FP6-004381-MACS

MACS

Multi-sensory Autonomous Cognitive Systems Interacting with Dynamic Environments for Perceiving and Using Affordances

Instrument: Specifically Targeted Research Project (STReP)

Thematic Priority: 2.3.2.4 Cognitive Systems

D2.1.1 Identification of architectural requirements of an affordance-based control

Due date of deliverable: December 31, 2004
Actual submission date of version 1: February 11, 2005
Actual submission date of version 2: October 10, 2005

Start date of project: September 1, 2004 Duration: 36 months

Middle-East Technical University (METU-KOVAN)

Revision: Version 2

Project co-funded by the European Commission within the Sixth Framework Programme (2002–2006)		
Dissemination Level		
PU	Public	**X**
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

EU Project MACS

Deliverable D2.1.1

Identification of architectural requirements of an affordance-based control

Erol Şahin, Ralph Breithaupt, Erich Rome, Patrick Doherty

Number:	**MACS/2/1.1**
WP:	2.1
Status:	draft, version 3
Created at:	January 15, 2005
Revised at:	October 11, 2005

FhG/AIS	Fraunhofer Institut für Autonome Intelligente Systeme, Sankt Augustin, D
JR_DIB	Joanneum Research Graz, A
LiU-IDA	Linköpings Universitet, Linköping, S
METU-KOVAN	Middle East Technical University, Ankara, T
OFAI	Österreichische Studiengesellschaft für Kybernetik, Vienna, A

This research was partly funded by the European Commission's 6th Framework Programme IST Project MACS under contract/grant number FP6-004381. The Commission's support is gratefully acknowledged.

Corresponding author's address:

Prof. Erol Şahin
Middle East Technical University
Dept. of Computer Engineering
Inonu Bulvari
TR-06531 Ankara, Turkey

Fraunhofer Institut Autonome Intelligente Systeme	Fraunhofer Institut für Autonome Intelligente Systeme Schloss Birlinghoven D-53754 Sankt Augustin Germany	Tel.: +49 (0) 2241 14-2683 (Co-ordinator) **Contact:** Dr.-Ing. Erich Rome
JOANNEUM RESEARCH	Joanneum Research Institute of Digital Image Processing Computational Perception (CAPE) Steyrergasse 9 A-8010 Graz Austria	Tel.: +43 (0) 316 876-1769 **Contact:** Dr. Lucas Paletta
LINKÖPINGS UNIVERSITET	Linköpings Universitet Dept. of Computer and Info. Science Linköping 581 83 Sweden	Tel.: +46 13 24 26 28 **Contact:** Prof. Dr. Patrick Doherty
	Middle East Technical University Dept. of Computer Engineering Inonu Bulvari TR-06531 Ankara Turkey	Tel.: +90 312 210 5539 **Contact:** Prof. Dr. Erol Şahin
Ö F A I	Österreichische Studiengesellschaft für Kybernetik (ÖSGK) Freyung 6 A-1010 Vienna Austria	Tel.: +43 1 5336112 0 **Contact:** Prof. Dr. Georg Dorffner

Contents

1 Preamble

This deliverable identifies the requirements of an affordance-based control architecture, extracted from our discussions on the scenario and on the concept of affordances. It is written as a guiding document to compile the discussions related to the architecture and provides *what* the requirements are for an affordance-based control architecture without addressing *how* those requirements should/could be implemented[1].

The initial version of this deliverable was initially submitted as a living document. In it, we had proposed to report its updated version at month 7 and its final version at month 18. However, the second iteration of the document was not released at the proposed date, since there were delays in related workpackages and a sufficient amount of progress to be projected onto this document was not made. This version of the document revises and extends the requirements incorporating progress so far.

2 Introduction

In this deliverable, we propose a list of architectural requirements for an affordance-based robot control. These requirements are derived from the discussions made regarding the perception, representation and learning of affordances as well as the demonstrator scenario. The list of requirements presented in this version of the document revises and extends the requirements presented in the first version of the document taking into account a) the multi-phase scenario described in the appendix of D6.4.1, b) the progress made on perception (reported in D3.1.1 and D3.1.2), learning (reported in D5.3.1) and representation (reported in combined D4.2.1 and D4.3.1) workpackages.

The architectural requirements for an affordance-based control architecture are presented in the next four sections. The first section presents the requirements that are general to the whole architecture and which do not fit in the three following sections on perception, representation and learning.

