Projects aiming to develop intelligent systems on the basis of powerful and general learning mechanisms starting from something close to a "Tabula rasa" risk being defeated by explosive search spaces requiring evolutionary time-scales for success. Biological evolution enables individuals to avoid this problem by providing large amounts of "innate" information in the genomes of all species. In the case of humans this seems to include meta-level information about what kinds of things are good to learn, helping to drive the learning processes as well as specific mechanisms, forms of representation, and architectures to enable them to work.

We shall avoid dogmatism on what needs to be innate, and explore various alternatives for amounts and types of innate knowledge and produce an analysis of the trade-offs.

2.4 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.

3 Scenario-based Research

Three scenarios have been identified for the study of systems integrating many functions within a single architecture. The first one *The Explorer* is concerned with a trainable robot able to learn how to find its way around a building or some terrain. The second scenario *The PlayMate*[1] is concerned with a robot that is able to manipulate 3-D structures, for instance in order to build a copy of a structure already built by someone else. The third scenario *The Philosopher* concerns the ability of a robot to reflect on what it has done, explain what is done and why, answer questions about why it did not do something and about what would have happened if it had done something different, or about what someone else has done wrong. The third scenario will be built on top of the first two: each of them will provide a test-bed for the mechanisms and representations proposed for acquiring and using reflective-understanding of both actions and thought processes.

In all of the scenarios we shall investigate various options for innate knowledge and capabilities and for kinds of learning that can arise out of and build on what is innate.

The first two scenarios make use of very different spatial properties and relationships because they involve different spatial scales and different relationships between percepts, body-parts and the

[1] Referred to as "CopyCat" in earlier drafts.

actions performed, implying very different requirements for understanding the structure of space and the positive and negative affordances relevant to the tasks. Different forms of representation may be useful (a) for thinking about an individual moving around in a (mostly) 2-D space and (b) for thinking about complex objects being simultaneously moved and rotated in a 3-D space by an agent which itself has spatial structure which changes during actions. Large-scale and small-scale spatial actions also have different requirements. Different learning processes are needed because of different time scales and different relationships to perceived and remembered information: the Explorer, unlike the PlayMate involves constantly relating a small region of space to a much larger enclosing region, whereas the PlayMate involves constantly relating visible surfaces to invisible surfaces and perceived spatial relationships to possible future spatial relationships. Different planning formalisms and strategies may be required.

Combining the two in a single scenario integrating the different forms of understanding of space and motion will be a major challenge. Even the uses of language in the two contexts will have interesting differences (e.g., different interpretations for "here" in different contexts: "Fetch a hammer from the store-room and bring it here", *vs* "Put the hammer here where I can reach it"). We also intend to explore the relationship between self-understanding and other-understanding in such contexts.

These scenarios raise many difficult problems whose solution will require interdisciplinary advances. The next section illustrates our approach to some of the problems.

4 Examples of sub-topics

4.1 Architectures

For many years, research in AI and computational cognitive science focused on forms of representation, algorithms to operate on them, and knowledge to be encoded and deployed or derived. In the last decade or two it has become clear that there is also a need to investigate alternative ways of putting pieces together into a complex functioning system, possibly including parts that operate concurrently and asynchronously on different sub-tasks, for instance, perception, action, reasoning and communicating. Unfortunately this has led to a plethora of architectures being proposed, and much ambiguity in the terminology used to describe them. One reason for this is lack of agreement on what the space of possible architectures is like, or on the terminology for describing architectures or on criteria for evaluating and comparing them.

We aim to produce a framework for describing and comparing architectures. A first draft and relatively simple example of such a framework is the *CogAff schema* described in [2], partly inspired by [1] and work by Minsky. The schema classifies components of an architecture in terms of their functional role, using different functional dimensions, including a crude three-way division between perceptual, central and action components, and another three-way division between components concerned with reactive, deliberative or meta-management functions. Superimposing those divisions gives a grid of nine types of components which may or may not be present in an architecture, and which may be connected in various ways to other components. Other ways of distinguishing architectural components will be needed. E.g. different sorts of developmental and learning processes, and also different types of motivational and emotional processes, will be associated with different sorts of components. Another architectural possibility is the inclusion of very fast reactive pattern recognition mechanisms connected to many parts of the architecture making it possible to detect problems that require rapid and global reorganisation of behaviour, e.g. freezing, fleeing, fighting, or pouncing on prey. Such an "alarm" mechanism could account for several types of emotions and is reminiscent of functions of brain-stem and amygdala.

