Towards Affordance-based Robot Control

Dagstuhl Seminar 06231
June 5–9, 2006

Erich Rome[1], Joachim Hertzberg[2], Georg Dorffner[3], and Patrick Doherty[4]

[1] Fraunhofer IAIS, Sankt Augustin, Germany; [2] University of Osnabrück, Germany; [3] Austrian Research Institute for Artificial Intelligence, Vienna, Austria; [4] University of Linköping, Sweden

Abstract. This article summarizes the objectives and the program of the Dagstuhl seminar 06231, "Towards Affordance-based Robot Control". It was held from June 5 to June 9, 2006, at the International Conference and Research Center for Computer Science Schloss Dagstuhl near Wadern, Germany.

1 Introduction

Today's mobile robot perception is insufficient for acting goal-directedly in unconstrained, dynamic everyday environments like a home, a factory, or a city. Subject to restrictions in bandwidth, computer power, and computation time, a robot has to react to a wealth of dynamically changing stimuli in such environments, requiring rapid, selective attention to decisive, action-relevant information of high current utility. Robust and general engineering methods for effectively and efficiently coupling perception, action and reasoning are unavailable. Interesting performance, if any, is currently only achieved by sophisticated robot programming exploiting domain features and specialties, which leaves ordinary users no chance of changing how the robot acts.

The latter facts are high barriers for introducing, for example, service robots into human living or work environments. In order to overcome these barriers, additonal R&D efforts are required. The European Commission is undertaking a determined effort to fund related basic, inter-disciplinary research in a line of Strategic Objectives, including the Cognitive Systems calls in their 6th Framework Programme (FP6, [1]). One of the funded Cognitive Systems projects is MACS ("Multi-sensory autonomous cognitive systems interacting with dynamic environments for perceiving and using affordances").

In Cognitive Science, an affordance in the sense of perceptual psychologist J.J. Gibson [2] is a resource or support that the environment offers an agent for action, and that the agent can directly perceive and employ. Only rarely has this concept been used in Robotics and AI, although it offers an original perspective on coupling perception, action and reasoning, differing notably from standard hybrid robot control architectures. Taking it literally as a means or a metaphor for coupling perception and action directly, the potential is obvious that affordances offer for designing new powerful and intuitive robot control architectures.

Perceiving affordances in the environment means perception as filtered through the individual capabilities for physical action and through the current goals or intentions, thereby coupling perception and action deep down in the control architecture and providing an action-oriented interpretation of percepts in real time. Moreover, affordances provide on a high granularity level a basis for agent interaction and for learning or adapting context-dependent, goal-directed action.

The main objective of the MACS project is to explore and exploit the concept of *affordances* for the design and implementation of autonomous mobile robots acting goal-directedly in a dynamic environment. The aim is to develop affordance-based control as a method for robotics. That involves making affordances a first-class concept in a robot control architecture. By interfacing perception and action in terms of affordances, the project aims to provide a new way for reasoning and learning to connect with reactive robot control. The potential of this new methodology will be shown by

Dagstuhl Seminar Proceedings 06231
Towards Affordance-based Robot Control
http://drops.dagstuhl.de/opus/volltexte/2006/725

going beyond navigation-like tasks towards goal-directed autonomous manipulation in the project demonstrators. All over, MACS aims at embedding its technical results into cognitive science.

Gibson's concept of affordances has a strong appeal. It has been used in design [3] and other areas. Reasons for the lack of usage of the concept in the Robotics literature probably include the non-technical way in which affordances are described in the Cognitive Science literature, making it hard to operationalize the concept in the context of a robot control program. In addition, the concept of affordances as a coupling of perception and action of an individual in its environment is not unanimously accepted in the Cognitive Science literature.

During the MACS proposal phase in late 2003, the idea of organizing an interdisciplinary Dagstuhl seminar related to the core MACS topics emerged. The planned purpose of the Seminar was threefold, namely 1) to disseminate the MACS project ideas and concepts into related scientific communities, 2) to receive feedback on and discuss these ideas, and 3) to discuss the usage of affordances in other research areas.

