Continuation of the Collaborative Research Center 1320

Everyday Activity Science and Engineering – EASE

Universität Bremen

Funding Proposal
2021 – 2025

Proposal for the Continuation of Collaborative Research Center 1320

EASE – Everyday Activity Science and Engineering

for

2021 – 2022 – 2023 – 2024 – 2025

Coordinating university:

Universität Bremen

Speaker:
Prof. Michael Beetz, PhD
Universität Bremen
Am Fallturm 1
28359 Bremen, Germany
phone +49 (0)421 218 64001
email beetz@cs.uni-bremen.de

Office:
Engelhardt, Verena
Universität Bremen
Am Fallturm 1
28359 Bremen, Germany
phone +49 (0)421 218 64049
email ease@uni-bremen.de

Bremen,

Prof. Michael Beetz, PhD (Speaker)

Bremen,

Prof. Dr.-Ing. Bernd Scholz-Reiter (Rector)

CONTENTS

List of Figures

List of Tables

List of symbols and icons

D	Definition (Box)
„“	Citation (Box)
i	Info (Box)

1 General Information

1.1 Key data

1.1.1 Statutory bodies of EASE

Speaker (leading scientist)

Prof. Michael Beetz, PhD
Universität Bremen
Am Fallturm 1
28359 Bremen, Germany
phone +49 (0)421 218 64001
email beetz@cs.uni-bremen.de

Vice speaker

Prof. Dr. Kerstin Schill
Universität Bremen
Enrique-Schmidt-Straße 5
28359 Bremen, Germany
phone +49 (0)421 218 64241
email kschill@cs.uni-bremen.de

1.1.2 Principal investigators

Principal investigators	Year of birth	Doctorate obtained in:	Home institute and location	Project
Prof. Dr., Alin Albu-Schäffer	1968	2002	Institute for Robotics and Mechatronics (DLR)[1], Sensor Based Robotic Systems and Intelligent Assistance Systems (SRA)[2]	R04
Prof. PhD, John Bateman	1957	1986	Applied English Linguistics (AEL)	P01, P02, P05-N
Prof. PhD, Michael Beetz	1961	1996	Artificial Intelligence (AI)	R01, R03, R04, R05, Z, MGK
Prof. Dr., Gordon Cheng	1968	2001	Institute for Cognitive Systems (ICS)[2]	R01
Prof. Dr., Vanessa Didelez	1971	2000	Leibniz Institute for Prevention Research and Epidemiology (BIPS)	H01
Prof. Dr., Rolf Drechsler	1969	1995	Cyber-Physical Systems (DFKI:CPS)[3], Computer Architecture (AGRA)	P04
Prof. Dr., Udo Frese	1972	2004	Multi-Sensor Interactive Systems (MSIS)	R02
Prof. Dr., Bettina von Helversen	1977	2008	Department of Psychology	H04
Dr.-Ing., Vladimir Herdt	1990	2020	Cyber-Physical Systems (DFKI:CPS)[3], Computer Architecture (AGRA)	P04
Prof. Dr., Manfred Herrmann	1959	1988	Neuropsychology and Behavioral Neurobiology (NBN)	H04
Dr. habil., Hagen Langer	1962	1989	Artificial Intelligence (AI)	INF, F
Dr.-Ing., Daniel Leidner	1986	2017	Institute for Robotics and Mechatronics (DLR)[1], Autonomie und Teleoperation	R06-N
Prof. Dr., Carsten Lutz	1971	2002	Theory of Artificial Intelligence (TAI)	P02
Prof. Dr., Rainer Malaka	1965	1996	Digital Media (DM)	P01, P05-N
Prof. Dr., Helge Ritter	1958	1988	Cluster of Excellence Cognitive Interaction Technology (CITEC)[4]	R05
Prof. Dr., Kerstin Schill	1958	1993	Cognitive Neuroinformatics (NI)	H01, H03
Prof. Dr., Tanja Schultz	1964	2000	Cognitive Systems Lab (CSL)	H03, H04
Dr., David Vernon	1958	1985	Artificial Intelligence (AI)	R01
Prof. Dr., Gabriel Zachmann	1967	2000	Computer Graphics and Virtual Reality (CGVR)	R03
Dr., Christoph Zetzsche	1956	2002	Cognitive Neuroinformatics (NI)	H01

Where not explicitly stated otherwise, the institutes are located at the University of Bremen.

[1] German Aerospace Center
[2] Technical University of Munich
[3] German Research Center for Artificial Intelligence
[4] Bielefeld University

1.1.2.1 Participating institutions

Faculty 3 Mathematics and Computer Science, University of Bremen (FB3)

- Artificial Intelligence (Prof. Michael Beetz, PhD)
  Universität Bremen, Am Fallturm 1, 28359 Bremen, Germany.
- Cognitive Neuroinformatics (Prof. Dr. Kerstin Schill)
  Universität Bremen, Enrique-Schmidt-Str. 5, 28359 Bremen, Germany.
- Cognitive Systems Lab (Prof. Dr. Tanja Schultz)
  Universität Bremen, Enrique-Schmidt-Str. 5, 28359 Bremen, Germany.
- Computer Graphics and Virtual Reality (Prof. Dr. Gabriel Zachmann)
  Universität Bremen, Linzer Str. 9A, 28359 Bremen, Germany.
- Computer Architecture (Prof. Dr. Rolf Drechsler)
  Universität Bremen, Bibliothekstr. 1, 28359 Bremen, Germany.
- Digital Media (Prof. Dr. Rainer Malaka)
  Universität Bremen, Bibliothekstr. 1, 28359 Bremen, Germany.
- Multi-Sensor Interactive Systems (Prof. Dr. Udo Frese)
  Universität Bremen, Enrique-Schmidt-Str. 5, 28359 Bremen, Germany.
- Theory of Artificial Intelligence (Prof. Dr. Carsten Lutz)
  Universität Bremen, Enrique-Schmidt-Str. 5, 28359 Bremen, Germany.
- Computer Architecture (Dr.-Ing.Vladimir Herdt)
  Universität Bremen, Bibliothekstr. 1, 28359 Bremen, Germany.

Faculty 10 Linguistics and Literary Studies, University of Bremen (FB10)

- Applied English Linguistics (Prof. John Bateman, PhD)
  Universität Bremen, Bibliothekstr. 1, 28359 Bremen, Germany.

Faculty 11 Human and Health Sciences, University of Bremen (FB11)

- General Psychology (Prof. Dr. Bettina von Helversen)
  Universität Bremen, Hochschulring 18, 28359 Bremen, Germany.
- Neuropsychology and Behavioral Neurobiology (Prof. Dr. Manfred Herrmann)
  Universität Bremen, Hochschulring 18, 28359 Bremen, Germany.

Leibniz Institute for Prevention Research and Epidemiology (BIPS)

- Leibniz Institute for Prevention Research and Epidemiology (Prof. Dr. Vanessa Didelez)
  Universität Bremen, Achterstraße 30, 28359 Bremen, Germany.

Cyber-Physical Systems, German Research Center for Artificial Intelligence (DFKI:CPS)

- Cyber-Physical Systems, DFKI (Prof. Dr. Rolf Drechsler)
  Universität Bremen, Bibliothekstr. 1, 28359 Bremen, Germany.
- Cyber-Physical Systems, DFKI (Dr.-Ing.Vladimir Herdt)
  Universität Bremen, Bibliothekstr. 1, 28359 Bremen, Germany.

Cluster of Excellence Cognitive Interaction Technology, Bielefeld University (CITEC)

- Cluster of Excellence Cognitive Interaction Technology (Prof. Dr. Helge Ritter)
  Universität Bielefeld, Universitätsstraße 25, 33615 Bielefeld, Germany.

Institute for Cognitive Systems, Technical University of Munich (ICS)

- Institute for Cognitive Systems, Technical University of Munich (Prof. Dr. Gordon Cheng)
  Technische Universität München, Arcisstraße 21, 80333 München, Germany.

Institute for Robotics und Mechatronics, German Aerospace Center (DLR)

- Institute for Robotics und Mechatronics, German Aerospace Center (Prof. Dr. Alin Albu-Schäffer)
  German Aerospace Center, Münchner Straße 20, 82234 Weßling, Germany.
- Institute for Robotics und Mechatronics, German Aerospace Center (Dr.-Ing. Daniel Leidner)
  Oberpfaffenhofen, Münchener Str. 20, 82234 Weßling, Germany.

Sensor Based Robotic Systems and Intelligent Assistance Systems, Technical University of Munich (SRA)

- Sensor Based Robotic Systems and Intelligent Assistance Systems, Technical University of Munich (Prof. Dr. Alin Albu-Schäffer)
  Technische Universität München, Boltzmannstraße 3, 85748 Garching, Germany.

1.1.3 Project groups and projects

<table>
<tr><td colspan="3">Research Area H: Descriptive models of human everyday activity
Academic disciplines: Neuroinformatics, Computer Graphics, Human Computation, Psychology, Statistics</td></tr>
<tr><td rowspan="2">H01</td><td colspan="2">Sensory-motor and Causal Human Activity Models for Cognitive Architectures</td></tr>
<tr><td>Schill
Didelez
Zetzsche</td><td>Cognitive Neuroinformatics, Bremen
BIPS, Bremen
Cognitive Neuroinformatics, Bremen</td></tr>
<tr><td rowspan="2">H02-E</td><td colspan="2">Mining and explicating instructions for everyday activities</td></tr>
<tr><td>Malaka
Bateman</td><td>Digital Media, Bremen
Applied English Linguistics, Bremen</td></tr>
<tr><td rowspan="2">H03</td><td colspan="2">Discriminative and Generative Human Activity Models for Cognitive Architectures</td></tr>
<tr><td>Schill
Schultz</td><td>Cognitive Neuroinformatics, Bremen
Cognitive Systems Lab, Bremen</td></tr>
<tr><td rowspan="2">H04-N</td><td colspan="2">Decision Making for Cognitive Architectures – Neuronal Signatures and Behavioral Data</td></tr>
<tr><td>Schultz
von Helversen
Herrmann</td><td>Cognitive Systems Lab, Bremen
Department of Psychology,
Neuropsychology and Behavioral Neurobiology, Bremen</td></tr>
</table>

<table>
<tr><td colspan="3">Research Area P: Principles of information processing for everyday activity
Academic disciplines: Linguistics, Theoretical Artificial Intelligence</td></tr>
<tr><td rowspan="2">P01</td><td colspan="2">Embodied semantics for the language of action and change: Combining analysis, reasoning and simulation</td></tr>
<tr><td>Bateman
Malaka</td><td>Applied English Linguistics, Bremen
Digital Media, Bremen</td></tr>
<tr><td rowspan="2">P02</td><td colspan="2">Ontologies with Abstraction</td></tr>
<tr><td>Bateman
Lutz</td><td>Applied English Linguistics, Bremen
Theory of Artificial Intelligence, Bremen</td></tr>
<tr><td rowspan="2">P03-E</td><td colspan="2">Spatial reasoning in everyday activity</td></tr>
<tr><td>Bhatt
Schultheis</td><td>Human-Centered Cognitive Assistance, City
Artificial Intelligence, Bremen</td></tr>
<tr><td rowspan="2">P04</td><td colspan="2">Validation of plan-guided robot behavior</td></tr>
<tr><td>Drechsler
Herdt</td><td>Cyber-Physical Systems, Bremen
Cyber-Physical Systems, Bremen</td></tr>
<tr><td rowspan="2">P05-N</td><td colspan="2">Principles of Metareasoning for Everyday Activities</td></tr>
<tr><td>Bateman
Malaka</td><td>Applied English Linguistics, Bremen
Digital Media, Bremen</td></tr>
</table>

Research Area R: ***Generative models for mastering everyday activity and their embodiment*** Academic disciplines: AI-based Robotics, Real-time Vision		
R01	CRAM 2.0 — a 2nd generation cognitive robot architecture for accomplishing everyday manipulation tasks	
	Beetz Cheng Vernon	Artificial Intelligence, Bremen TUM, Institute for Cognitive Systems, Munich Artificial Intelligence, Bremen
R02	Multi-cue perception supported by background knowledge	
	Frese	Multi-Sensor Interactive Systems, Bremen
R03	A knowledge representation and reasoning framework for robot prospection in everyday activity	
	Zachmann Beetz	Computer Graphics and Virtual Reality, Bremen Artificial Intelligence, Bremen
R04	Cognition-enabled execution of everyday actions	
	Beetz Albu-Schäffer	Artificial Intelligence, Bremen DLR, Robotik und Mechatronik, Oberpfaffenhofen
R05	Episodic memory for everyday manual activities	
	Ritter Beetz	Bielefeld University, CITEC, Bielefeld Artificial Intelligence, Bremen
R06-N	Fault-Tolerant Manipulation Planning and Failure Handling	
	Leidner	Institute for Robotics and Mechatronics, Oberpfaffenhofen

Area Z: ***Additional subprojects***		
MGK	Integrated research training group	
	Beetz	Artificial Intelligence, Bremen
INF	Information infrastructures	
	Langer	Artificial Intelligence, Bremen
F	Laboratory infrastructure support	
	Langer	Artificial Intelligence, Bremen
Z	Project management and central services	
	Beetz	Artificial Intelligence, Bremen

1.2 EASE research profile

1.2.1 Summary of the research programme

State-of-the-art robot agents can perform everyday manipulation activities such as loading a dishwasher or setting a table. However, they can do so only within the narrow range of conditions for which their control programs have been specifically designed or trained, and they are still far from achieving the human ability to autonomously perform a wide range of everyday tasks reliably in a rich variety of contexts.

The vision behind the collaborative research center EASE is the creation of cognition-enabled robots capable of accomplishing human-scale everyday manipulation tasks in varied domestic environments without explicit task-specific programming and given only general high-level instructions. To this end, EASE has established the research area “Everyday Activity Science and Engineering”: the study of the design, realization, and analysis of information processing models that enable robot agents (and humans) to master manipulation tasks that may appear simple and routine, but that are, in fact, complex and demanding.

EASE takes the perspective that the mastery of everyday activity can be formulated as the computational problem of deciding how robots have to move their bodies in order to accomplish underspecified manipulation tasks and that these decisions should be based on knowledge and reasoning. The unique approach that EASE takes is that we investigate and develop complete robot agents that perform end-to-end context-driven manipulation tasks by leveraging (a) explicitly-represented knowledge, (b) explicit inherently-adaptable generalized action plans, and (c) powerful prospection mechanisms based on machine-understandable inner world models.

