1 Summary – to be removed

Foo

Karlsruhe Institute of Technology

Communicating, Planning, and Executing Object-Action Complexes

Diploma Thesis
of

Nils Adermann

At the Department of Informatics
Institute for Anthropomatics (IFA)
Humanoids and Intelligence Systems Laboratories (HIS)

Advisor: Prof. Dr.-Ing. Tamim Asfour

Duration: June 1st, 2014 – November 30st, 2014

Contents

2 Motivation

The majority of robots in use today solve only one or few specific problems in easily predictable constraint environments. Rigid requirements on predictability may suit industrial robots working in assembly lines but they pose a problem for general purpose service robots. A service robot collaborating with humans in environments made for humans like a household will encounter a much wider variety of objects and activities. A service robot operating in these environments will without doubt have to cope with unexpected situations.

Artificial intelligence has been a research subject for more than 50 years and many results have been found using simulation or simple robots. But only recent advances in hardware and mechanics have made it feasible to build general purpose robots like humanoid robots. The open environments these operate in and inaccuracies of their sensors pose additional problems for artificial intelligence. Decisions need to be made relying on uncertain data, and unexpected situations need to be analyzed and understood to react appropriately.

This thesis intends to demonstrate that a robot agent can adapt and plan its activities in accordance with dialog held with a human agent or another robot agent using the concept of Object-Action Complexes to communicate its perceptions and plans. It will further show that Object-Action Complexes are a suitable mechanism for learning how to apply existing knowledge to new situations using information provided through dialog with another agent.

The robot is fitted with a dialog system to allow a user to provide the robot with information about its environment and its capabilities. The user can instruct the robot to perform complex tasks. The robot uses a symbolic task planning system to arrive at a series of actions to be executed to achieve the goal selected by the user. The task planner is also used to plan further dialog contributing information required to execute steps of the plan in order to achieve the goal.

The robot's knowledge of its own capabilities and objects in its environment is represented in Object-Action Complexes. Object-Action Complexes are symbolic representations of sensorimotor experience grounded in real-world experience of the robot. They are designed to formalize adaptive and predictive behaviors at all levels of a cognitive processing hierarchy Geib et al. (2011). Thus the robot can communicate with other agents about abstract Object-Action Complexes which are still grounded in actual sensorimotor experience. As Object-Action Complexes additionally provide sufficient formalization to use them in high-level symbolic task planning, they allow the robot to execute complex tasks involving communication.

Throughout this thesis we assume that the robot has already gained various sensorimotor experiences recorded in Object-Action Complexes. The goal is now to find suitable mechanisms to generate abstractions of these Object-Action Complexes which aid communication. These abstractions categorize various Object-Action Complexes into more generic groups suitable for translation into spoken language. These categories enable the robot to predict that a particular new Object-Action Complex might be executable if it can place the description of the Object-Action Complex in the same category as existing Object-Action Complexes. As a result the robot can learn new Object-Action Complexes through spoken language which it can still actually execute by transfering sensorimotor experiences from similar Object-Action Complexes. Experimentation then allows the robot to refine these new Object-Action Complexes according to specifics not directly clear from just the category of Object-Action Complexes it is a part of.

To evaluate the proposed mechanism for learning categories of Object-Action Complexes the execution of Object-Action Complexes is combined with dialog in a scenario that involves tightly coupled interaction between two agents. The agents need to cooperate in order to achieve a complex goal together. They can only reach their goal by communicating sensorimotor information including perceived force exerted by the other agent. The symbolic task planner is responsible for deciding which Object-Action Complexes to execute and which information needs to be relayed to the other participating agent in the form of abstract Object-Action Complexes.

Goals:

- Execution of OACs chosen through learned categories of OACs
- Application of OACs to unknown (on semantic level) objects assumed to be applicable because of spoken dialog and OAC categorization
- Acquiring knowledge about physically executable OACs from dialog
  - Example for grounding of spoken communication about plans & OACs in physical embodiment

3 State of the Art

- Formal definition of OACs and previous uses of OACs
  - Relationship between OACs and language
  - Grounding & Executing OACs
  - Learning OACs
- Dialog systems
  - Speech recognition
    - one4all (Waibel)
  - Natural Language Generation
  - Maintaining conversational state
- Planning in AI
  - PDDL
  - PKS
  - Execution of symbolic plans on real robots (PACO-PLUS/Studienarbeit)
- Planning dialog with symbolic planners
  - Combining Dialog Planning with Task Planning
  - Dialog Planning with PKS

Questions: Where do predicates come from? How are predicates and continuous sensor data linked - discretization processes from attributes / perception How can a task specific domain be chosen? How is the level of abstraction chose? Are there even clear hierarchies of OACs? Can one plan on a higher level without considering a lower level? Excluding spatial information? Motion planning -> task planning -> but even humans don't consider what they will do 20 steps ahead when performing an action, unless they have learned from experience to consider that particular other action far in the distance -> how does this relate to working memory & attention?

Memory Models: LTM/Working Memory How to represent knowledge of actions & objects in memory How to use memory in the context of planning How are planning domains related to LTM? How is plan monitoring related to working memory validation?

4 A Cognitive Architecture for Object-Action Complexes

4.1 MemoryX

- Overview
- Working Memory Segments
  - Object Instances
  - Relations
  - Active OAC
- Long Term Memory Segments
  - Object Instance Classes
  - Relation Classes (?)
  - OACs
  - Episodes

4.2 The Central Executive Agent

- Processes
  - Perception
  - Execution / Simulation
  - Task Selection
  - Planning
  - Learning

4.3 Planning

- PDDL / PKS
- Interface

4.4 Evaluation and Discussion

5 Integration of a dialog system with Working Memory

5.1 Evaluation and Discussion

6 Planning with Dialog

6.1 Evaluation and Discussion

- Examples (show that it works at all)
  - ARMAR wipes a table with cloth and learns about sponges
  - ARMAR and human carry a large board together, ARMAR gives instructions (through the door, left right, duck)
- ???? categorization of OACs through generated dialog (Peter) taking place in simulation
- Integrating new version of PKS with ARMAR based on SPOAC
- Working Memory system by Alexey
  - Compare to SPOAC
- Changes to SPOAC to be able to detect new "unknown" objects
- Designing PKS domains for examples
- Extending SPOAC perceptors with all sensor data required for the examples
- SPOAC perceptor for force feedback
- Defining an encoding for information about tightly coupled interaction in OACs
- Integration with speech recognition system for demos
- Creating a new main scenario controller which deals with language input and starts planning only when appropriate

memoryx introduction - alexey reference framework for segmentable memory new segments: oacsegment, activeoacsegment, episode segment

episodic memory implementation

PDDL, pddl parser / AST / PKS transformation / execution - some papers on building parsers/compilers/vms

gui: step by step simulation (like debugger)

simulation allows specifying particular failures by providing override PDDL effect for particular OACs in particular situations

7 Conclusion

Bibliography

C. Geib, J. Piater, and R. Petrick. Object-Action Complexes: Grounded abstractions of sensory-motor processes. *Robotics and*, 2011. URL `http://www.sciencedirect.com/science/article/pii/S0921889011000935`.