3 Architectural Requirements

The ideas that we report in this document stems from our discussions on the concept of affordances, and on the demonstrator scenario. As reported in other deliverables (D3.1.1, D3.1.2, combined D4.2.1 and D4.3.1, D5.3.1 and D6.4.1), our literature survey showed that there is a rather broad, partly contradictory meanings that are all associated with the concept of affordances. In some articles, the notion of affordances is used as a mere description means, with no indication of an explicit embedding of affordance-based control. Gibson's own work on affordances does not elaborate on the action part, that is, the usage of affordances. So one important task within the project is to find a working model of affordances in the context of robot control. This model must allow an implementation, including explicit representation of affordance-related information:

- perceptual: features that describe affordances

[1] How these requirements can be implemented in an architecture is the topic of our on-going work in the task of "The development of an affordance-based control architecture". This task is in progress and we already have developed a draft architecture which will be reported in D2.2.2 on month 18.

- self-model: a description of the robot's capabilities (sensor and actuator properties)
- usage/behavioral part: descriptions of action sequences for acting upon affordances
- deliberation: mission plans that make use of affordances, both for normal mission execution and recovery in abnormal situations

From these basic requirements, it is clear that we must extend the original notion of affordances in several new ways, namely by defining a model, by representing affordances explicitly (which contradicts the most disputed concept of direct perception) and by defining ways to make use of affordances for robot control. Our affordance model will be determined by our control architecture, the design of which is driven by the ultimate goal of implementing a complete, working affordance-based robot control that as a proof of concept. We identified some of the essential requirements for embedding the affordance concept in a control architecture.

During our discussions, after the first version of this document, a multi-phase scenario was accepted for the development of the architecture. According to this approach, at each phase a different competence is aimed. Although some of the details on the competences required are described in the appendix of D6.4.1, we have also used the presentations made by partners during the Linköping and Ankara meetings to revise and extend the architectural requirement lists.

We should note that these requirements are presented as a seed that will eventually evolve into an affordance-based control architecture and that they should *not* be considered as the targeted specifications for the architecture to be developed. The specification is a different deliverable.

We have a common understanding that the architecture fall in the hybrid-control paradigm, including both reactive and deliberative control components. The architecture is thought to consist of five components as shown in Figure 1

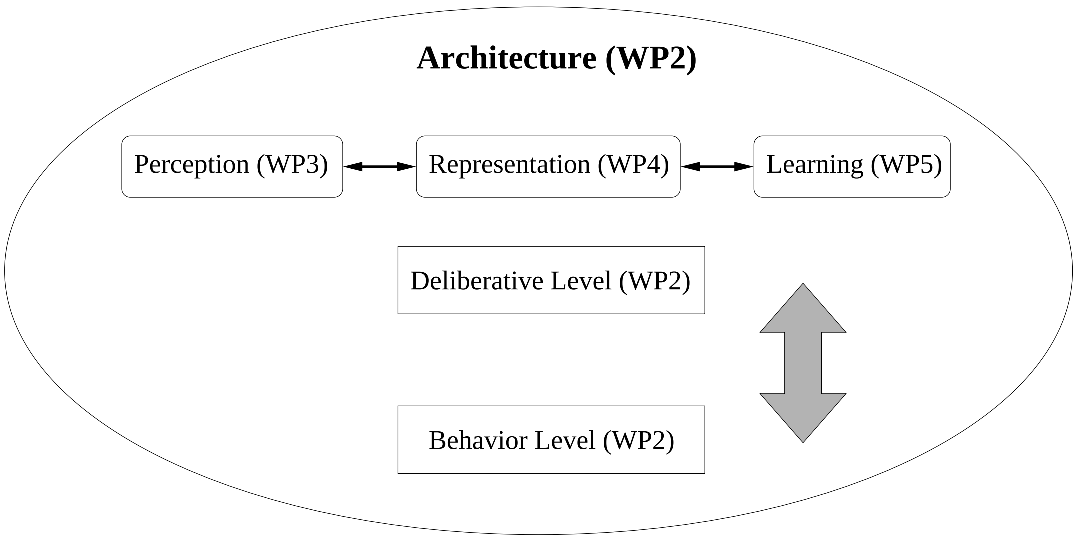

Figure 1: Five components of the affordance-based control architecture.

Each of the upper three components, namely perception, representation and learning have a corresponding workpackage and we discuss the related requirements in the following sections. Regarding the behavior-level component, we would like point out the following requirements:

- A set of behaviors that can act on the objects to achieve a desired affordance is required. During our discussion of other requirements, it will become obvious how these behaviors would be connected to the rest of the architecture. Although we have preliminary notions on how these behaviors can be defined an implemented (one idea would be to use perceptual (standard), affordance (new) and motor (standard) schemas), we do not commit to a specific notion yet.