The core of our approach lies in designing, building and analyzing generative models for accomplishing everyday household tasks. A generative model provides the basis for a mapping from the desired outcomes of a task to the motion parameter values that are most likely to succeed in generating these outcomes. Such a model can be viewed as a joint distribution of motion parameter values and the corresponding task outcomes. In EASE, the generative model is realized through knowledge representation & reasoning, which is based on the robot’s tightly-coupled symbolic and sub-symbolic knowledge about the tasks it is performing, the objects it is acting on, and the environment in which it is operating. These generative models are used to simulate various task execution candidate strategies before committing to one particular strategy to be performed in the physical world.

The research into generative models of everyday activities is inspired by investigations of the manner in which humans master their everyday manipulation tasks, the results of which provide the computational mechanisms that can then be used to replicate these human abilities in cognitive robots.

The first research phase of EASE focussed on “understanding by building” all the elements necessary for physical robot agents that can accomplish everyday manipulation tasks flexibly and robustly in a broad range of contexts, demonstrating them in a complete integrated system. The second phase aims at developing a next-generation robot cognitive architecture that will provide the unifying representational and processing framework necessary for the cognitive development of robot manipulation skills in everyday activities. This will be based on the rational reconstruction of insights obtained from answering the main research questions of the first funding phase, questions such as“how are specific everyday experiences, encapsulated in episodic memories of the robot’s previous experiences, used to generate flexible, context-sensitive action policies?”

1.2.2 Detailed presentation of the research programme

1.2.2.1 Starting point and key research questions

Robot agents are now capable of **performing everyday manipulation activities** such as loading a dishwasher or setting a table. While these agents successfully accomplish specific instances of these tasks, they only perform them within the narrow range of conditions for which they have been carefully designed. They are still far from achieving the human ability to autonomously perform a wide range of everyday tasks reliably in a wide range of contexts. In other words, they are far from **mastering everyday activities**. The human ability to produce efficient, flexible, and reliable complex, goal-directed behavior for vaguely specified tasks in open environments is still one of the most fundamental mysteries of nature. Our goal in EASE is to solve this mystery.

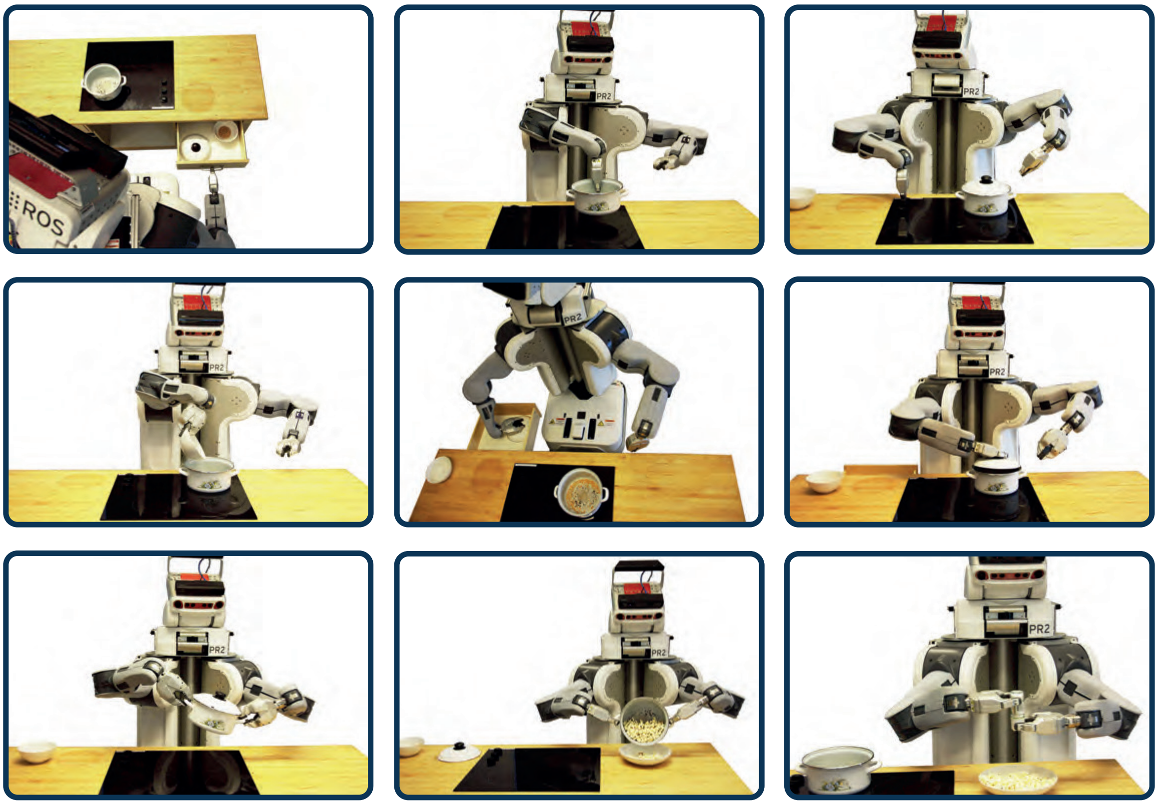


Figure 1.1: A robot agent making popcorn; see also the accompanying video.[5]

Consider, for example, the task of making popcorn. Figure 1.1 shows how this task is accomplished by a robot agent. Making popcorn requires the context-specific sequencing and execution of a variety of behaviors. Many of these behaviors implement different ways of accomplishing the same actions, which depend on the object acted on, the purpose, the physical object state, and the situational context. For example, when lifting the empty pot out of the drawer to place it onto the stove (see Figure 1.1, top-center), the robot grasps the pot by the rim and holds it without having to be concerned about its contents, while it later picks up the pot from the hot plate more carefully, grasping it by the two handles (see Figure 1.1, bottom-left), before rotating the pot, pouring the popcorn into a bowl and salting it.

As none of this is specified in the task, the robot has to infer how to execute actions from what it knows, including factual knowledge, commonsense knowledge, experience, and foresight.

[5] https://www.ease-crc.org/link/video-popcorn-experiment

Everyday Activity

is “a) a complex task that is both common and mundane to the agent performing it; b) one about which an agent has a great deal of knowledge, which comes as a result of the activity being common, and is the primary contributor to its mundane nature; and c) one at which adequate or *satisficing*[6] performance rather than expert or optimal performance is required.”
— Definition by Anderson (1995)

Key research questions for developing a better understanding of everyday activities and how they can be accomplished include: How can an agent, be it a human or an autonomous robot,

... perform the appropriate actions with the appropriate objects in the appropriate ways when given an underspecified task such as “place the pot to the left of the hot plate”?

... perform everyday tasks even in unfamiliar environments with unfamiliar items such as making popcorn in another kitchen or warming up soup instead of making popcorn?

... act competently and efficiently despite the large amount of knowledge and reasoning required to perform the task?

... answer questions about what it is doing, why it is doing it, how it is doing it, and what it expects to happen when it has finished doing it?

Everyday Activity Science and Engineering (EASE) attempts to find answers to these questions by designing, building, and analyzing a **generative model** for accomplishing everyday household tasks, such as unloading the dishwasher or preparing a simple meal. The purpose of investigating the model is not only to enable robot agents to master everyday manipulation tasks but also to better understand how humans perform these tasks.

Generative Model

Actions are decomposed into primitive motions, each of which has a set of parameters that determines the exact nature of the motion. A generative model is a joint distribution of motion parameter values and the associated effects of performing these motions. It provides the basis for a mapping from desired outcomes of an action to the motion parameter values that are most likely to succeed in accomplishing the desired action. In EASE, the generative model is realized through knowledge representation & reasoning, based on a robot’s tightly-coupled symbolic and sub-symbolic knowledge representation, the tasks it is performing, the objects it is acting on, and the environment in which it is operating.

The mastery of everyday tasks is an essential capability of humans. This ability is learned, starting in childhood, through extensive experience, teaching, and demonstration, and often declines with age and brain-related diseases. The ability of people to carry out everyday tasks is tightly connected to their independence. Healthcare professionals have ways to assess people’s (in-)ability to perform activities of daily life (Hartigan, 2007), such as brushing teeth, making coffee (Giovannetti et al., 2007) and preparing meals. Such assessments, like the Naturalistic Action Test (NAT) (Schwartz and Buxbaum, 1997), judge the ability of people “to select actions and objects at the right time and in the right order,

[6] *Satisfice* is a term coined by Simon (1956), which combines the verbs satisfy and suffice. Satisficing is a decision-making strategy or cognitive heuristic that entails searching through the available alternatives until an acceptability threshold is met and explains the behavior of decision makers under circumstances in which an optimal solution cannot be determined.

and to engage in self-monitoring and error correction" (Lawton and Brody, 1969; Schwartz et al., 2002). The NAT specifically addresses everyday tasks that require the use of objects, the sequencing of multiple steps, and the achievement of nested goals (Giovannetti et al., 2002). While investigating the effects of cognitive decline on people's ability to accomplish everyday activities is very important in its own right, EASE seeks instead to make use of the characterization of these activities by health professionals to inform its model of everyday activities for cognitive robots.

D

Everyday Activity Science and Engineering (EASE)

is the study of the design, realization, and analysis of information processing models that enable robot agents (and humans) to master complex human-scale manipulation tasks that are mundane and routine. EASE not only investigates action selection and control but also the methods needed to acquire the knowledge, skills, and competence required for flexible, reliable, and efficient mastery of these activities.

1.2.2.2 Challenges of everyday manipulation tasks

We investigate everyday activities in particular in the context of EASE robot days and years (see Figure 1.2). An EASE robot day consists of the preparation of three simple meals for breakfast, lunch, and dinner. We take heating a frozen pizza or microwaving a curry as examples of a simple meal preparation task. For each meal, in addition to preparing and cooking the food, the robot also has to set the table, clean the table, and load & unload the dishwasher. These tasks are all of the type "put things where they belong" and can be realized through fetching and placing actions.

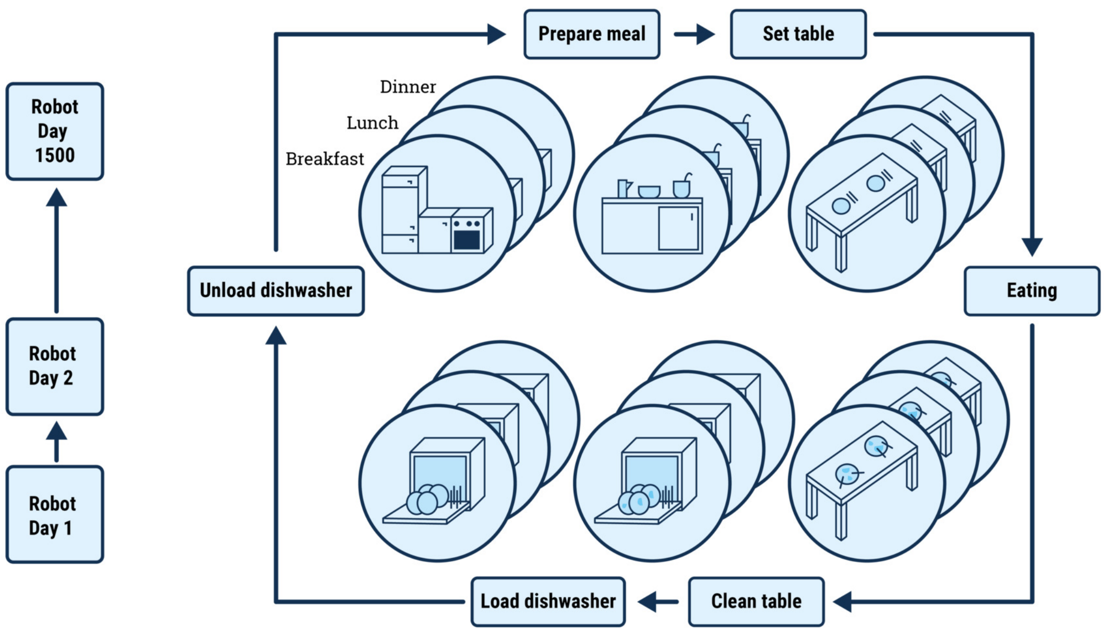


Figure 1.2: The EASE robot housework challenge.

There are a number of challenges in the development of generative models for mastering everyday manipulation tasks:

- **Requests[7] for accomplishing everyday tasks are underdetemined.** A request such as "set the table," "load the dishwasher," and "prepare breakfast" typically does not spell out the intended goal state. Yet the requesting agent has specific expectations about the results of the activity. The agent carrying out the everyday activity needs to acquire the missing knowledge and competently disambiguate information to accomplish the task in the expected manner. To set the table, an agent needs to generate subtasks, such as putting each object that is missing on the table where it is expected to be. The instruction does not spell out which items are missing, where they can be found, or how they have to be arranged on the table. More specifically, the start and the goal state are underdetermined, uncertain, and ambiguous.
- **Accomplishing actions requires flexible, context-specific behavior.** The underdetermined instruction "fetch the next missing object and put it where it belongs" has to generate different behaviors depending on the object, its state, the current location of the object, the scene context, the destination of the object, and the task context. The behavior has to be carefully chosen to match the current contextual conditions, and variations in these conditions require adaptive behavior.
- **Competence in accomplishing everyday manipulation tasks requires decisions based on knowledge, experience, and prediction.** The knowledge required includes common sense, such as knowing that the tableware to be placed on a table should be clean and not in use and that clean tableware is typically stored in cupboards. It also requires intuitive physics knowledge, e.g. that objects should be placed with their center of gravity close to the support surface to avoid them toppling over. Domain knowledge might include knowledge such as the fact that plates are made of porcelain, which is a breakable material. Experience allows the robot agents to improve the robustness and efficiency of their actions by tailoring behavior to specific contexts. Prediction enables the robot to take likely consequences of actions into account, such as predicting that a specific grasp would require the object to be subsequently regrasped in order to place it at the intended location.
- **Accomplishing everyday manipulation tasks requires agents to reason at the motion level.** A large part of the decision-making has to take place at the motion level, reasoning about the way the parameterization of motions alters the physical effects of these motions, and thereby identifying the best way to achieve the intended outcomes and avoid unwanted side effects. For example, in order to pour popcorn onto a plate, an agent has to infer that it has initially to hold the pot horizontally and then tilt it. To do this, it has to grasp the pot with two hands such that the center of mass is in between the hands. It cannot grasp the pot by grasping the rim because the rim is hot and because it would be hard to tilt the pot when grasping the rim. More specifically, the tilting motion is the easiest when grasping the handles and tilting the pot around the axis between the handles. The agent should also remove the lid before pouring because it could fall down when tilting it.
- **Mastery requires agents that "know what they are doing"**. By this we mean that they can answer open questions about what they are doing, why they are doing it, how they are doing it, what they expect to happen when they do it, how they could do it differently, what are the advantages and disadvantages of doing it one way or another, and so on. This applies both to making decisions and reasoning about motions. It is important that agents know what they are doing so that they can assess the cost of not doing it effectively or failing to do it. This understanding also allows them to discover possibilities for improving the way they currently do things.