The organizers saw researchers in four broad areas (philosophy and logic, artificial intelligence and computer science, psychology, and economics and game theory) addressing highly related (in some cases, the same) problems, in which work in one area in all likelihood would benefit research in another. Hence for the Dagstuhl seminar, the organizers felt that there would be valuable interactions and contributions that could be anticipated by bringing people together from these areas.

The remainder of this summary will provide a brief description of the affordance-based robot control idea, followed by a description of the seminar contents and results.

2 Affordance-based robot control

An important aspect of Gibson's theory of affordances is that

> "... to perceive an affordance is not to classify an object." [4, p. 134].

Gibson goes on to state that

> "... If you know what can be done with a graspable object, what it can be used for, you can call it whatever you please. ... The theory of affordances rescues us from the philosophical muddle of assuming fixed classes of objects, each defined by its common features and then given a name. ... But this does not mean you cannot learn how to use things and perceive their uses. You do not have to classify and label things in order to perceive what they afford." [4, p. 134].

Thus, objects and affordances are complementary in the sense that one object class may offer a multitude of affordances, and one affordance may be offered by a multitude of object classes.

An example for the latter statement is the following. A beverage can affords to be picked up, to be opened, to be emptied, and to be thrown away, to name just a few. But each of these affordance is also offered by a multitude of other things that humans experience in their everyday environments. Thus, affordances encompass a *function-centered view* on the environment.

How can one exploit this for robot control? First, we can state that a function-centered perception approach will realize a view of the environment that is orthogonal to object-centered perception. Such function-centered perception would potentially allow a robot to find more alternatives for acting in its environment. A robot mission that requires to find—based on appearances only—and use certain objects in the environment will fail if one or more of these objects cannot be found. But often the identity or appearance of an object may not be relevant for completing a task. A task could, for instance, also be completed if the robot finds an alternative object that offers the same functions as the original one.

An affordance-inspired robot control with a robot-specific function-centered perception would allow a robot more flexibility in plan execution and thus increase the likelihood of successfully completing a mission. Thus, it would enhance a robot's abilities to perceive and utilize the potential for action that the environment offers, i.e. enable a robot to make use of affordances. This is the central hypothesis of MACS.

2.1 The Seminar

2.2 Goals and Content of the Seminar

The aim of the seminar was to bring together researchers from Robotics, Informatics and the Cognitive Sciences to exchange their experiences and opinions, and generate new ideas regarding the following questions:

- How could or should a robot control architecture look like that makes use of affordances as first-class items in perceiving the environment?
- How could or should such an architecture make use of affordances for action and reasoning?
- Is there more to affordances than function-oriented perception, action and reasoning?

The answers to these questions are currently widely open. Two points can be stated with certainty, however. First, an affordance-based robot control architecture cannot simply be an extension (an "added layer", so to speak) to existing modern control architectures. The reason is that affordances would spring into existence in low-level perception, would have to pass filters in the control, such as attentional mechanisms, in order not to flood the robot's higher processing levels, and serve in some explicitly represented form of a structured result of perception as a resource for action selection, deliberation, and learning. So if there is such a thing as an affordance-based control architecture, affordances will have to play a role in all of its layers.

Second, the answers to the seminar questions do not depend on whether or not the Cognitive Sciences agree that Gibson is "right" in the sense that affordances exist in biological brains or minds or exist in the interaction between biological individuals and their environment. The point is, if Gibson's description of phenomena of functional coupling between perception and action is correct, then it is of high interest for robot control designers, independent of how it is best understood according to Cognitive Science standards. Therefore, the seminar would profit from either proponents or opponents of the affordance model. The aim here was discussion and exchange, not unanimity.