In recent years many researchers have attempted to design robots using only reactive mechanisms, arguing that either features of the environment or emergent interactions between many individuals will produce effects that were thought to require deliberative and other mechanisms. Others have argued that this suffices only for simple organisms and insect-like robots. Instead of engaging in such battles we shall try to understand under which conditions the various types of architectural components are useful.

A difficult challenge will be designing different parts of the architecture so that they can interact unexpectedly while running. A visual system may need to switch from looking ahead for a gap in a fence to looking down at uneven terrain, so as to guide walking actions. Likewise detailed walking actions may have to be modulated or redirected on the basis of high level perceptual processes, e.g. noticing evidence of a slippery surface. Likewise speech may need to be modulated or re-directed on the basis of visual processes that detect puzzlement in the face of the listener or notice something that answers a question before it is fully formulated.

One of our tasks is to explore whether the self-understanding that most AI systems lack can be based on an architectural layer permitting self-observation, classification, evaluation and possibly some control of internal states and processes, especially deliberative processes that are capable of getting stuck in loops, wasting resources by repeating sub-tasks or not noticing opportunities. An important form of learning might include detecting such cases and finding out how to reduce their effects. This is related to notions of "executive function" used in psychology and psychiatry. Empirical research on executive functions in humans and their development may both illuminate and be illuminated by exploratory designs of artificial cognitive systems with similar functions.

4.2 Representations

Recently some researchers have claimed that animals or robots need no representations. Our response is that all organisms use sensory information to determine how to select actions that use internal energy. Biological evolution discovered many variations on that theme, depending on the kind of information acquired, how it is processed, how it is used, when it is used (e.g. long-term storage may be required), how it is transformed, how it is combined with other information, and how it is communicated. In all cases there is some *medium* used for the information, but there are great differences between different media, including whether they are discrete or continuous, one-dimensional or multidimensional, what sorts of structures they can have, and so on. We can avoid disputes about whether some of them are or are not *really* representations by investigating *what kinds* of representations they are, and what their costs and benefits are to the organism.

Our proposal emphasises perception of affordances, namely the ability of an agent not merely to see what *already exists* (objects, handles, surfaces, gaps, holes, etc.) but also to see the *possibilities* for action and the constraints on possible actions, e.g. movements, grasping, folding, joining, separating, lifting, dropping, etc. This leads to novel requirements for perceptual mechanisms. Most work on perception considers how to represent the entities that exist and are perceived, whereas affordances are concerned with what does not not exist but might exist. We need to find ways of perceiving them without generating combinatorial explosions of possibilities. This may require new forms of representation of possibilities and constraints on possibilities.

Part of the research will be on *requirements* for representations and the trade-offs between different forms of representation in different parts of an integrated system. There has already been much investigation of representations suitable for very specific tasks (e.g. extracting structure from motion in order to produce a graphical display of a scene from novel viewpoints), but the task of designing representations for systems with multiple requirements (e.g. supporting verbal descriptions of the scene, or controlling actions, or aiding causal understanding, or allowing

performance to improve with practice) may lead to new, more challenging, requirements.

4.3 Learning

In a complex architecture, there may be different kinds of learning mechanisms in different components. Current theories of learning will need to be substantially extended to explain, for instance, kinds of learning that extend the individual's ontology for perceiving and thinking about the environment, and kinds of learning that develop fluency and speed in motor performance, e.g. because reactive components are *trained* by processes in deliberative components. For the system as a whole we shall investigate different sorts of learning within our planned scenarios, including *tutor-driven* learning where a tutor gives various kinds of tasks, explanations, demonstrations, corrections, etc., *tutor-supervised* learning where the learner (the robot) takes most of the initiative and requests help or advice when difficulties are encountered, and *exploratory learning*, where the robot notices new phenomena and categorizes them using its previously acquired knowledge, using whatever mode of categorization (as a type of object, a type of event, a type of difficulty, a type of solution, etc.) is appropriate.

Requirements for continuous, incremental, open-ended, life-long learning will be analysed. These requirements rule out forms of learning which separate a training phase from a phase in which information is used. Humans, like many other animals, continue extending and refining skills of many kinds for many years. Our robot should be able to do the same. That implies that the early knowledge, both about the environment and about oneself, while useable is incomplete in many ways. This requirement for indefinite learning will probably provide important clues as to the nature of some of aspects of self knowledge. Obviously not everything improves over time: you know exactly how many arms, hands and fingers you have at a relatively early stage, whereas developing ball-catching, stone throwing, berry picking, tool-manipulating and violin-playing skills may go on for a long time thereafter. Such continuous improvements in precision and speed might be produced by feedback-driven partly probabilistic adaptive mechanisms. However some kinds of learning involve development of new large scale 'chunks' that are re-usable, such as the actions appropriate to a particular tool, or playing a particular chord on the piano, or fluently typing a certain syllable, or a whole word, on a keyboard, where each chunked action allows quite a lot of variation in detailed movements according to context. Such re-usable chunks require at least two distinct types of learning (a) whatever has to be learnt in order to perform them, and (b) whatever has to be learnt in order to make plans in advance of performing them.