[7] Requests can originate both from other agents, such as people with whom the robot is working, especially in the next phase of EASE, and from the agent itself as self-generated high level goals.

1.2.2.3 EASE vision and mission

The EASE Vision
is the creation of robot agents that can accomplish everyday manipulation tasks at the competence level of humans.

In the EASE vision, *competence* means that robot agents are able to translate underdetermined action requests into the appropriate behaviors and adapt their behaviors spontaneously to new situations and demands, allowing them to assist humans reliably in a wide variety of settings. Robots will have to act fluently without hesitation, understand what they are doing, communicate the reasons for their choice of behaviors,[8] and improve performance by learning from experience, by reading, by observing, or by playing.

The term *manipulation* indicates that we focus on actions that generate physical motions in order to change objects and substances in goal-directed ways and avoid unwanted side effects. Performing manipulation actions flexibly, robustly, and competently requires intuitive physics and commonsense reasoning in order to translate desired effects into the motion parameterizations that can achieve them.

The term *everyday task* refers to activities that are complex, yet common and routine, such as various household chores. Everyday tasks are made routine through the accumulation of actionable commonsense and naive physics knowledge acquired through experience, reading, observing, dreaming (i.e. mental rehearsal), and playing.

EASE selects everyday activities as its target domain because they allow robots (1) to structure their activities such that they exhibit regularities that can be exploited for better performance, (2) to continually acquire readily actionable commonsense and intuitive physics knowledge, and (3) to improve performance by specializing general actions through the exploitation of task constraints, structure, and regularities.

The EASE Mission
EASE intends to achieve its vision by establishing and developing a new multidisciplinary research field: Everyday Activity Science and Engineering. The goal of this research field is to understand how agents can orchestrate their knowledge and combine it with reasoning processes to accomplish manipulation tasks with a high degree of competence.

EASE facilitates high-impact research, organizes resources, and provides education, technology transfer, and training for programming robots that have autonomous manipulation capabilities for complex and natural tasks in open environments. To achieve this, EASE builds on extensive previous research at the University of Bremen while also integrating a growing community of scientists from international institutions.

The concept of open research is actively encouraged to facilitate collaboration and make faster progress on core research challenges, and to support the democratization of robotics technologies (see Section 1.2.4).

EASE aims to remain at the forefront of the research field, as indicated by research outputs, creating autonomous robot agents based on EASE models of agency, and providing openEASE, an open knowledge and data service for the research community. EASE also provides open-source software components for perception, knowledge processing & acquisition, cognition-enabled robot programming, and semantic object manipulation.

[8] In this context, the term behavior is taken to mean an action strategy, i.e. a way to achieve a desired goal.

1.2.2.4 Long-term research agenda

The EASE research enterprise is organized in three 4-year phases. EASE research takes the cognition-enabled control framework as a starting point (Beetz et al., 2012). Each phase uses and extends the results from the previous phase. For example, the concepts and results from Phase 1 will play a key role in constructing the framework in Phase 2. The long-term research program is depicted in Figure 1.3.

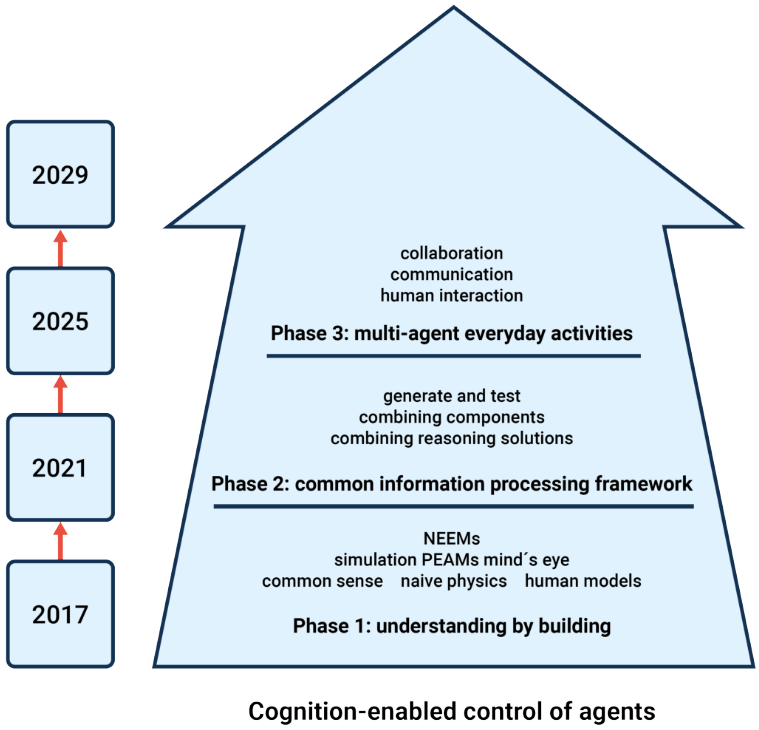


Figure 1.3: Long-term perspective of the EASE research program.

The three 4-year phases have the following research focal points.

Phase 1: Understanding by building (2017-2021) The focus of the first research phase was on understanding everyday activity by building a generative system for selected categories of everyday manipulation tasks, in particular fetch and place tasks.[9] At the same time, we collected data on how humans accomplish the same categories of everyday tasks in real-world and virtual reality scenarios, recording episodes of everyday activities. These episodes provide the basis for obtaining, in a data-driven manner, a better understanding of the structures in human everyday activity. We also built a common knowledge representation across the three different EASE research areas of cognitive robotics, cognitive neuroscience in humans, and knowledge representation & reasoning. This provides a common representation of episodic memories of activities and elements of a formally represented ontology.

The priority in developing the generative model was to ensure that the model can generate all fetch and place actions for setting the table, cleaning the table, loading and unloading the dishwasher, in realistic scenarios.

[9] In EASE we use the term fetch&place in addition to the more common term pick & place. The reason is that we want to emphasize that, in the context of everyday activities, getting an object and putting it somewhere is more complex than in typical pick and place tasks. In everyday activities, it includes finding the right object, sometimes opening containers to access objects. This requires more complex and sophisticated control programs to generate the actions required in everyday contexts.

The main goal of EASE in Phase 1 was to accomplish this on the basis of underdetermined task descriptions, in a robust manner,[10] thereby achieving long-term autonomy.[11] In order to facilitate the rational reconstruction of the generative model in phase 2, i.e. the re-casting of the technical results and scientific insights in a unified and extensible computational framework, we aimed at coding the generative model in a modular and transparent manner. Consequently, the decisions that are derived by the robot agent from the generative model as it adapts its behavior to the current context are explicit and subject to introspection. Furthermore, since the generative model is itself a process in EASE, it can be interpreted and thereby extended and generalized. The goal of this effort is to have, for the first time, a model implemented as a computer program that can be further inspected, studied, and investigated, not only by observers, but by the robot itself. This opens up the real possibility of computational cognitive development through metacognition implemented by re-entrant self-programming. The goal will be taken up in the second focus of Phase 2, as outlined in the next section.

EASE subprojects in this first phase collected pools of episodic memories of common everyday activities that include ones generated from robot and human activity as well as ones generated by simulations, and playing Games with a Purpose about these activities. To achieve maximal synergies between the subprojects, the episodic memories, which we call NEEMs (Narrative-Enabled Episodic Memories), are represented in the knowledge representation and reasoning (KR&R) framework. They are made accessible in the open web-based knowledge service OPENEASE and linked to a common ontology used by all partners. In addition, a software infrastructure was created for the collection, representation, and compression of experiences, and for the extraction generally applicable knowledge from them. Furthermore, initial investigations into discovering structures of everyday activities, called PEAMs (Pragmatic Everyday Activity Manifolds), and their role in mastering everyday activity were studied in various subprojects. As we will see in Sections 1.2.3.3–1.2.3.8, we were successful in achieving the Phase 1 goal of constructing a complete system that uses the generative model to perform a subset of the EASE robot day challenges planned for the end of the project after Phase 3.

Constructing this complete system necessitated the development of a cognitive architecture. However, rather than design it in the usual manner based on established desiderata (Sun, 2004) that focus on achieving a complete unified theory of cognition (Newell, 1990), we adopted a more pragmatic approach that focussed on designing and integrating only the components that were necessary to achieve the main goal of EASE in Phase 1, i.e. successful demonstration of the EASE generative model of robot agency. This approach allowed us to investigate and identify the representations and reasoning & learning processes that were needed to achieve the goal of cognition-enabled robot manipulation but without being constrained by requiring everything fit together perfectly. This cognitive architecture, CRAM, is described in detail in Section 1.2.3.2. It provides the foundation for the realization one of the main the goals of Phase 2: a much-extended version 2.0 of the CRAM cognitive architecture based on the situation model framework (Schneider et al., 2020).

Phase 2: A cognitive architecture for robot agents and everyday manipulation activities (2021-2025) The first research focus of the second funding phase of EASE will be the design and implementation of a second generation cognitive architecture for robot agents that provides and extends the capabilities of the EASE generative model developed in phase 1. A cognitive architecture provides a software framework that orchestrates the core cognitive abilities required for the mastery of everyday activities, including perception, action, anticipation, learning & adaptation, autonomy, attention, action selection & refinement, memory, reasoning, and metacognition. The architecture also specifies the formalisms for knowledge representation and the types of memories that are used to store them, the cognitive processes that act upon the knowledge, and the learning mechanisms that acquire it. The second generation cognitive architecture will build on the initial proof-of-principle CRAM cognitive architecture developed in phase 1 which implements the initial version of the EASE generative model.

[10] Robustness refers to the extent that the goals are successfully achieved over many instances of the task in varying circumstances.

[11] Autonomy refers to the self-reliance of the robot in carrying out tasks, achieved, in part, because of the robot's robustness.

As already noted, the CRAM cognitive architecture is described in Section 1.2.3.2.

A second focus will be the design, implementation, and investigation of a comprehensive knowledge representation & reasoning system for robot agents and everyday activities. This system will represent the cognitive architecture, its operation, and its representational structures in a rigorous way using an ontology such that the representations and the key computational processes become machine-understandable. Doing this will turn the robot agent into one that knows what it is doing, i.e. one that is able to answer open questions about what it does, why, how, what it expects to happen, and what alternatives exist. This will enable the automated generation of learning problems, their solution, and the integration of the solutions into the robot control programs. The representation and reasoning mechanisms will also help to automate large parts of the investigation of the mastery of everyday activities by humans.

The third focus is to continue to expand our understanding of how humans accomplish their everyday activities. This understanding is advanced by (a) taking insights from human mastery of everyday activities and investigating their potential for improving the capabilities of the EASE generative model, and (b) investigating aspects of human activity by testing hypotheses that derive from taking the EASE generative model as the axiomatic basis of a model of cognitive behaviour. In this way, the third focus on human mastery of everyday activities provides both the push and pull to drive the research agenda forward.

Taken together, the three focus points address one of the main goals of the next phase of EASE: the development of techniques that provide the extended functionality of the CRAM 2.0 cognitive architecture. Specifically, CRAM 2.0 will feature significantly enhanced metacognitive abilities, as well as flexible, context-sensitive cognitive behaviours — both fast habitual behaviors and slower but more adaptive deliberative behaviors — and prospective cognitive motion control during action execution. This will be achieved by adopting the situation model framework recently introduced by Schneider et al. (2020). With CRAM 2.0 we aim at a consistent design that reflects what we learned in the first phase and also meets the additional new requirements: flexibility that goes beyond generalized plans, fast execution of habitual tasks (system1 and system 2), cognition in motion execution by the action executive, a KR&R framework for prospection, and the synergistic integration of data-intensive machine learning and, in particular, deep learning.

Phase 3: Multi-agent everyday activities (2025-2029) The focus of the third phase will be the investigation of everyday activity in scenarios where multiple agents interact. These agents can be humans whom the robot agents serve, or humans or other robot agents they cooperate with.

In this phase we will investigate how the structure of everyday activity facilitates the cooperation between different agents. This involves a transition from a focus on goals, intentions, and actions, to shared goals, shared intentions, and joint action, requiring the use of powerful mechanisms such as implicit communication. For example, consider a robot assisting a person icing a cake. While the person pipes the decorative icing on the cake, the robot slowly rotates the cake on the revolving cake turntable at just the right speed for the person doing the icing. Another example would be a scenario in which a robot, after serving the food to its human user, possibly an elderly person, notices that she or he is not eating and so looks around to find out why. The robot notices that the person dropped their fork on the floor and it swiftly brings another one. EASE will investigate learned models of activities and the knowledge abstracted from them by observing human behavior to understand and replicate such competent cooperation and implicit coordination patterns. In this, the theory of mind (Meltzoff, 1995) that is often required to understand the intentions of other agents will be confined to taking a perspective on the shared goals which can be accomplished without having to interpret the psychological or affective disposition of the other agent, possibly communicated through micro-gestures.

1.2.3 Research progress in the first phase

Significant progress has been made on many fronts in Phase 1. The main result is the generative model that was designed, implemented, and investigated in the first funding phase. This model was demonstrated in an internal EASE milestone experiment in September 2020 to enable physical robot agents to set and clean a table given vague task requests. This generative model only requires a carefully designed, **generalized action plan** for fetching and placing objects, which is autonomously contextualized by the model for each individual object transportation task. Thus, the robot autonomously infers the body motion that achieves the respective object transportation task and avoids unwanted side effects (e.g. knocking over a glass when placing a spoon on the table) depending on the type and state of the object to be transported (be it a spoon, bowl, cereal box, milk box, or mug), the original location (be it the drawer, the high drawer, or the table), and the task context (be it setting or cleaning the table, loading the dishwasher, or throwing away items). The body motions generated to perform the actions are varied and complex and, when required, include subactions such as opening and closing containers, as well as coordinated, bimanual manipulation tasks.

We were able to show that the competence of the generative model can be increased by asserting additional generalized domain, commonsense, and intuitive physics knowledge and reasoning, and that substantial parts of such knowledge can be acquired by the robot itself through experience, observation, and taking advice. In addition, the model exhibits impressive introspective capabilities that enable the robot agents employing it to answer questions about what they are doing, why, how, what they expect to happen, and so on. In simulation, we accomplished this scenario in even more variations, such as different kitchen setups with different furniture arrangements, on different robot platforms, and we also applied our generalized fetch and place plan in different domains, specifically retail and assembly domains.