2.3 Seminar Organisation

The organizers brought together 32 researchers from different scientific communities, including Computer Science, AI, Robotics, Computer Vision, Cognitive Science, and Biology who attended the seminar. Given that the scientific background of the participants was not homogeneous, and that there was only little technical work that directly fit the seminar topic (as remarked above, there are only relatively few examples of using explicitly the concept of affordances), the program (cf. [5]) was composed of:

- Six overview talks (ca. 60') centered around the state of the art rather than the presenter's own work and achievements, serving to inform the heterogeneous audience;
- Thirteen shorter presentations (ca. 30') of mainly young researchers working in related areas in order to both give the participants an impression of the range of work topics that was present in the audience and to feed the discussion groups; and
- smaller work and discussion groups on focused technical and conceptual issues – this served to involve more participants in active exchange than would have been possible in the full group.

As last part of the seminar, the results from the work in the groups have been summed up in a plenary discussion.

2.4 More Scientific Questions, Work Group Topics

The following topics have been handled by the working groups:

"Learning" Should affordances in a robot be programmed or learned? (Can they be programmed in the first place?)
"Representation" What about an affordance needs to be represented in a robot, and how?
"Perception" How and where in the architecture would attention, intention, or other internal states filter affordances that were perceived on a low level?
"Architecture" How would affordance-based control go together with behavior-based and plan-based control? Is it complementary? Redundant? Inconsistent? And how can affordances be used for reasoning and action?

3 Discussion groups

3.1 Perception

The initial guiding question of the discussion group on simulating the perception of affordances was: "How and where in the architecture would attention, intention, or other internal states filter (aspects of) affordances that were perceived on a low level?"

The group work started with a discussion of Gibson's concept of direct perception. Gibson claimed that the environment contains all of the information needed for accurate perception, and that perception is immediate and spontaneous, and therefore, it does not use any "unconscious inference" (in Helmholtz' sense), that is, no (mental) representations or reasoning are required. Gibson's third claim related to direct perception is that perception and action cannot be separated. Perception serves to guide action, this action generates additional sensory information to be picked up by the perceptual system, and this in turn influences the next action or movement. So an animal is continually exploring its environment, detecting invariant relations, and perceptually learning.

The first two claims are the most debated [6] aspects of the concept of direct perception. There is evidence for more complex, indirect perception in humans, which include memory and (object) recognition processes. In order to account for this evidence, Gibson's view must be extended to a perception model that includes an object-binding view. Such a model has first been proposed by Neisser [7].

Given that there are no object recognition processes involved in affordance perception, the question then is what is the nature of the perceived information? For biological beings, affordances are often dependent on the size of their own bodies. Perception then is relative to this measure. It changes with age, and it must be learned by self-experience. Most probably, some "cues" in the environment signal an agent the existence of an affordance. Such cues could be described in terms of features relative to the perceptual capabilities of that agent.

For designing an artificial affordance-inspired agent, the set of features must be determined that its experimentation or operation environment offers, appropriate sensors must be chosen, and relevant feature detectors implemented.

There are literally thousands of affordances that an adult human has learned. However, our everyday experience is that we do not get drowned in the affordances that the environment offers. Instead, we somehow are able to concentrate on those affordances that best satisfy our current needs or goals. When we are hungry, the perception of edible things is dominant, and when we are tired, the perception of things that we can rest upon is dominant. Humans can employ attentional processes to focus on the perception of particular affordances.

Similarly, artificial attentional processes would support an agent's focussing on those affordances in the environment that are suited to implement the agent's current goals. They help constrain the perception on aspects or features of the environment that are relevant for the task at hand.

As a major benefit of simulating affordance perception in a robot control architecture, the group expected a reduction of complexity in sensing. To be more specific, if a robot designer has a clear understanding of the affordances a robot typically encounters in its operational environment and which cue features are most commonly used in the affordance representations generated by the learning methods, then the number of feature detectors may be significantly reduced. It might also turn out that certain sensors are redundant.

3.2 Representation

The discussion group on learning affordances started with the guiding question: Question 1: “What about an affordance needs to be represented in a robot, and how?”

General considerations. The reason of introducing the affordance-inspired robot control architecture is to enable a robot to solve its tasks in a more robust way by enabling it to take advantage of what the environment offers. Affordances for a robot do not necessarily have to be intelligible for humans. Trying to apply the affordance idea directly from a human perspective (e.g. social affordances) can easily lead to handwaving.