- The behaviors should be modulated by the sub-symbolic features of the objects they are acting on.

- The behaviors should be coded during the initial phases of the project, however discovery of new behaviors leading to new affordances will remain an elusive target for us.

4 Perception

The perceptual requirements for the architecture can be best described in terms of requirements for other parts of the architecture. Therefore we split these requirements into three categories.

4.1 Perceptual requirements related to representation

Perception should be able to provide a spatio-temporal representation of the world from the perspective of the robot. Perception of the world should be filtered based on the goal of the robot, both to prevent the robot "drowning in affordances" and to activate only the perceptual processing schemas that are relevant for the goal for real-time response.

Perceptual system should utilize both bottom-up and top-down mechanisms to generate a spatio-temporal representation of the world. These systems should interact with each other and should evolve during the project according to the needs of the phase that is being studied.

- During the first phase, perceptual system should not have an "object notion". Instead, features should be hierarchically aggregated and, generalization and categorization of these features should be left to later phases.

- Perception should only be able to operate in structured environments initially and its operating range should be extended in later phases. Assumptions regarding the operating environments can be implemented by top-down perceptual mechanisms. For instance, the bottom-up attention can be used to focus color blobs in camera images, which can then be used to segment the object and extract its properties.

- Perception should have a mechanism to dissociate the interacted "object" from the background. Couple of clarifications are needed here. First, the term "object" merely binds coherent set of features associated with the part of the environment that is being interacted with. In this sense, our objects should really be considered as pseudo-objects without making any claims about the highly-debated object concept.

Second, such a dissociation would implicitly require focus-of-attention (FOA) and segmentation mechanisms whose functioning may be modulated by higher-levels.

- Perception should contain a set of spatio-temporal feature filters/detectors to create a spatio-temporal representation of the object being interacted with. These filters/detectors should only act on the object excluding the background.
- Perception should be able to provide object representations at both symbolic and sub-symbolic levels.
- Perception should have a mechanism to activate only the set of filters that are relevant for the detection of objects that are most likely to support a desired affordance. The advantages of such a mechanism is two-fold. First it would prevent the perception system to run fast on a given computational resource, by disabling the execution of filters that are not relevant for the task at hand. Second, it would minimize the feature representation of an object also minimizing the computations in the consequent steps of processing.
- Perception should be able to perceive the inter-spatial relationships between objects, such as an object being in front of another. Such perception abilities will be essential for tasks such as moving a box occluding an entrance.

4.2 Perceptual requirements related to learning

Perception is required to have the following interactions with the learning mechanism.

- Perception should be able to evaluate whether the testing of an affordance succeeded or not. During learning, the robot will interact with different objects and will have to test whether they possess a certain affordance. In order to learn the associations between objects interacted and the affordances they support, the robot needs a "post-trial" evaluation of its behaviors by observing the world.
- Perception should be able to detect the beginning and the end of the actions/events.
- Perception should be able to detect the consequences of its actions by comparing before and after perceptual representations of the environment and detecting the changes.

4.3 Perceptual requirements related to behaviors

- The behaviors of the robot needs to take into account the specific details of the object being interacted, both at the symbolic and the sub-symbolic level. For instance, the opening of a robot hand should be modulated by the size of an object. Such modulation can be essential to directly couple perception with action as in Gibson's optical flow example.

Deliverable D3.1.1 proposes an outline for the feature vector hierarchy of objects and D3.2.1 presents preliminary results obtained experiments for recognizing affordances from visual cues.

5 Representation

The representation part will integrate perceptual representations of the objects, and the inter-object spatial relationships between the objects being interacted. These representations will then be used as the base for learning and they should also allow deliberation by higher levels.

The following requirements can be put forward:

5.1 Representation requirements related to perception

- Both sub-symbolic and symbolic features of objects, as perceived by the perceptual system, should be represented. It is still unclear how the symbolic representations can be derived from raw sub-symbolic features. Initial discussion on how and at what phase the symbolic representations should come into play have not yet been decided.
- Symbolic representation of objects should enable easy reasoning.
- Object representations should also be able to accommodate "hidden features" which may be known from previous interactions with the object. One example of such a hidden feature would be the weight of a coke can whose content was consumed in the past.
- Abstract representations of objects, minimal feature sets that are essential for an object to be able to support of an affordance, should also be accommodated. These representations would be different object instance representations that are created by the perception.
- Representations of multiple objects should be able to stored and integrated along with the inter-object spatial relationships among them. In a way, this should resemble to the working memory concept which is very limited (3-6 objects) in size and often uses deictic representations for object instances.
- Abstract relationships between the objects should be represented.
- Time series data, recording the perceptual trace of experiments should be represented.