The results of the internal EASE milestone experiment in September 2020 are documented through

- a short video showing the milestone experiment;[12]
- interactive access to the experiment data in openEASE;[13]
- downloadable experiment data in the standardized NEEM format from the NEEM-hub;[14] and
- the code of the main software components of the generative model as open-source software.[15]

The design of the generative model was informed by interpretation and abstraction of high-volume and high-dimensional human everyday activity data. These include 100 multi-modal recordings of varied table setting episodes that contain muscle activity, eye tracking, motion capture, video, audio, concurrent and retroactive think-alouds as well as brain activity from mirror perception. The models and episode recordings are transformed into a machine-understandable representation, semantically rooted in a common ontology of robot and human agency and common representation of activities, which were also developed as a key result of the first funding phase. Highlights of the automatic and semi-automatic acquisition of models of human everyday activity are summarized in a video.[16]

[12] https://www.ease-crc.org/link/video-ease-robot-day

[13] https://data.open-ease.org/QA?neem_id=5fd0f191f3fc822d8e73d715

[14] https://neemgit.informatik.uni-bremen.de/neems/ease-2020-pr2-setting-up-table

[15] Links to the software project websites and the corresponding GitHub repositories are as follows.
Perception: http://robosherlock.org, https://github.com/robosherlock/.
Motion control: http://giskard.de/, https://github.com/semroco/.
Plan executive: http://cram-system.org, https://github.com/cram2/.
Knowledge processing: http://knowrob.org, https://github.com/knowrob.

[16] https://www.ease-crc.org/link/video-h-achievements

In the following, we identify and explain the highlights of our achievements in advancing cognition-enabled robotics. We begin in Section 1.2.3.1 by describing in detail the EASE generative model of robot agency. We follow this in Section 1.2.3.2 with an overview of the CRAM cognitive architecture. Sections 1.2.3.3 – 1.2.3.8 provide details of six key results in Phase 1. We follow this in Section 1.2.3.9 with a discussion of the key scientific insights resulting from Phase 1, concluding in Section 1.2.3.10 with an overall evaluation of the Phase 1 achievements.

The six key results are linked to the structure and operation of the CRAM cognitive architecture. It has five major subsystems: (a) the CRAM plan language (CPL) executive, (b) the KNOWROB2.0 knowledge representation & reasoning system, (c) the ROBOSHERLOCK perception executive, (d) the GISKARD action executive, and (e) the COGITO metacognition subsystem for transformational planning and learning (these subsystems are described in more detain on Section 1.2.3.2).

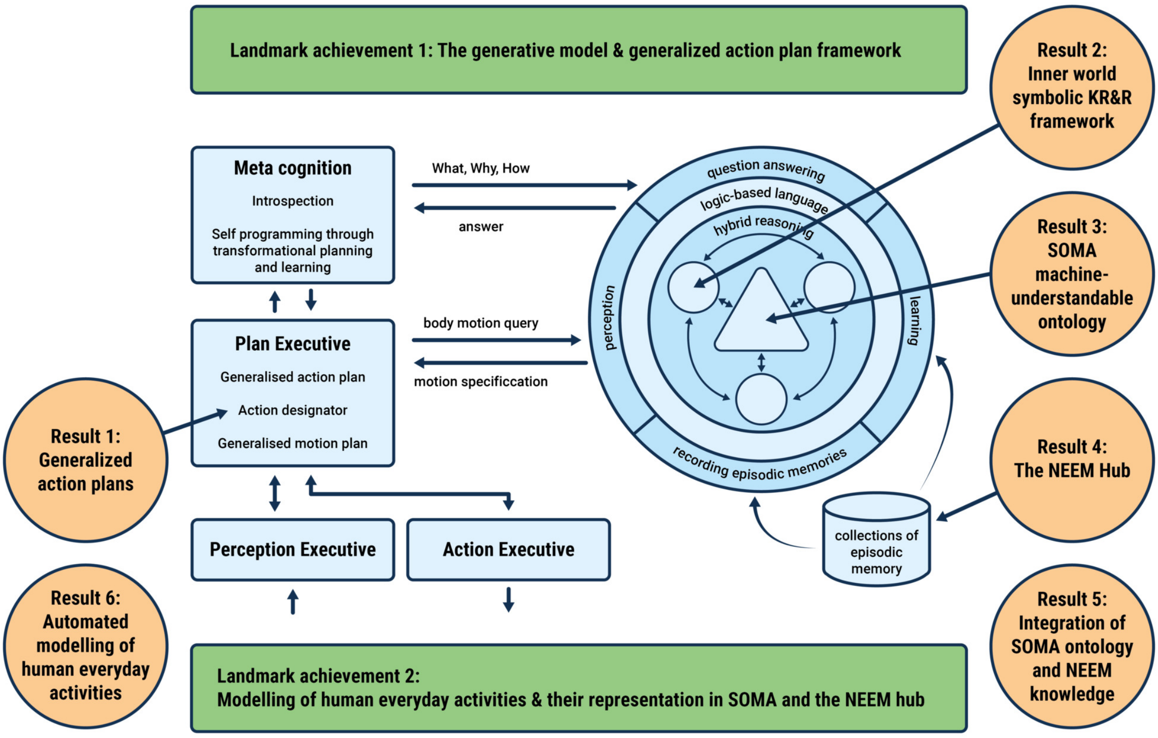


Figure 1.4: Key results and landmark achievements of the first EASE phase and the components of the cognitive architecture to which they apply.

By way of preview, the key results in Phase 1, visualized in Figure 1.4, are as follows.

1. The generalized action plans in the CRAM Plan Language (CPL) Executive.
2. The inner world symbolic knowledge representation & reasoning framework in KNOWROB2.0.
3. The SOMA (Socio-physical Model of Activities) machine-understandable ontology of all EASE knowledge (and data structures & processes) in KNOWROB2.0.
4. The NEEM-HUB: CRAM's generalization of episodic memory — encapsulating sub-symbolic experiential episodic data, motor control procedural data, and descriptive semantic annotation — and the accompanying mechanisms for acquiring them and learning from them in KNOWROB2.0.
5. The integration of the SOMA ontology and the NEEM knowledge in a hybrid symbolic / sub-symbolic framework for observation and interpretation of activities in KNOWROB2.0.
6. Automated modelling of human everyday activities (this result also helped inform the research that led to the achievement of results 3-5).

The conceptualization, articulation, and design of the EASE generative model and generalized action plan framework — realized through the CRAM cognitive architecture — is one of the two landmark achievements of Phase 1 of EASE; see Figure 1.4 (top).

The automated modelling of human everyday activities, i.e. key result 6, and the representation of these activities in the NEEM-HUB and the SOMA ontology is the second landmark achievement of Phase; see Figure 1.4 (bottom).

The second achievement was only made possible by the strong collaboration between EASE research areas H, P, and R. In Phase 2, we will continue this line of research as well as investigating aspects of human activity by testing hypotheses that derive from taking the EASE generative model as an axiomatic basis for a model of cognitive behavior.

Returning to the EASE generative model of robot agency, we divide the explanation in the next section, Section 1.2.3.1, into three parts.

In the first part, Section 1.2.3.1.1, we describe the foundations of the approach and highlight the importance of manipulation in cognitive development in the context of everyday activities. This part also sets out the fundamental goal of creating an activity description programming language that allows manipulation actions to be carried out successfully, simply by saying what action is to be carried out but without having to say how it has to be carried out. We refer to this as an **underdetermined action description**.

D

Underdetermined Action Description

EASE targets the use of high-level abstract specification of the robot actions required to carry out an everyday task. These action specifications are framed in incomplete terms, i.e. they don't provide all knowledge required to complete the task. For example: "fetch the milk and pour it in the bowl." Such an incomplete specification is referred to as an **underdetermined action description**. Think of it as a vaguely-stated instruction, just like the ones people typically give when they are asking someone to do something. The knowledge required to complete the action is acquired when the underdetermined action description is being performed. Later, we will see that an underdetermined action description has two counterparts in the computational mechanisms involved in performing it: a **generalized action plan** and a **high-level action designator**.

In the second part, in Section 1.2.3.1.2, we describe in detail the EASE approach to the goal of creating an activity description programming language. This approach uses reasoning and contextual knowledge to identify the missing information in the underdetermined action description and deploying prospection to identify the robot motions that are most likely to result in the successful execution of the required action. Prospection is achieved using a high-fidelity virtual reality physics engine simulation of a digital twin of the robot and its environment. Collectively, these processes constitute the **generative model**: the mapping between motion parameter values and the outcome of a successful action.

The cornerstone of the EASE approach — and one of the key results of the research in Phase 1 — is the concept of a **generalized action plan**: a computational encapsulation of an underdetermined action description that can be deployed in many everyday contexts.

Generalized Action Plan

This is a plan for a specific category of underdetermined action descriptions. It is a high-level plan and must be expanded into a low-level motion plan before it can be executed by the action executive, Giskard. Like the action description, it is underdetermined: not all the knowledge required to execute the plan is specified. The required knowledge takes the form of the values of the parameters of the motion primitives into which the generalized action plan is expanded. These motion parameter values maximize the likelihood that the associated body motions successfully accomplish the desired action. They are provided by the **generative model**. The expansion and motion parameter value identification is a process referred to as **contextualization** which operates on an element of the generalized action plan known as an **action designator**.

We also describe how, after first selecting the generalized action plan that matches the underdetermined action description, this generalized action plan is then expanded in a process referred to as *contextualization*. In this process, the action-specific arguments required for a **high-level action designator** in the generalized action plan are inserted. The plan is then extended by including the parameters needed to execute the associated motion plan. Finally, using the EASE generative model, an executable motion plan that will successfully achieve the desired action is created by determining the required motion parameter values.

High-level Action Designator

A designator is a placeholder for yet-to-be-determined information. This information is determined at run time based on the current context of the task action. There are four types of designator: action, object, location, and motion (i.e. elementary movement) designators. The identification of the required information is referred to as designator **resolution** and we speak of **resolving a designator**. There is a hierarchy of action designators. Resolving a high-level action designator is accomplished by expanding it into its constituent motion designators. These comprise the elements of the motion plan. The motion parameter values are provided by the **generative model**. Resolving a high-level action designator is also referred to as **contextualization**.

Finally, in the third part, Section 1.2.3.1.3, we summarize the EASE approach to robot agency and we discuss the power of this approach — the generative model and the associated generalized action plans — and its capacity for automatic open-ended extension through metacognition, including self-programming and transformational planning and learning.

Three Complementary Perspectives on Successfully Accomplishing an Action

We speak of performing an **underdetermined action description**, executing (i.e. interpreting) a **generalized action plan**, and resolving the **high-level action designator** in a generalized action plan. The three terms are effectively equivalent. We use them in different contexts: performing an underdetermined action description when focussing on the action, executing a generalized action plan when focussing on the way action is accomplished with the aid of the generative model, and resolving a high-level action designator when focussing on the implementation by resolving the designator into its constituent designator components.

1.2.3.1 The EASE generative model of robot agency

1.2.3.1.1 The EASE perspective on robot agency EASE takes the "brain as a computer program" perspective. As a computer program, the brain is a huge multi-purpose program that — among many other things — can program itself through learning and planning. This allows it to solve complex problems it has never seen before, take vague instructions and do the right thing. The ability to self-program also allows the brain to solve specialized versions of computational problems that are in their general form intractable or even unsolvable, and do so with impressive efficiency. The research question that EASE tackles is: how can we design, realize, and understand a computer program with these properties, enabling it to generate flexible, robust, and context-sensitive high performance behavior? To increase the likelihood of success in this challenging endeavour, EASE focusses on one problem domain: goal-directed object manipulation. The reason for this choice is that goal-directed object manipulation is one of two main driving forces behind the evolution of intelligence (McGinn, 2015).[17] Furthermore, in goal-directed object manipulation the consequences of the computations are more immediately observable.

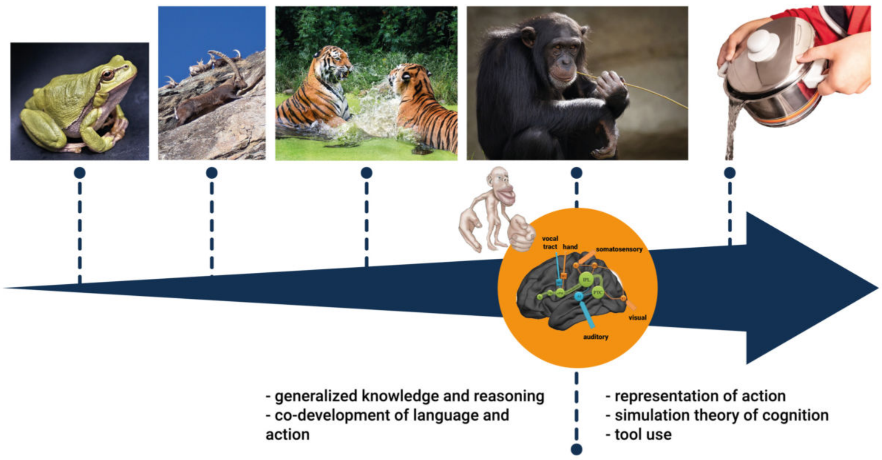


Figure 1.5: The step change in cognitive capability achieved through and necessary for goal-directed object manipulation.

Extension and enhancement of manipulation abilities is a driver of cognitive development
Figure 1.5 sketches the evolution of the brain with respect to the cognitive capabilities needed for physical agency. In this evolution step, one can see radical changes in the structure and the cognitive functions of the brain. The new information processing paths that are formed produce a jump in the capabilities such as tool use, simulation, representation and reasoning, co-development of language and manipulation capabilities. A key impetus for step changes in cognitive capabilities is the deep-seated drive of an agent to extend its repertoire of actions and its predictive control of these actions, especially goal-directed object manipulation capabilities.

New interface layers trigger scientific progress
In computer science, accelerated progress in research and quantum leaps in innovation have often been achieved through one of the most impactful categories of computer science inventions: the introduction of a new interface layer. Domingos et al. (2006) argues: "If we look at other subfields of computer science, we see that in most cases progress has been enabled above all by the creation of an interface layer that separates innovation above and below it, while allowing each to benefit from the other. Below the layer, research improves the foundations (or, more pragmatically, the infrastructure); above it, research improves existing applications and invents new ones." One example of such

[17] Many argue that social interaction is the second one.

interface layers are relational database systems that separate the organization of data for maintaining and querying information from the implementation issues of database systems. Another example is high-level programming languages that let you state computational processes at a conceptual level and which provide compilers that ensure that the programs can be efficiently executed by computer systems. Other examples include GPUs and HDLs (hardware description languages).