A transfer of the affordance idea from biological beings to man-made artefacts (agents) could start with the following adapted definition:

Definition 1 ((Agent) affordance). *An* agent affordance *is a relation between an agent and its environment which affords a capability. The agent/environment relation affords a capability if the agent:*

1. *has the capacity to recognize that it is in such a relation between itself and its environment, and it*
2. *has the ability to act to bring about that capability.*

Two things should be noted here: First, perceiving an affordance does not mean that the agent has to act upon it. But acting upon an affordance is required when an effect shall be achieved. And second, affordances do not fail, but actions can. A certain relation between a robot and an environment is either present or not. Actions fail when the necessary relations are not present during execution. If the action finishes and the necessary relation was present during the execution (it might have been changing in progress of execution), then the effect is achieved.

Minimum requirements for representing aspects of affordances in a robot control architecture would be:

– A simple mechanism for learning affordances in a robotic agent. It is based on time series of sensor data – both raw and processed –, time series of actuator states and of action states.
– In order to bring about a certain effect when performing a certain action, the presence of specific values in data streams is required.

More specific considerations. In order to take advantage of affordances in a robotic architecture we need the following:

– Representational structures necessary to identify whether a robot is in a certain relation with its environment, that is, if it offers support of an action.
  a) Which data (raw/complex) streams are necessary, i.e. represent invariance?
  b) What are [ranges of] values and their temporal relationship?
– Relations of actions and effects —for planning and learning additional information—must be represented:
  (affordance [cue], action, effect)

Architectural consequences. The following architectural elements are necessary in an affordance-inspired robotic architecture:

– An attention mechanism (equivalent to matched filters) for pre-selecting relevant perceptual features or cues.
– A monitoring mechanism checking if before and during executing an action the robot remains in the required relation with the environment.
– An execution mechanism which bases its action decisions on responses of the monitoring mechanism.
– A planning mechanism which takes advantage of the knowledge about the triple: (affordance [cue], action, effect).

3.3 Learning

The initial guiding question of the discussion group on learning affordances was: "Should (aspects of) affordances in a robot be programmed or learned? Can they be programmed in the first place?"

Necessity of learning, learning approaches. As far as the necessity of learning for an affordance-based agent is concerned, the standard argument 'pro' learning that applies to other control approaches applies here, too: Learning is necessary for an affordance based approach, because a programmer cannot foresee every situation, every action outcome, etc.

Learning and adaptivity could expose a principal strength of an affordance-inspired control approach: If the physical abilities of a robot change (malfunction, accident, "growing up" ...) the affordances change and the agent would be able to adapt to these changes.

Usually robots do not evolve but are rather designed. Every design decision (choice of sensors, actuators, behavior system, learning structure etc.) determines and limits the set of affordances that the robot is able to perceive, use and learn. In order to investigate affordance-inspired robot control, the designer has to realize co-evolution of a robot and its "ecological niche"—the experimentation or operational environment—in a nutshell. The robotic agent needs at least the capability to learn from experience. The group agreed that in principle it is possible to learn affordances independent from other beings or agents. However, learning by imitation is a complex research topic on its own.

Affordances and Objects. There are different pathways in the brain to process affordance-related information and to process objects. As Neisser [7] claimed, the affordance processing pathway is evolutionary older. Objects can – to a certain extent – be detected or separated from the background even if the affordances of such objects are not known. Thus objects seem to add benefit to the affordance concept, but do not seem to be a prerequisite to start with. To learn about the existence of objects could be very useful (cf. conclusion) and even necessary when dealing with hidden affordances. Hidden affordances are affordances that cannot be detected by perception alone, but require reasoning and/or action, e.g. probing objects in the environment or memorizing that a box does contain smaller objects that offer certain affordances. Perception of such a box alone does not suffice to derive the affordances. The box must be opened to verify the affordances of contained objects.