A major challenge is detecting and removing, or preventing inconsistencies, for instance where learning occurs at different levels in abstraction hierarchies. It may be that in view of the explosive combinatorics the system will have to tolerate some undetected inconsistencies and take remedial action only when contradictions are discovered.

These and other considerations suggest that different forms of learning about the same objects and actions may happen in different parts of the architecture. In particular, it may be useful to have different perceptual and learning processes going on concurrently in a reactive layer, in a deliberative layer and in a meta-management (self-reflective) layer that includes observation of the processes in the other layers. Part of the challenge to be address is how these different processes share the same physical sensors and motors for their different purposes.

All these requirements constrain the sorts of ontologies that develop, the sorts of representations that facilitate learning, and mechanisms for making results of learning usable in different tasks. This should provide new ways of testing and evaluating previous theories and mechanisms.

4.4 Perception-Action Modelling

State-of-the-art approaches to perception, action, and planning, fall into two broad classes: abstract relational representations of the effects of actions as used in classical AI planning and probabilistic models of action effects in continuous spaces as used in robot localisation and mapping. The former are general, and non-task specific, but assume either powerful symbolic perceptual systems, or information provided from elsewhere, and if not carefully designed can lead to explosive search spaces. The latter capture uncertainty in both action and observation, and for some-problems they can converge to solutions without massive search. They are, however, typically tied to geometric representations of space, and to specific types of sensors used and specific uses of perceptual information. One of our tasks is to connect these different types of representation in such a way that updates to one representation can be propagated to other representations. We may find that neither mode of representation as currently used is adequate for some of the tasks, for instance representation of affordances, which involve multiple possibilities for changing relationships, or coping with problems requiring significant extensions of the robot's current ontology – e.g. learning new concepts of physics or chemistry, or learning to think about goals and thoughts of other agents.

4.5 Continuous Planning and Acting in Dynamic Multiagent Environments

Realistic dynamic and partially observable environments pose great difficulties. Other agents' actions as well as naturally occurring events (e.g. sunset) may change the agent's surroundings in ways it cannot foresee, control or even perceive. So plan-execution must be modulated in the light of perceived changes (e.g. stop moving when your path or your line of site is blocked). With increasing dynamics of the world an agent's knowledge will become less accurate, and its plans more likely to require modification during execution — yet not all not all plans can easily be repaired during plan execution. Switching to purely reactive forms of planning is no solution since there are situations in which how best to react cannot be decided without thinking several steps ahead. Constructing conditional plans that work under all possible circumstances is both computationally explosive and may require unrealistic prophetic capabilities.

One solution may be to allow agents to postpone resolution of contingencies and handle them only if they occur. The robot may be able to learn which actions are not worth planning in great detail, and how to use planned and unplanned acquisition of new information during execution to check the applicability of plans, to fill gaps in abstract plans and to help with plan revision. This requires an architecture in which unplanned-for perceptual processes can cause current external and internal behaviours to be interrupted or modulated. This should include the ability to detect new malfunctions in sensors or motors which may require either repair or use of alternative strategies.

4.6 Collaborative planning and acting

Further complications and further opportunities arise when other agents are in the environment. They can produce many surprises. In general it is difficult or impossible to predict everything that other intelligent systems will do. However, friendly others may be willing to give advice, provide useful factual information or collaborate either in forming plans or executing them, or both. All this requires *communication*. However, different groups of agents may have different ways to communicate. Groups of artificial agents can communicate using special-purpose formal languages, while human-robot interactions should allow the human to use more convenient methods.

The project will investigate requirements for various kinds of communication in different sorts of contexts and will analyse trade-offs between different solutions, including trade-offs concerning forms of representation to be used within individual agents and forms of communication

between agents in multi-agent scenarios of different kinds. For instance, requirements for agents collaborating on "Explorer" tasks that require moving between different rooms of a building where contact may be temporarily lost are different from the requirements for "PlayMate" tasks where two or more agents are jointly building some structure where they differ in which parts and relationships are visible and what sorts of actions they are performing at any one time, e.g. picking up, putting down, holding together, holding something out of the way, etc. In the latter, communication may require more subtle inferences about what the other can see or do.