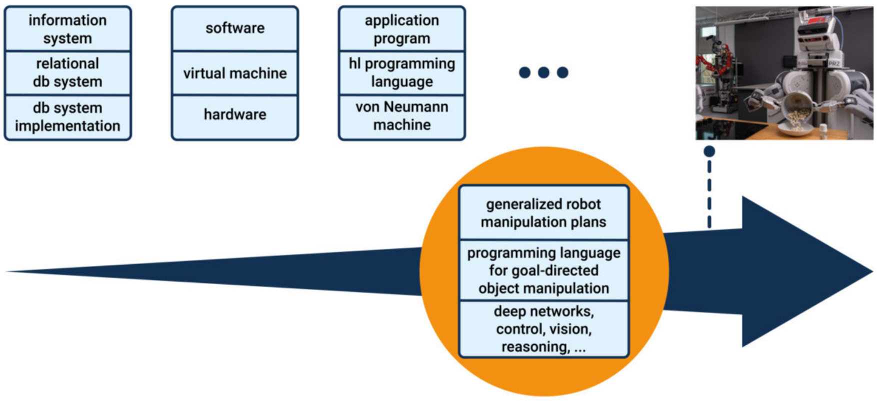


Figure 1.6: The step change in computer science innovation achieved through the introduction of a new interface layer for cognition-enabled object manipulation, highlighting the impact that can be achieved through the introduction of a re-programmable interpreter for a plan programming language for goal-directed object manipulation.

The EASE perspective

Developing the argument from the previous paragraph, EASE takes the following perspective. Suppose the jump in cognitive capability that the brain made in terms of goal-directed object manipulation could be rationally reconstructed as the invention of a software interface layer that facilitates goal-directed object manipulation by creating a re-programmable interpreter for a plan programming language; see Figure 1.12. In this scenario, we want to find answers to the following research questions.

- How can we design a programming language whose main purpose is to generate body motions for underdetermined manipulation tasks, a programming language that can allow a robot to leverage abstract knowledge, understand what it is doing, accomplish novel manipulation tasks, and learn from very few examples? What are the mechanisms for data/knowledge abstraction and procedural abstraction that facilitate such capabilities? What are the basic computational steps?

- How could an execution system for such a language be implemented? Exploiting the enormous progress in computational techniques including deep learning, physics simulation, rendering engines, and big data information systems, what are the most promising ways to implement such an interface layer? Considering what we know about the human brain — the only system that we know being capable of meeting the requirements — what are the computational mechanisms we can incorporate into the execution system?

- Leveraging the programming language for robot agents that can accomplish human-scale everyday manipulation tasks, can we implement robot agents that accomplish EASE robot days or even EASE robot years?[18] Can the robot agents substantially improve over their lifetime? Can the robot agent answer questions about what it is doing, why it is doing it, how it is doing it, and what it expects to happen as a result?

The value of the EASE research enterprise lies in the transformative insights and inventions that the answers to these questions provide, and which we will use for the realization of robot agents that accomplish complex manipulation tasks. These insights and inventions will first of all target the research

[18] Refer again to Section 1.2.2.2 for an explanation of the concept of robot days and robot years.

fields of autonomous robotics and artificial intelligence. In robotics it will mainly be a cognitive architecture of unprecedented capability embodied in complete robot agents. In AI, we anticipate that the breakthroughs will be mainly in knowledge representation and reasoning (KR&R) where we expect the principal contribution to be a hybrid symbolic/subsymbolic knowledge representation that can directly interface with vision-based perception and motion control to enable reasoning capabilities based on internal simulation with perceptuomotor mental imagery. In addition, the KR&R system will include a deeply integrated episodic memory system and mechanisms to generalize commonsense and intuitive physics knowledge. Finally, EASE will generate hypotheses about models that enable agents to accomplish everyday manipulation tasks that will trigger and inform research focussed on achieving a better understanding of human cognitive behavior.

1.2.3.1.2 The EASE approach to robot agency in detail

EASE formulates the mastery of everyday manipulation tasks as the computational problem of deciding how robots have to move their bodies in order to accomplish a natural abstract, i.e. vague and underdetermined, task request and postulates that these decisions should be based on knowledge and reasoning.

The unique approach that EASE takes is that it investigates complete robot agents that perform end-to-end[19] manipulation tasks leveraging **explicitly represented knowledge and behavior prescriptions** and **powerful prospection and memory mechanisms** based on a machine understandable inner world model.

Investigating **end-to-end manipulation tasks** means that the scope of the research activities is the whole computational process that transforms natural abstract task requests into flexible context-sensitive behavior, i.e. body motions, including the physical effects that the body motions may cause. This process includes the mapping of sensor data to motion commands.

Figure 1.7 shows a schematic diagram of the main components and processes involved in realizing the EASE approach to robot agency: the plan executive, the knowledge representation & reasoning executive, the NEEM episodic memory, the perception executive, the action executive, and metacognition. Please refer to this as we walk through each of these elements in the following text. We will refine this schematic diagram in Section 1.2.3.2, specifically by showing how each element forms a part of the CRAM cognitive architecture (see Figure 1.13).

Symbolic knowledge for introspection and self-programming

Leveraging **explicitly represented knowledge** means that the framework is equipped with, maintains, and extends tightly-coupled symbolic and subsymbolic representations of the acting robot, the tasks it is performing, the objects it is acting on, and the environment it is operating in. It uses these representations to make better-informed decisions about the intended course of action. Collectively, this is referred to as the EASE generative model which is realized by the knowledge representation & reasoning executive, NEEMs, and perception executive in Figure 1.7. Specifically, and as we will see in subsequent sections, it is the means by which the motion parameter values corresponding to a vague and underdetermined action description are generated such that the likelihood of success in the task or action is maximized.

Because part of the knowledge is represented symbolically, other knowledge that is implicitly entailed by (i.e. can be inferred from) this explicitly-represented knowledge can be computed effectively through symbol manipulation.

Robot agents that are equipped with knowledge bases will generate a knowledge structure if and only if the knowledge represented by the structure is entailed by the knowledge base. In particular,

[19] Here, the term end-to-end is used to convey the core idea that the manipulation task is accomplished by mapping a vaguely-stated high-level goal (requested either by another agent or self-generated) to the specific low-level motions required to accomplish the goal. It also includes the idea that any constraints that arise at the general level are propagated to the low-level execution in that mapping. Finally, end-to-end also means that the success in accomplishing the task is evaluated in the same space as that in which the goal was formulated, i.e. the perceptual space comprising observations of the world before starting to do the task and observations of the world after completing it, and every instant in between.

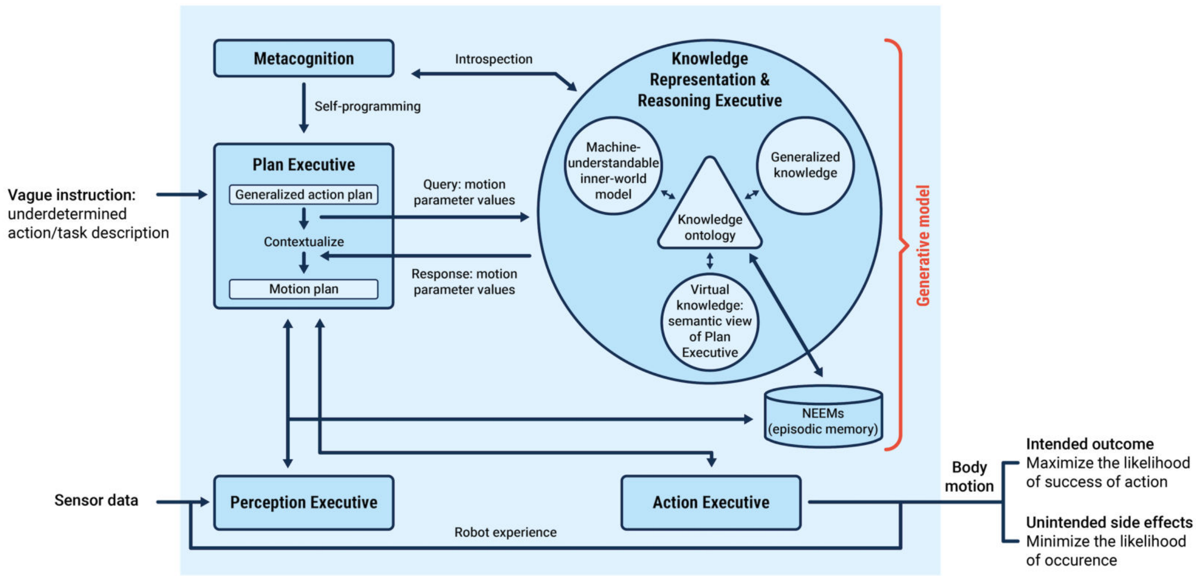


Figure 1.7: Schematic diagram of the main components and processes involved in realizing the EASE approach to robot agency. This schematic is refined further in Figure 1.13 depicting the CRAM cognitive architecture. Note the generative model at the right-hand side.

several knowledge bases can be used to formally assert the relationship between data structures that the robot control system uses and the information that these data structures implement. This enables robot agents to use symbolic inference to automatically answer questions about what the robot is doing, why it is doing it, how it is doing it, and what the consequences might be. This gives the robot agents introspective capabilities, which turn them into “systems that know what they are doing” (Brachman, 2002).

Leveraging **explicitly-represented behavior prescriptions** means that the control program is stated as a plan: where a plan is understood in the sense of a program that can not only be executed but also be reasoned about and manipulated. Plans are represented as symbolic knowledge structures that mirror the structure of the behavior they generate. This plan representation facilitates self-programming and, consequently, development and self-improvement through automatic revision of plans and the generation of new ones.

Symbolic/subsymbolic knowledge for internal simulation and learning generalized knowledge

The robot agents are also equipped with **powerful prospection and memory mechanisms based on a machine-understandable inner-world model**; see Figure 1.7. Again, this inner-world is a hybrid symbolic/subsymbolic knowledge structure with sufficient detail to generate a realistic visualization of the robot’s beliefs about the environment as it simulates the execution of its plans. In addition, the robot records episodic memories of manipulation activities. These episodic memories can subsequently be visualized and analyzed using photorealistic videos and they are coupled to a symbolic knowledge base that represents the activity as a story that tells what happened in the episode. This collection of episodic memories constitutes the experiences of a robot that can be used to learn generalized knowledge.

The combined symbolic/subsymbolic representation is one of the foundations of the EASE approach to robot agency. By coupling a robot’s subsymbolic representation of experience — encapsulated in Narrative Enabled Episodic Memories (NEEMs) — with an ontology-based abstract symbolic representation of that experience (see Figure 1.7), EASE makes it possible to reason about these experiences and to reenact them. Since experience is a temporal event, the NEEM subsymbolic representation and the ontological symbolic representation are periodically synchronized at key points that have perceptuomotor significance, e.g. when the robot hand touches the handle of a pot or when the refrigerator door opens. It also provides great flexibility for learning generalized knowledge through the use of NEEM interpreters and NEEM generators. To see how, note that experiences in EASE can originate from four different sources, each with different sets of sensorimotor modalities. These sources are:

1. a robot's physical actions in the real world;
2. a robot's simulated actions in its inner world;
3. a human's physical actions in the real world;
4. a human's simulated actions in a high-fidelity photorealistic virtual reality environment.

Each of these four sources has an associated generator to capture the experience in the EASE NEEM format. Similarly, each has a NEEM interpreter to map the NEEM to the EASE ontology, thereby creating the symbolic/subsymbolic knowledge structure. The interpreter also provides a powerful way to index the NEEM-encapsulated experiences and select those experiences that match certain semantic requirements, stated symbolically. This allows, for example, NEEM data for certain types of actions or NEEM data with certain temporal characteristics to be extracted, analyzed, and generalized.

Generalized action plans

We come now to the cornerstone of the EASE approach: **generalized action plans**. The robot agent is equipped with a generalized action plan for each action category, which typically corresponds to action verbs such as *fetch*, *place*, *pour*, and *cut*. A generalized action plan specifies an action schema, i.e. a template of how actions of this category can be executed. An action plan is invoked with a request to perform an **underdetermined action description**.[20] Given a request to perform an action description, the plan executive interprets the generalized action plan in order to generate a body motion that is to achieve the goals implied by the request. The plan executive does this by refining the action description into a parameterized motion. This motion plan then generates a body motion that achieves the desired effects. The goal of the refinement process is to identify the values of motion plan's parameters that maximize the likelihood that the associated body motions successfully accomplish the desired action. The refinement is informed through reasoning with knowledge and perceiving the robot's environment.

Interpreting[21] a plan also yields an experience, a representation of what the robot did and what happened when it executed the plan. Thus, there are two results from interpreting a generalized action plan. The first result generates the required body motions to accomplish the action successfully. The second result generates an episodic memory of what the robot experienced as it was accomplishing the action.

[20] As noted earlier, we can speak of *performing an underdetermined action description* or, alternatively, in the specific terms of CRAM, the computational framework in which EASE is implemented, we can speak of *resolving a high-level action designator* in a generalized plan for that particular action. In effect, high-level action designators encapsulate high-level action descriptions and resolving a high-level action designator effectively performs an action description. Designators are placeholders and exploit the current context of the task action when they are resolved at run time. There is a hierarchy of action designators so that the resolution of an action designator can involve the instantiation and resolution of other action designators. The action designators at the lowest level in the hierarchy are referred to as atomic action designators. Ultimately, all action designators are resolved into more primitive motion, location, and object designators. Specifically, atomic action designators are resolved directly into the motion designators that form the elements of the motion plan. Designator resolution is accomplished either by querying knowledge embedded in the plan, by querying knowledge in the KnowRob2.0 knowledge base, or by accessing sensorimotor data through the perception executive. Resolving a motion designator results in motion of the robot body. For the remainder of this proposal, we will speak of performing underdetermined action descriptions and interpreting the associated generalized action plans for that action category, while being aware that we could also speak about resolving high-level action designators.

[21] A plan written in the CRAM plan programming language is executed by an interpreter of this language. Therefore, we also refer to the process of plan execution as *plan interpretation*.

In the next section, we will provide a detailed example of the way in which an underdetermined action description is refined into the parameterized motion plan, specifically by showing how the associated generalized action plan for that category of action is interpreted. To set the scene for this example, we first explain the concept of action description refinement and generalized action plan interpretation in more detail.

The action associated with an underdetermined action description is performed by first selecting the generalized action plan for the action category corresponding to the action description. This is followed by a process we call **contextualization**. This process has three steps, as follows.

The Three Steps of Contextualization

1. *Instantiate* the selected generalized action plan by inserting the arguments required for the specific action to be performed. One of these arguments is the action type, e.g. *fetch, place*, *fetch&place*, *pick up*, *pour*, or *cut*. Others include the type of the object to be manipulated or the destination location. These arguments are typically designators of some kind, e.g. an action, object, or location designator. This result is referred to as the contextualization plan.
2. *Extend* the instantiated generalized action plan by adding the parameters needed to execute the motion plan, e.g. which arm to use, what grasp pose to use.[22]
3. Create a *query* for the values of the motion parameters that will produce robot body motions to achieve the goal of the underdetermined action description (and, equivalently, the associated instantiated and extended generalized action plan).