5.2 Representation requirements related to learning

Learning is expected to take place on the representations of the world. Therefore representations should provide support as listed below.

- Representation should support an episodic memory for learning affordances. This can be done by storing the time series of perceptual data. This support is essential since, even the discovery that an object supports a particular affordance spans a long time-span and such episodes need to be stored to enable learning using the "post-trial" perceptual evaluation of the behavior. Such a support also seems to be essential for tackling more complex affordances.

- Representation should provide means for associating object features with the affordances they support.

- Representation should support the learning of abstract relationships of objects and among objects. One example of such an abstract relationship is the concept of "object permanence".

Detailed specification proposals to tackle some of these requirements are proposed in the combined deliverables D4.2.1 and D4.3.1. In this deliverable, some architectural components are also proposed for affordance support.

6 Learning

The architecture should embed various learning mechanisms throughout the architecture. An essential point that should always be kept in mind is that affordance should be addressed in all different mechanisms.

In accordance with the multi-phase approach, learning is expected to be developed in different phases. In its presentations, ÖFAI outlined the preliminary phases of learning. During the first phase of learning, basic affordances, such as grippability, should be learned. During this phase, actions (or behaviors) are assumed to be atomic and that affordances being learned should be related to holistic objects.

During the second phase, complex affordances, affordances that consist of a sequence of atomic actions are proposed to be learned. An example of this is the affordance of stackability. In this phase, the affordances are related to specific object parts. Learning mechanisms developed at this level should be incrementally appended to the mechanisms developed for the first level.

Requirements for the learning mechanisms can be listed as follows:

- For the task of learning, "the relationships between the agent's action possibilities and the environmental entities" are considered to be the affordances to be learnt. The robot should learn these relations by experimenting its repertoire of actions on the entities in the environment observing their positive/negative outcomes.

- Learning should use the time series representations of actions/events extracted by the perceptual system. For this, the learning mechanisms should utilize suitable distance functions for comparing time series representations of different experiments made by the robot. It should also utilize suitable distance functions for comparing combinations of sensor channels.

- The learning mechanisms should to learn to select and configure perceptual filters essential for the task to be tackled without losing essential information.

Below we list the requirements for learning related to perception and representation.

6.1 Learning requirements related to perception

- The architecture should have learning mechanisms to detect and learn coherent (or invariant set) of spatio-temporal features that would support a desired affordance.

- Within the architecture, learning mechanisms should be triggered both by the existence and non-existence of features of an object. Sometimes, the lack of a feature, not the existence of it, can be essential for an affordance.
- Learning mechanisms should be able to discover new affordances through exploration.
- Mechanisms developed for learning should also be extensible for learning new affordances through imitation (which is beyond the goals of this project) and programming.
- The inconsistencies that may occur during learning (e.g. two objects which have the very same representation support conflicting affordances) should be able to trigger the learning (creation of new or fine-tuning of existing) perceptual filters/detectors.
- Learning should accommodate noise during training. The noise can stem from perceptual processing, behavioral imperfection as well as uncertainties within the environment.

6.2 Learning requirements related to representation

- Learning should not only use object/world/context representations as its database, but it should also be able to insert its deductions, such as a new abstract object representation, into the representation part of the architecture.

In D5.3.1, an affordance-learning architecture component, called ABACUS, which took into account the requirements put forward above, is presented. In this document, three phases of learning is also described in detail in line with the scenario presented in D6.4.1.

7 Conclusions

We summarized a set of requirements that are identified as essential for an affordance-based control architecture resulting from the second iteration. The relationship of these requirements with respect to the existing robot control architectures in the literature were reviewed in D2.2.1. Progress has been made on the perception, representation and learning components of the architecture and these are being reported in related deliverables.

Using the architectural requirements provided in this document, a draft affordance-based control architecture is already proposed. This architecture is implemented on the MACSIM simulator and initial results from experiments that are similar to the ones proposed in the appendix of D6.4.1, were conducted with promising results. The development of the architecture is planned to be completed on month 18 and will be reported in D2.2.2.

The final version of this document will be submitted at month 18.