4.7 Further requirements for an active robot

An active robot has to be able to control and make use of its own body, and the relation of its body to the environment. In the PlayMate scenario the agent has to be able to move its arm to a target position, do eye-hand coordination, and do these things irrespective of whether the arm is carrying a load, impeded by some obstacle, or moving in unusual conditions such as injury or mechanical, or sensor dysfunction. How should such an agent represent information about its own body? It is unlikely that animals have full 3-D geometric models of their bodies. One possibility is to use dynamically changing affordances, i.e. information about possibilities for and constraints on, possible actions as a kind of knowledge combining things in the environment and oneself. So different kinds of self-knowledge will be relevant in the Explorer and the PlayMate scenarios.

4.8 Meaning, language and social interaction

In all the scenarios the robot will have to be able to acquire, manipulate, store, combine and use information, about the environment and about itself and other agents. Some information may be expressed only in internal forms, others in external communications and some in both forms. This raises deep questions about how it comes about that internal or external structures can be treated by the robot as having semantic content. This is sometimes referred to in AI circles as the problem of 'symbol-grounding', but is much older in the history of philosophy.

We expect to show that no simple answers are correct, since in order to be able to do anything at all, including being able to perceive and learn, the robot, like a new-born animal, will require some "innate" information which implies that not all information can come from perceiving and acting in the world. However, it is clear that animals do learn about new things through interacting with the world so that the innate mechanisms must allow bootstrapping of new ontologies driven at least in part by interacting with the environment.

It is likely that several different kinds of semantic development will be required, including discovery of new hierarchies of sub-categories through self-organising classifiers, and also high level conceptual extensions through discovery of structural inadequacies in an existing ontology — e.g. the need to explain why two things that appear very similar to the senses behave in very different ways, perhaps because they are similar agents with different beliefs and desires, or because they have different, unobservable, physical structures or internal mechanisms. Another process that can drive ontological extension is discovering bugs and features in the agent's own planning and thinking strategies that are not objects of ordinary perception.

Finally, language-based social processes can drive semantic development, as happens when humans learn school and university subjects using ontologies that extend far beyond what the learner can sense or act on. Sometimes a precursor for this is learning a new more appropriate form of representation, e.g. in learning to use mathematical notations, circuit diagrams, chemical formulae, maps of various kinds, and most recently programming languages. It can even include

acquiring explicit knowledge about the language used for communication, after the knowledge has been acquired implicitly through learning to use the language.

Several aspects of social interaction implicit in our descriptions above, will need to be made explicit as the project progresses. For instance, if the robot is to be able to communicate effectively whether as slave, pupil, collaborator, negotiator, or teacher it will have to acquire an understanding of a number or facts about language users, such as that they have percepts, beliefs, desires, intentions, preferences, principles, etc. that they have a variety of types of knowledge and skill and can differ in their capabilities and also differ over time as a result of learning. It will also have to understand various kinds of dialogue structures and how they can be used (or abused) to achieve various kinds of goals. The fact that fairly abstract dialogue structures (e.g. requesting clarification before answering a question) can coexist with other kinds of processes (e.g. listening to and watching other agents, completing some action, planning the next sentence, noticing that the hearer looks puzzled) helps to determine requirements for the architecture. Very young children cannot do such things at all, let alone do them concurrently. Yet they seem to develop those abilities over time. This is one of many indications that the information-processing architecture itself develops over time. How to achieve that in our robot is one of the hard questions to be investigated.

4.9 Software tools

Success of a project like this will depend on tools that support rapid-prototyping for exploratory construction of complex architectures with many interacting, concurrently active components performing different tasks, possibly at different levels of abstraction. Existing toolkits are mostly either committed to a particular sort of architecture or else aimed at multi-agent systems composed of lots of relatively simple agents perhaps distributed over many machines. More general and open-ended toolkits will be needed, including tools for developing mechanisms that allow self-observation and self-criticism during program execution (meta-management), and tools that support design and implementation of architectures that develop within an individual, something not achieved by current learning mechanisms.

5 Conclusion

We do not claim that we can achieve our long term targets within the scope of this project – or even a large subset. However, unless researchers at least try to assemble all the various pieces of the puzzle that they have been mostly studying in isolation they will fail to see even the trees properly because they don't see the larger wood of which they are part.

The problems are so difficult that many will regard even thinking about them as a waste of time. Our answer is that by carefully analysing the long term goal and working back from it to intermediate goals we can define short-term and intermediate objectives that are attainable and take us in the right direction.

References

[1] NILSSON, N. *Artificial Intelligence: A New Synthesis*. Morgan Kaufmann, San Francisco, 1998.

[2] SLOMAN, A. Beyond shallow models of emotion. *Cognitive Processing: International Quarterly of Cognitive Science 2*, 1 (2001), 177–198.