The third query step in this contextualization process lies at the heart of the EASE working hypothesis: for every action category we can specify a motion plan schema with a small number of motion parameters that is sufficient to achieve the desired outcome and avoid unwanted side effects, and that we can do so for a large variety of objects and tasks by leveraging the constraints imposed by the current context. This query is answered by the knowledge representation & reasoning (KR&R) executive in its role as the implementation of the generative model. The answer identifies the motion parameter values that produce robot body motions that are most likely to succeed in accomplishing the desired action. Since there are two sets of variables in play here — (a) the set of motion plan parameters generated by a given action plan category and its associated generalized action plan, and (b) the set of physical effects (and, in particular, the robot's observations of the physical effects) — an underdetermined action description is best understood as a request to sample a **joint probability distribution of motions**[23] **and the physical effects these motions cause.** The query effectively samples this joint probability distribution to identify the motion parameter values that are most likely to succeed in accomplishing the desired outcome for the underdetermined action description.

This contextualization is accomplished by **reasoning**, exploiting the constraints of contextual knowledge & current perceptual information, and **prospection**, using the robot's inner world to simulate plan execution.[24] Once the query has been answered by the KR&R executive and the parameter values have been determined, the motion plan is executed adaptively by the action executive.

Before proceeding with the detailed example of how an underdetermined action description is refined into a parameterized motion plan (by three-step contextualization of the generalized action plan associated with that category of action), we will first draw out the power of the conceptual view of an

[22]These motion parameters are identified by resolving the high-level action designator into the motion plan's constituent motion designators, by way of the action designator hierarchy.

[23]The motions are generated by the associated generalized action plan.

[24]We remarked in a previous footnote that we can speak of *performing an underdetermined action description* or, alternatively, of *resolving a high-level action designator* in a generalized action plan. We now remark that the process of refining an underdetermined action description, i.e. contextualizing a generalized action plan, is equivalent to the process of resolving the high-level action designator.

underdetermined action description as a joint probability distribution of (a) motions that the generalized action plan generates and (b) the physical effects these motions cause.

The unique EASE approach leverages explicitly-represented knowledge & reasoning and inner-world simulation-based prospection in order to sample this joint probability distribution. It does not use an explicit graphical model or Bayesian network to infer the motion parameter values but it does not preclude it either. However, the joint probability distribution allows us to conceptualize what we seek in the knowledge-based contextualization process: we seek to maximize the utility of the selected motion parameter values, i.e. maximize the likelihood of the action being successful. This corresponds to a probability distribution that has low entropy and is highly informative: in other words, one that exhibits sharp peaks across the probability distribution, indicating motion parameter values that have a high probability of achieving the desired outcome.

It is the exploitation of knowledge that shapes the probability distribution in this manner and decreases the entropy of the distribution, tuning the distribution to the context based on knowledge of the environment in which the activity is being conducted, knowledge provided by perception and, significantly, knowledge provided by prospection through inner-world high-fidelity virtual reality and physics engine simulation. Thus, when querying the KnowRob knowledge base for the motion parameter values that are most likely to succeed in accomplishing the desired action outcome, reasoning, perception, and prospection yield a probability distribution with sharp peaks and the parameter values corresponding to the peak with the highest probability are selected as the response to the query. By analogy with Bayesian reasoning, EASE uses evidence supplied by prior knowledge, reasoning, perception, and prospection to select the motion plan parameter values that are maximally-likely to result in a successful action.

A generalized action plan example

A sketch of the generalized action plan for fetch&place is depicted in Figure 1.8. It is an action schema that provides a template for how to transport any object to any destination. The first step in this schema states that the object is to be picked up in an appropriate manner, while the robot agent is to stand at the appropriate location. What is appropriate depends on the object, the robot capabilities, the surroundings, and the task context. The plan language lets the programmer (or the eventual human agent collaborating with the robot) state underdetermined action descriptions, which have to be contextualized by the robot agent at execution time. For example, in the underdetermined description of the *fetching* action, it is not specified with which arm to grasp the object, with what kind of a grasp, with which force to squeeze the object when grasping it, etc. It is also not stated where the object is located in the environment. All these motion parameters have to be inferred from the robot's knowledge and supplemented at run time with perceptual information.

def-plan fetch&place (*?object*, *?destination*,
?loc-for-fetch, *?loc-for-place*)

1. **with-robot-at-location** *?loc-for-fetch*
   perform (an action
   (type fetching)
   (object *?object*))

2. **with-robot-at-location** *?loc-for-place*
   perform (an action
   (type placing)
   (object *?object*)
   (destination *?destination*))

Figure 1.8: A generalized action plan for fetch&place.

The *with-robot-at-location* construct of the plan language ensures that during the action the robot base is located appropriately. For example, for performing the *fetching* action the robot should stand in a location where it is able to perceive and reach the object. If the robot is not at the appropriate location, performing the action is suspended, the robot repositions itself and only then continues the action.

The body of the plan specifies the high-level logical structure of the plan. The plan body consists of two steps. The first step tells the robot to fetch a specific object while standing at an appropriate location for fetching. The second step tells the robot that the object has to be placed at the given destination, and that the robot should ensure that it is standing at the appropriate location from where the destination is reachable. The underdetermined nature of the plan implies that if the exact position of the object is not known then the robot has to search for it. The search process will be more efficient if the robot knows places where the object is likely to be. The robot can infer this knowledge from its knowledge base. Similarly, the manner in which the object has to be picked up can only be decided after the object is found and its geometry and state as well as the scene context are known.

The challenge of contextualization

The task of the plan executive is to contextualize the underdetermined action descriptions[25] in order to adapt the action descriptions to the specific context, including the object to be acted on, any tool to be used, the current situation, the current environment, the user preferences, and the robot capabilities. The challenge of action contextualization is that the space of behaviors and the space of possible task contexts in everyday activities are huge and open-ended. Consequently, it is a very demanding reasoning exercise to generate a candidate behavior that can achieve the desired outcomes while avoiding unwanted side effects.

In order to contextualize action descriptions, the plan executive generates a sequence of queries for the knowledge base and the perception executive. For example, a query might be “how should I grasp the object that I intend to pick up?”. The reasoner then has to infer that grasping the handles of the object from both sides would be appropriate because the object has two handles and is too heavy to be grasped with one hand. Another way of contextualizing actions is through perception queries. To this end, the plan executive asks the perception executive to detect the object that is to be transported based on the underdetermined description, such as the object that contains the soup. This perception task then returns the required information, e.g. the geometry and pose of the object, which are needed to parameterize the motion.

Motion plans

The resolution of high-level action designators generates *motion schemata* that tell the plan executive how to execute the constituent motions. The motion schemata are represented as **motion plans** that structure the motion into motion phases. For example, the motion plan for picking up an already detected object in the robot’s reach and placing it at another location is shown in Figure 1.9.

A motion plan for fetch&place actions comprises four different motion phases: *reaching*, *lifting*, *transporting*, and *releasing* (Flanagan et al., 2006).[26] Each motion phase has a goal. When the goal is achieved, the start of the subsequent motion phase is triggered. Goals can be force-dynamic events, e.g. the robot finger coming into contact with the object to be grasped, or other perceptually distinctive events, e.g. a milk carton becoming visible when a fridge door is opened. Each motion phase comes with several parameters that can be set in order to adapt the motions to the current context. For example, when reaching for an object, the parameters could be constrainted for the reaching motion, the grasp type, and the position of the robot grippers on the object.

We note again that the EASE working hypothesis is that for every action category we can specify a motion plan schema with a small number of motion parameters that is sufficient to achieve the desired outcome and avoid unwanted side effects, and that we can do so for a large variety of objects and tasks by leveraging the constraints imposed by the current context.

[25] Recall: performing an underdetermined action description is accomplished by contextualizing the action description at run time. This is the equivalent of resolving the associated action designators.

[26] The paper by Flanagan et al. (2006) refers to *action phase* rather than *motion phase*. We use *motion phase* here to connote a lower level of abstraction and avoid confusion with the higher level of abstraction entailed by a generalized action plan and an action designator.

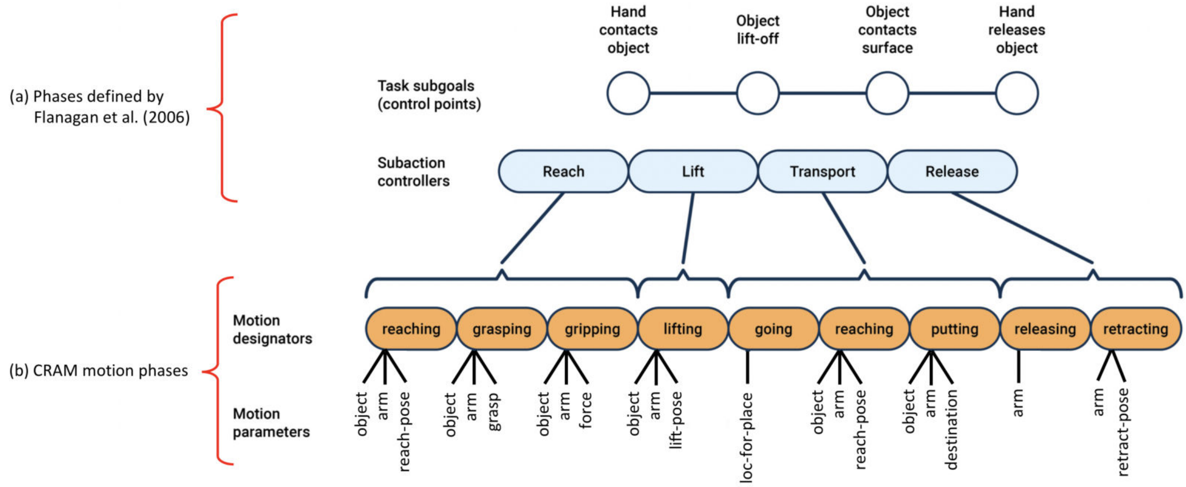


Figure 1.9: Motion plan for the *fetch&place* generalized action plan. (a) The motion phases of a pick and place task as described in (Flanagan et al., 2006): motion phases, each governed by a subaction controller, are depicted in blue boxes while the control points depicted by circles represent the task subgoals; refer also to footnote 26. (b) The realization of this approach in CRAM, showing the motion designators that comprise each motion phase depicted in orange boxes, together with their associated motion parameters.

Interpretation of action plans

Action plans are invoked through requests in the form of action descriptions such as “fetch a cup and put it on the kitchen table” in the activity context of setting the table. This can be formalized as follows:

perform (an action
    (type *?category*)
    (*?key-1 ?value-1*)
    (*?key-2 ?value-2*)
    ...)

For example, filling in *?category* and the key-value pairs in the description, we have:

perform (an action
    (type fetch&place)
    (object (an object
        (type cup)))
    (destination (a location
        (on (an object
            (type kitchen-table)))))))

As noted already, a request to *perform (an action (type ?category) ...)* translates into a query to sample the joint probability distribution of motions that the action plan for *?category* generates and the physical effects that these motions cause. The purpose of the reasoning by the plan executive is to refine action descriptions to exclude motion parameterizations that do not achieve the desired outcomes or might cause unwanted side effects.

To find these motion parameterizations, the plan executive first instantiates the request in the body of the contextualization plan as shown in Figure 1.10 (left). The plan executive next extends the generalized action plan by adding the parameters needed to execute the motion plan; see Figure 1.10 (right). The plan executive then queres the generative model for values of the parameters of the motion plan that would generate a body motion to achieve the goal of the action description; see Figure 1.11.

1. **with-robot-at-location** *?loc-for-fetch*
   perform (an action
   (type fetching)
   (object *?object*))
2. . . .

1. **with-robot-at-location** *?loc-for-fetch*
   perform (an action
   (type fetching)
   (object *?object*)
   (arm *?arm*)
   (grasp *?grasp*)
   (lift-pose *?lift-pose*)
   ...)
2. . . .

Figure 1.10: (left) Instantiating the *fetch&place* generalized action plan for the underdetermined action description; (right) extending the instantiated generalized action plan by adding the parameters needed to execute the motion plan.

Choosing the appropriate motion parameterization requires the robot agent to exploit its knowledge through reasoning. Some of the knowledge will be based on previous experience and some knowledge will be based on prospection using internal simulation.

query-variables (*?loc-for-fetch*, *?arm*, *?grasp*, *?lift-pose*, . . .) **to-succeed** (
  with-robot-at-location *?loc-for-fetch*
    perform (an action
      (type fetching)
      (object *?object*)
      (arm *?arm*)
      (grasp *?grasp*)
      (lift-pose *?lift-pose*)
      . . .)
)

Figure 1.11: Body-motion-query: what are the values of *?loc-for-fetch*, *?arm*, *?grasp*, *?lift-pose* that produce robot body motion that maximizes the likelihood that the requested action will succeed?

The use of a generalized action plan implies that the ability to accomplish everyday manipulation tasks is the result of reasoning with knowledge and that some of this knowledge encapsulates constraints imposed by context, supplemented by perceptual knowledge acquired at run time. The knowledge needed to refine action descriptions can be represented as generalized and modular knowledge chunks that can be composed through a reasoning engine to derive appropriate motion parameterizations for novel tasks, situations, and contexts. Thus, the robot can infer parameter values for the motion plan that are likely to achieve the desired outcome of the task and avoid unwanted side effects (see Figure 1.11).

1.2.3.1.3 Summary of the EASE approach to robot agency

EASE is founded on the concept of a generative model and a conceptual framework for accomplishing everyday manipulation tasks. These are based on the following three hypotheses:

- We can provide for each manipulation action category such as *fetch*, *place*, *pour*, *cut*, *wipe* a general motion plan schema with motion phases and phase-specific motion parameters that can generate a range of body motions to achieve the respective goal of the action in a large variety of contexts.
- A request for performing actions can be represented by declarative, symbolic, and underdetermined action descriptions that are to be refined through knowledge, reasoning, and perception in order to infer the motion parameter values that generate a body motion to achieve the goal of the action description.
- The knowledge needed to refine action descriptions can be represented as generalized and modular knowledge chunks that can be composed through a reasoning engine to derive appropriate motion parameterizations for novel tasks, situations, and contexts.

To perform an underdetermined action description, we contextualize it by (a) instantiating the associated generalized action plan, (b) extending it by adding the parameters needed to execute the motion plan, and (c) creating a query for the parameter values that maximize the likelihood that the associated body motions will successfully accomplish the desired action. The action executive then adaptively executes the body motions.