3.4 Architecture

The initial guiding question of the discussion group on affordance-inspired architecture was: "How can (aspects of) affordances be used for reasoning, action, and plan-based robot control?" The group started with converting the guiding question into more specific questions:

1. What are the main elements of an affordance-inspired architecture?
2. What are necessary requirements and constraints of those elements?
3. What are dependencies and interfaces between those elements?
   (*this question could not be discussed in the available time*)

What are the main elements of an affordance-inspired architecture? The group agreed on typical elements like a perception component, a behavior system, a learning component and a component for realizing deliberation. As a consequence of the latter, representational mechanisms should be integrated into the architecture. Similar to a purely reactive system—in the sense of the current understanding of insect control mechanisms—the main execution loops through perception and behaviors. In order to have an architecture that is more general and supports a variety of more complex tasks, a close integration of deliberation, representation and learning is needed, too. Furthermore, it emerged from the discussions that a dedicated mechanism would be needed that monitors the presence of affordances as well as the results of the agent's acting upon affordances. This component has been termed a *monitor*.

– The perception component would handle the sensory data and perform perceptual filtering.
– The behavior component consists of different types of behaviors: e.g. reflexes as well as basic skills which are more explicitly affordance related.
– Ideally, the deliberation would have planning and reasoning capabilities. In order to enable reasoning about affordances, a formal definition of the agent's abilities would be required. Deliberation must have mechanisms for selecting those affordances, i.e., their representations, that are suited for solving the (sub-)task at hand.
– A representational framework must encompass at least representations for actions and affordances. It was disputed whether the perceptual part of a representation would refer to objects or rather to features or cues. A model of the agent's environment might be part of the framework. Mechanisms equivalent to episodic and short term memory have been identified as useful for implementing the storage of affordance representations.
– The learning component would realize adaptive control. When the agent adapts to new or changed affordances, the learning component must update the representation accordingly.

What are necessary requirements and constraints of those elements? The required properties of the affordance-inspired architectural components as identified by the work group are summarized below.

Perception. The properties of a perception component are characterized as follows:

– A perception element should have direct connections to the behavior component. *Attention mechanisms* help to detect salient cues bottom-up and to top-down focus on searched cues.
– *Invariance detectors* would be especially suited to support a learning element.
– *Affordance filters* for goal- and context-directed perception for action should be included.
– In order to support proper design of the agent's perceptual abilities, a perception component should be *reconfigurable and modular*, i.e. encompass many different *filters* as well as parameterizable filters are. Technically, such a component would run on multiple loops and time cycles, dependent on the complexity of the filters.

Behaviors. The following requirements for a behavior component of an affordance-inspired architecture have been identified:

– Active perception behaviors must be included.
– Behaviors should be parameterized, their parametrization guided by perception.
– The creation of action chains for execution must be supported.
– A balancing method for reactive and deliberative behaviors should be incorporated.

Deliberation. A deliberation component should include a planner. However, classical strategic planning has been estimated suboptimal for affordance-inspired planning, since it is too inflexible and would probably compensate the potential advantages of affordance-based control.The planning process should be dynamic and adaptable to the context and the situation at hand, especially to the available and perceived affordances, thus allowing to use alternatives. A prerequisite for finding alternatives is “affordance matching”, i.e. selecting one suitable affordance for the task at hand from a repertoire of available, equivalent affordances. For cases where alternatives cannot be found automatically, interactive planning should be allowed. Generally, a plan should be structured in units of (expected) effects resulting from selected actions, based on their linked representation (affordance [cue], action, effect). The planning process need not necessarily be performed online in real time, although this is desirable as a far reaching goal. An initial plan could be generated “off-line”. The representation of goals and of task-dependent context needs to be addressed, but has not been discussed in the group. The deliberation component should also be able to configure the top-down attention mechanism. As a top-level component in a robot architecture, deliberation also includes mechanisms for interfacing to or communicating with a human user or other agents.

Representation. Representations of aspects of robot affordances are grounded in perception. Especially invariants of all sorts should be adequately representable. The question here is how the invariants are practically determined or measured (cf. section on learning). Representations should be re-usable, that is, adaptable to changed affordances, for instance. The direct usability for deliberation, i.e. planning and reasoning, requires a symbolic representation of robot affordances. In order to represent the affordance link between perception and action, representation should have the basic form of (affordance [cue], action, effect).