The use of a generative model also caters for the cases where refinement through contextualization is not possible. It does this by providing the robot agent with the ability to reprogram the generalized action plan for specific task variations and contexts. Cases where such reprogramming is necessary are, for example, closing a door by pushing with the elbow or the foot instead of grasping the handle and moving the hand according to the articulation model of the door. The generative model accomplishes this reprogramming through transformational planning and learning.[27]

The plan executive and the action plans are designed to facilitate introspective reasoning: the inference and perception tasks are represented in the plan in a modular and transparent form and the queries generated during plan execution and their answers are recorded together with the success and failure of the corresponding action. This provides the robot with a form of computational awareness: the means to reason about the inferences (and the reasoning that produces them) when inferring what constitutes appropriate context-specific behaviour. Consequently, the robot can answer queries about the inference tasks it needs to solve to perform an action, about any information that is missing when determining the appropriate behavior, about the inferred proposed behavior, about whether the behavior would achieve the desired outcome, and about any unwanted side effects.

In the EASE generative model, this awareness of body motion reasoning is an essential factor for the cognitive development of everyday manipulation capability because it enables the robot agent to assess the reasoning mechanisms and substitute inference mechanisms with better ones, if necessary. Thus, the robot can improve its manipulation capability by improving its reasoning capability.

In Section 1.2.3.1, we introduced the idea that a quantum leap in innovation is often brought about by the introduction of a new interface layer in a complex system. In Section 1.2.3.1.2, we introduced two ideas. First, everyday underdetermined actions can be accomplished by viewing them as a problem

[27] In our previous work we have investigated sophisticated techniques for transformational planning and learning (Beetz, 2000; Belker and Beetz, 2001; Müller et al., 2007). We have recently started to apply these techniques to control programs that autonomously control manipulation tasks of real robots (Kazhoyan et al., 2020b).

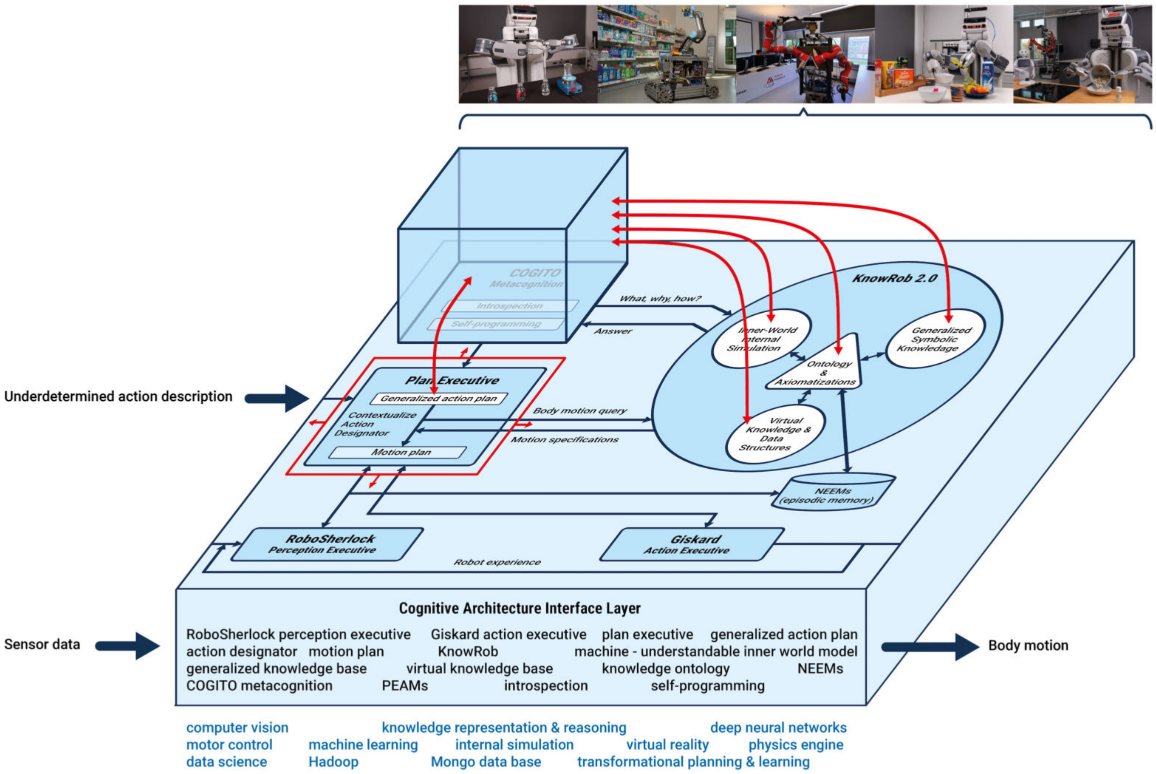


Figure 1.12: The cognitive architecture viewed as an interface layer exposing a self-programmable generalized action plan interpreter that can be reprogrammed by the metacognition system (above the interface) to effect generalization and specialization of plans. This results in an extended generalized plan language (exposed by the interface and depicted in white) and an extended interpreter (shown by the red rectangle). Abstract knowledge in the knowledge ontology, the machine-understandable inner-world model, the virtual knowledge base, and the generalized knowledge base (all exposed by the interface and depicted in white) are updated accordingly. This abstract interface allows different robots to be used for different tasks in different domains. All the implementation processes and representations associated with the contextualization of the generalized plan language are hidden below the interface layer.

of finding the robot motions that would be most likely to succeed in achieving the required action outcomes without any undesired side effects. Second, this problem can be solved by sampling of a joint distribution of (a) the motion parameters associated with the action category, and (b) the physical effects that these motions cause.[28] Here we will draw together these two ideas to explore the full power of the EASE perspective on robot agency.

Recall that we require a programming language and interpreter to generate robot body motions for manipulation tasks, given underdetermined action descriptions or, equivalently, generalized action plans. This programming language must leverage abstract knowledge, allowing the robot to understand what it is doing, accomplish novel manipulation tasks, and learn from very few examples. This language must be implemented in some execution system. We suggest that the programming language is the generalized action plan language and that the cognitive architecture is an interpreter for this language. Thus, the cognitive architecture can be viewed as the interface layer beneath which lie the complex mechanisms that (a) map instantiated generalized action plans to likely-to-succeed motion parameter values and (b) execute these motions adaptively; see Figure 1.12. Above this interface layer lie the

[28] In particular, the robot's observations of those effects.

mechanisms by which generalized action plans are flexibly instantiated and extended. The role of the cognitive architecture interface layer, then, is to provide an abstract interface to the execution system which encapsulates the cognitive mechanisms required (a) to identify the motion parameter values that will produce the robot body motions most likely to succeed in achieving the required actions and then (b) to adaptively execute these motions.

The cognitive mechanisms operate on the joint distribution of motions and their effects, of which there will be many (at least one for each action category). These joint distributions are generative internal models which, when subjected to the constraints imposed by reasoning with contextual knowledge, yield the sample in the joint distribution that maximizes the likelihood that the action will succeed. This sample has the maximum expected success measure over all internal models that are relevant in the current context. These generative internal models can be learned from experience and are composable so that they can be recombined to yield novel action strategies.

The interface layer, i.e. the cognitive architecture, interprets the generalized action plan by executing the three steps of contextualization and then, by deploying the action executive, adaptively executing the parameterized motion plan using parameter values produced in the third step in the contextualization process. However, because the cognitive architecture is a plan interpreter, it is also a symbolic program. This means that metacognitive processes above the interface layer can re-program the intrepreter to extend the general action plan language it interprets. This extension improves the ability of the cognitive architecture to identify the robot body motions that are likely to succeed in accomplishing an underdetermined action description. This extended language expresses new generalized action plans which are generated by the metacognitive processes that implement transformational planning and learning. The metacognitive processes also generate specialized action plans that exploit PEAMs (pragmatic everyday action manifolds) to achieve feasible solutions to otherwise intractable problems by identifying the constraints that knowledge of everyday activies and the environment can bring to bear; see Section 1.2.3.2.2.5.

The EASE approach to robot agency implements this metacognition in a unique way. While most cognitive architectures view metacognition as a separate independent module responsible for the oversight of the performance of the cognitive architecture, the CRAM cognitive architecture implements the metacognitive functions (of plan generalization, plan specialization, and extension of the plan language and plan language interpreter) using the plan executive itself, recruiting the KR&R executive to effect the necessary inference mechanisms. It achieves this by the introduction of a third knowledge base in the KR&R executive: the **virtual knowledge base** (see Figure 1.7). This knowledge base is in effect a view (in the technical sense) of the plan executive, i.e. a dynamically-instantiated abstract representative of the CRAM plan language interpreter. This virtual knowledge is also axiomatized in the knowledge ontology to expose the semantics of the interpreter. This means that the knowledge representation and reasoning executive can then be used by the plan executive to reason about itself and thereby achieve the metacognitive functions mentioned above. In this way, the plan executive effectively self-programs and metacognition can be considered to be a logical extension of the plan executive with reentrant processing.

In summary, the plan executive then has three distinct responsibilites:

1. Plan execution via contextualization using the generative model.
2. Plan recovery via plan monitoring and failure handling.
3. Plan language extension via metacognition: plan generalization and plan specialization.

The KR&R executive is involved in all three of these responsibilities.

Planned Extensions

One of the main goals of the next phase of EASE is to target the development of techniques to leverage the adaptive capacity of CRAM. We will do so in three different aspects of the operation of CRAM: (a) the metacognition process, (b) the contextualization process, and (c) the action execution process, all of which will provide CRAM with the flexibility and context-sensitivity that are characteristic of cognitive behavior (Schneider et al., 2020).

Some of these developments will be based on research in computational knowledge-based transformational planning and transformational learning in the EASE research area R, while others will be based on research in areas H and P, all leveraging, among other things, the power of the recently introduced situation model framework (Schneider et al., 2020). We describe these research plans in detail later in the proposal in Sections 1.2.5 and 1.2.6. Here, we wish to motivate this research by highlighting further evolution of the CRAM cognitive architecture described below in Section 1.2.3.2. To set the scene, we will describe very briefly some of the main features of the situation model framework; for a more complete and detailed summary, see Section R01.4.2.1.

The Situation Model Framework The Situation Model Framework was introduced by Schneider et al. (2020) as the basis for understanding how cognitive behaviour, in general, and flexible context-sensitive cognitive behaviour, in particular, is realized in humans, animals, and machines. It is based on five main concepts: (i) behavioral episodes, (ii) the two-systems approach to thought and action, (iii) the capacity limitation of working memory, (iv) the need for attentional control, and (v) the representation known as a cognitive map. These concepts are bound together in a single framework, specifically the framework of situation models. We focus here on behavioral episodes, the two-systems approach, and cognitive maps.

Situation models build on the new concept of a behavioural episode which links the functional elements of perception, long-term memory, and motor control for action. The elements that comprise a behavioural episode are objects, scenes (arrangements or layouts of objects), actions, and action outcomes. Collectively, this set of four elements comprises the behavioural episode and it captures a temporal causal relationship between its four elements.

There are two classes of behaviour, one involved in carrying out routine habitual actions, and one involved in carrying out actions that required deliberation. These behaviors are achieved using two complementary systems: system 1 and system 2 (Kahneman, 2011); see, also, (Norman and Shallice, 1986). System 1 is used for habitual actions and system 2 is used for flexible actions. In habitual actions, system 1 retrieves a number of behavioral episodes, subjects them to a winner-take-all competition, and selects one winning behavioral episode. This winning behavioral episode then controls the action by filling in the required sensor or motor information in real time. In system 2, the behavioral episodes can be recalled (and optionally modified) or newly constructed and then simulated to assess the outcome of the action. Thus, system 1 involves reactive or "automatized" (Schneider et al., 2020) control of action based on previous experience while system 2 involves prospective control of action. Both use behavioral episodes.

A cognitive map is a representation of the associative link between (perceptual) cues, actions, the outcome of those actions (and, strictly speaking, the perception of these outcomes). It also encapsulates any other relevant environmental (i.e. contextual) information that either conditions the link or perhaps captures any second-order effect of the action. It is significant that a cognitive map represents a change in state. Specifically, it captures the causal relationship that binds together perception, action, and perceived action outcome, in a manner that is analogous to behavioral episodes. Cognitive maps differ from behavioral episodes, however, in that they are usually associated with events and behaviours that take place over extended periods of time.

Metacognition Since behavioral episodes and, more globally, cognitive maps which capture the spatial and temporal structure of assemblies of behaviour episodes both encapsulate the sensorimotor and causal relationship among objects, scenes, actions, and action outcomes, EASE research area H will target sensorimotor and causal modelling, among other topics. The results of this research will facilitate flexibility in metacognitive expansion of existing capabilities, effectively generating new knowledge and new action capabilities, and allowing the robot to operate in unexpected or novel situations, adapting both the plan language and the generative model.

Through metacognition, we seek to expand action capabilities of a cognitive robot. To support this, research area H will also target the development of hybrid discriminative and generative models of human activity (exploiting contex-free grammars and probabilistic action units and deep multi-modal networks) to identify new ways of describing the temporally-extended hierarchical organization of motion primitives that comprise complex actions. By covering both discriminative and generative modelling, this research will leverage the respective strengths of each: (a) directly learning the posterior distribution that characterizes the space of action sequences comprising everyday activities and (b) inferring the posterior distribution by learning the joint distribution over motion values and complex actions (Ng and Jordan, 2001).

Humans exhibit flexibility and context-sensitivity in purposeful and goal-directed activities. In order to understand how they do this, and thereby inform the computational metacognitive process in a cognitive robot, research area H also plans to model human learning and decision-making processes, focussing on what decisions are necessary and when they are necessary to master complex everyday activities. This is particularly important when generalizing behavior to new decision situations. The objective is to find the optimal trade-off between exploiting object- and situation-specific knowledge and abstract knowledge by understanding how humans acquire knowledge that allows generalizing beyond the distribution of the data which characterizes the situation in which they committed errors and learned, even with very sparse experience.

Contextualization The research on sensorimotor and causal modeling in area H will also provide the basis for the flexibility in the assembly, recombination, and construction of new behavioral episodes required for the effective operation of system 2 through more effective sampling of the joint distribution in the generative model. Causal modelling will also benefit attentional processes, both internal and external.

The hybrid discriminative and generative modelling of human activity in area H will provide essential insights into the compositionality of behavioral episodes in system 2 and the deliberative process of recombining and constructing new behavioral episodes. So too will the research on modeling human learning and decision-making processes, especially when a robot is presented with ambiguous situations, uncertainty about action plans, and the processing of interfering information.

Action Execution The research on sensorimotor and causal modelling in area H will also contribute to the achievement of the flexibility that is necessary during action execution for adaptive movement generation.