Learning. Learning requires perception. All knowledge generated or picked-up is based on the agent’s perceptual capabilities. Learning invariants of percepts or of perception-action relations is an essential capability of an affordance-inspired learning component. The component should learn frequently and use “sequences of actions” instead of fixed states. In order to establish the affordance link between perception and action, the component must support supervised or unsupervised learning of action effects. Based on this capability, the robot could even learn to predict the consequences of its actions. Life-long learning is desirable, as well as the ability to recognize failures of actions and to learn thereof.

Execution Monitor. An open question is the amount of responsibilities and competencies of the monitor component. Basically, it has to watch over the execution of a plan and has to decide wether or not a replanning is necessary, and it can select sub-plans or other alternatives based on the current situation and the goal at hand.

4 Conclusion

The concept of affordances has a strong appeal, since it seems to be intuitively understandable and applicable to a variety of areas. Several groups and researchers have been inspired by the concept of affordances. Affordances have been used in design of human-computer interfaces, in the development of new approaches for robot control, and in investigations of human wayfinding strategies in large man-made infra-structures [8].

In all these areas, the major problem for utilization is to find a model that is suitable for the particular usage or implementation of the affordance concept. One major difficulty for finding operational models of the affordance concept is the vast generality of Gibson’s affordance definition which he simply defined for all “animals”. The questions arose whether it is really applicable to beings as different as crickets and humans, and whether is applicable to animals at different levels of individual development.

Given that the creation of a suitable model and implementation of affordance-inspired robot control can be achieved, the question then is: What will change in robot control if one introduces affordance support? Interestingly, the work group on architecture came to similar conclusions as the MACS project. Affordance support has to be introduced on several levels of a robot control architecture, especially in the perception parts and the behavior system. For the perception part, the relationship of “matched filters” and affordances has been identified as an interesting research topic. Reasoning about affordances and using them for goal-oriented action requires some form of symbolic representation. And finally, affordance-related representations should be learned by a robot.

Regarding the benefit of affordance-inspired control, the hypothesis is that it will provide a systematic way to detect agent-specific possibilities and alternatives for action based on function-oriented perception. A working implementation would enable a robot to find more action alternatives than pure appearance-based perception approaches. However, there are many situations where recognition capabilities are required. Neisser [7] proposed an approach that includes both affordance-related perception and object recognition. However, to date this approach has not been realized in robot control either. Thus, as a long term research question, the interaction between affordance perception and object recognition seems to be worthwhile to pursue. Investigating the little explored affordance-inspired perception and control is a prerequisite for a combined system along Neisser's considerations.

5 Acknowledgements

This work 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. The organizers express their gratitude to the Dagstuhl foundation for their support and for enabling this seminar in their exceptional facilities, and to the participants for their contributions and for making the seminar successful and enjoyable.

References

1. IST priority European Commission. 2nd call, cognitive systems, work programme 2003-4 reference. http://cordis.europa.eu/ist/workprogramme/en/2_3_2_4.htm, 2003.
2. James J. Gibson. The theory of affordances. In R. E. Shaw and J. Bransford, editors, *Perceiving, Acting, and Knowing*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1977.
3. Donald A. Norman. *The Psychology of Everyday Things*. Basic Books, New York, NY, 1988.
4. James J. Gibson. *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston, MA, 1979.
5. Dagstuhl Foundation. Towards affordance-based robot control. http://www.dagstuhl.de/06231/, 2006. Dagstuhl Seminar 06231.
6. J. Fodor and Z. Pylyshyn. How direct is visual perception? Some reflections on gibson's 'ecological approach'. *Cognition*, 9:139–196, 1981.
7. U. Neisser. *Cognitive psychology*. Prentice-Hall, Englewood Cliffs, NJ, 1967.
8. M. Raubal. *Agent-based Simulation of Human Wayfinding: A Perceptual Model for Unfamiliar Buildings*. PhD thesis, Institute for Geoinformation, Vienna University of Technology, Vienna, Austria, 2001.