Epilogue to the EASE generative model of robot agency

In motivating the planned extensions above, we referred mostly to the impact of work in EASE research area H. However, EASE research areas R and P will contribute equally to the realization of the planned advances in the metacognition, contextualization, and action execution processes by leveraging insights from the situation model framework, and, in particular, by developing the associated computational formalisms that will underpin the next generation of the CRAM cognitive architecture. In the sections that follow, specifically Sections 1.2.3.2 – 1.2.3.10, we set out the capabilities of the current version of the CRAM cognitive architecture and the key results from the first phase of EASE, before addressing the research goals and research plans for the second phase in Sections 1.2.5 and 1.2.6, respectively.

1.2.3.2 The CRAM cognitive architecture: a realization of the generative model and generalized action plan framework

Having described the first landmark achievement of the first phase of EASE, i.e. the conceptualization, articulation, and design of this generative model and generalized action plan framework, we now proceed to visit the other results of Phase 1. As we already noted, we will do this with reference to the CRAM cognitive architecture, which is the way in which the generative model and generalized action plan framework are realized in EASE. To set the scene, we begin below with an overview of CRAM. This is followed by Sections 1.2.3.3 – 1.2.3.8, which describe in detail the six key results previewed in the introduction to Section 1.2.3 and summarized in Figure 1.4, including a description of the way in which we have advanced our understanding of how humans accomplish their everyday activities. We conclude the presentation of the results of EASE Phase 1 with a summary of the key scientifc insights in Section 1.2.3.9.

1.2.3.2.1 A cognitive architecture for robot agents

The success of EASE depends on robot agents being equipped with software that enables them to perceive their environments and produce competent actions. For robot agents to physically perform manipulation activities as complex as required by the EASE robot days and years is particularly challenging. Below we will explain the current state of the EASE software infrastructure.

> The robot control system developed in EASE is unique worldwide because of:
> - the complexity of the fine-grained manipulation tasks that it tackles;
> - the pervasive integration of AI technology including knowledge representation and reasoning;
> - the depth of integration of AI technology into perception and manipulation and control;
> - the methods for automated logging; and
> - the commitment to providing important software components open-source.

EASE uses CRAM (Cognitive Robot Abstract Machine) as a cognitive architecture for implementing the control systems for robot agents. CRAM integrates perception, motion control, plan-based control, and other cognitive capabilities in a coherent software framework that facilitates its embodiment in robot agents. It has been extended by EASE researchers and this development has been continued through the first funding phase of EASE. The main extensions introduced in EASE include the introduction of an episodic memory framework, the provision of an integrated framework for robot learning, and extensions of its components for perception, knowledge representation and reasoning, plan-based control, and motion control, which are described below.

The current version of the CRAM cognitive architecture is shown in Figure 1.13, which also implements the generative model of the first EASE phase discussed in Section 1.2.3.1. The main functional components of the architecture are (a) the plan excutive of CRAM, (b) the knowledge representation & reasoning (KR&R) framework KNOWROB, (c) the perception executive ROBOSHERLOCK, (d) the action executive GISKARD, and (e) the metacognition component COGITO. The plan executive interprets the generalized plan schemata in order to accomplish an underdetermined request. It does so by analyzing the difference between what knowledge is required by the motion plan and what is known about the action so far in order to formulate and issue the query for the body motion parameter values to fill the knowledge gaps. These queries are then answered through the KR&R system KNOWROB and the perception executive ROBOSHERLOCK. The sufficiently refined motion plans are then executed by the action executive GISKARD. In addition, the meta-reasoning component COGITO enables the robot agent to adapt the structure of its high-level and motion plans through transformational learning and planning.

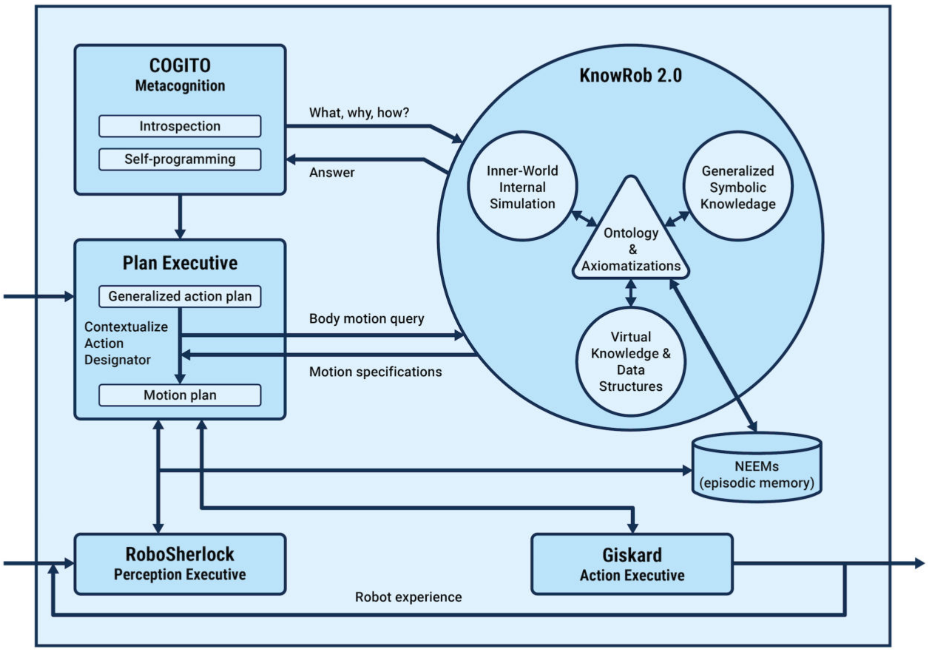


Figure 1.13: The current version of CRAM, the EASE cognitive architecture.

One of the main goals of the second funding phase of EASE will be the design, implementation, and investigation of the next generation of CRAM, which will orchestrate the cognitive capabilities for mastering everyday manipulation tasks and for evolving the robot capabilities towards such mastery.

An implementation of the CRAM (Beetz et al., 2010b) open-source software toolbox can be found at `http://www.cram-system.org`. Several prizes have been awarded in relation to the research in cognition-enabled control and its components: The PR2 beta program awarded an autonomous mobile manipulation platform for the CRAM project. Related publications received several best paper awards (AAMAS, ICRA, ICAR) or were finalists for such awards (IROS, ICRA).

A relevant source for the assessment of CRAM as a tool for realizing cognition-enabled robot agents is the H2020 coordination action RockEU2, which among other topic items conducted activities in market observation, technology watch, innovation support, analysis of funded proposals, regulations assessment, and standardisation support. In RockEU2 the catalogue of cognitive systems capabilities aims at describing the state of the art of the corresponding research area. In the catalogue, CRAM is assessed to provide nine of the top eleven industrial requirements[29] for cognitive robots (for details see (Vernon and Vincze, 2017) page 11-13) and fulfill all the cognitive functional requirements that have been identified in the same study.[30]

[29] The top eleven industrial requirements for cognitive robots according to Vincze and Vernon (2017) are (1) safe, reliable, transparent operation, (2) high-level instruction and context-aware task execution, (3) knowledge acquisition and generalization, (4) adaptive planning, (5) personalized interaction, (6) self-assessment, (7) learning from demonstration, (8) evaluating the safety of actions, (9) development and self-optimization, (10) knowledge transfer, and (11) communicating intentions and collaborative action.

[30] The cognitive functional requirements for cognitive robots identified in the H2020 coordinating action RockEU2 are perception, declarative knowledge & memory, procedural knowledge & memory, planner & plan executive, reasoning & inference mechanisms, metacognition, environment model, internal simulator, goal representations, declarative learning, procedural learning, attention mechanism, and real-time action controller.

The catalogue states: *"... one of the very few open source cognitive architectures that has the demonstrated potential to be deployed as a complete cognitive robotic framework:* CRAM *(Cognitive Robot Abstract Machine). We don't suggest all cognitive robots should be based on* CRAM *but it does prove an excellent exemplar of a maturing system of interoperable AI tools and techniques that can be freely used by the community, especially as a vehicle for industrial deployment."*

1.2.3.2.2 Functional components of the cognitive architecture

In the following sections, we will summarize the main aspects of the five functional components of CRAM: the plan executive, the KR&R framework KNOWROB, the perception executive ROBOSHERLOCK, the action executive GISKARD, and the metacognition component COGITO.

1.2.3.2.2.1 Plan executive

The plan executive interprets and executes plans written in the CRAM plan language. CRAM plans are concurrent, event-guided control programs that specify how the robot has to respond to sensory events in order to successfully perform an action. What turns the programs into plans is that they are represented explicitly, transparently, and modularly such that the robot agent can reason about the programs and modify them at execution time.

One of the key achievements in the first EASE phase was a novel design and implementation of a generalized fetch&place plan that can generate the context-specific behavior to accomplish table setting and cleaning tasks and facilitates the improvement of behavior through planning and learning (see the experiments in Section 1.2.3.3) (Kazhoyan and Beetz, 2017, 2019a). The distinctive capabilities of the plan design are the complexity and the variability of fetch and place behavior that a single plan schema can generate. These variations include active object search, optional support actions such as opening and closing containers in order to fetch objects, and dynamic behavior adaptations including the skipping of unnecessary subactions. In addition, the plan schemata include sophisticated methods for failure detection and handling, as well as the context-dependent continuation of the primary activity after the failure recovery.

To enhance failure tolerance, Bauer et al. (2020) propose to predict the effects of robot actions by augmenting collected experience with semantic knowledge and leveraging realistic physics simulations by considering semantic similarity of actions in order to predict outcome probabilities for previously unknown tasks. The physical simulation is used to gather simulated experience that makes the approach robust, even in extreme cases, and can be used to predict action success probabilities.

In addition, the reasoning methods for plan-based control have been complemented with a powerful plan projection method (Kazhoyan and Beetz, 2019b), which can be seen as very fast built-in physics simulation. Also, the reasoning tasks have been represented explicitly in the plans which facilitates experience-based learning through self-specializing plans (Koralewski et al., 2019). Furthermore, Kazhoyan et al. (2020b) have made substantial steps toward the realization of a comprehensive and powerful framework for transformational learning and planning on real-world robot plans. These research activities are to be continued to invent a fully functional plan transformation framework for self-programming that were so far only possible in fairly simplified forms of simulated robot agency (Müller et al., 2007; Beetz, 2000).

Another distinctive property of the CRAM plan executive is that it represents the behavior generating plans, the computational processes they initiate, the motions they generate, and the physical effects that the motions cause as explicit symbolic knowledge structures (Mösenlechner et al., 2010). It also represents the causal relationships between these knowledge structures. This recorded knowledge structure enables the robot agents to answer queries about what the robot does, why it does it, how it does it, and what is happining. It also allows the robot agent to diagnose its behavior, by inferring answers to questions such as: "Could the goal of the action be achieved?" "Did the robot not attempt to transport an object because it has not seen it?" Finally, through these knowledge structures the robot can identify the subplans that are responsible for certain effects, which enables powerful plan transformation and debugging methods.

Meywerk et al. (2019) formally verified a general fetch and place plan from a simple CRAM shopping scenario. It involves a robot that is to move objects between shelves and a table, satisfying the given constraints. This verification works by translating the generalized plan into more lightweight intermediate plan representations to simplify formal reasoning. The intermediate representations constitute a compact, yet powerful language that can capture the simulation semantics of basic cognition-enabled robotic plans and can be verified by the verification tools proposed in Subproject P04.

1.2.3.2.2.2 Knowledge Representation and Reasoning Framework

EASE uses the KnowRob knowledge system as a key component for representing knowledge and reasoning about it. KnowRob is currently one of the most advanced knowledge processing systems for robots. It has enabled robot agents to accomplish complex manipulation tasks such as making pizza, conducting chemical experiments, and setting tables. The KnowRob knowledge base appears to be a conventional first-order time interval logic knowledge base, but it exists to a large part only virtually: many logical expressions are constructed on demand from data structures of the control program, computed through robotics algorithms including ones for motion planning and solving inverse kinematics problems, and from log data stored in noSQL databases. KnowRob enables robots to acquire open-ended manipulation skills and competence, reason about how to perform manipulation actions more realistically, and acquire commonsense knowledge.

We will examine KnowRob in more detail in Section 1.2.3.4 as it is the second key result of Phase 1 of EASE.

1.2.3.2.2.3 Perception executive

An important source of knowledge that the robot agent needs to access in order to accomplish its manipulation tasks are the images captured by the robot's camera. In order to refine underdetermined action descriptions, the perception executive needs to be able to detect objects in a scene based on partial object descriptions and extract specific information about the object to be manipulated, the scene context, and the environment structure.

Robot perception in EASE is realized using the perception executive RoboSherlock[31] (Beetz et al., 2015a) (Best Service Robotics Award ICRA 2015).

RoboSherlock provides a symbolic language that enables the robot agent to specify a large variety of perception tasks that need to be solved in order to accomplish underdetermined everyday manipulation tasks. In this language, perception tasks are stated in terms of *object descriptions*, *object hypotheses*, and *task descriptions*. Using these descriptions, a robot agent can describe a red spoon using the following construct: *(an object (category spoon) (color red))*. The command to *detect* an object description asks the perception system to detect objects in the sensor data that satisfy the description and return the detected *hypotheses*. In more detail, an object detection task has the following form:

detect (an object
    (category *?category*)
    (*?key-1 ?value-1*)
    (*?key-2 ?value-2*)
    ...)

The attributes that can be used in object descriptions include *shape*, *color*, *category*, *location*, *pose*, *CAD-model*, and *part-of*. In particular the *category* attribute is very expressive as it allows for the application of self-defined categories. Suppose you have a classifier that can infer the affordances of object hypotheses, then you could ask the perception system to detect objects in a scene that afford a given action. Also, by combining visual detection with knowledge-enabled reasoning and other forms of computations, such as computing volumes, RoboSherlock can also accomplish perception tasks such as "is there a container that can hold more than half a liter"?

[31] `http://robosherlock.org/`

Query results	Perception task
	**detect** (an object (category 'FoodOrDrinkOrIngredient')) **examine**[a] ?id [:size, :pose] **detect** (an object (color yellow) (shape round)) **detect** (an object (location (a location (next-to (an object (category 'ElectricalDevice')))))) **examine** ?id [:geom-primitive, :logo] [a]Examine tasks can only be called for existing object hypotheses
	**detect** (an object (category 'PancakeMaker') (location (a location (on (an object (category 'CounterTop')))))) **detect** (an object (color black) (shape round) (category 'ElectricalDevice'))
	**detect** (an object (category 'CookingUtensil') (location (a location (in (an object (type container) (category drawer#3)))))) **detect** (an object (shape flat) (color black) (location (a location (in (an object (type container) (category drawer#3))))))

Figure 1.14: ROBOSHERLOCK as a vision-based question answering system.

Thus, we measure the perceptual capabilities provided by our generative model in terms of the perception tasks that can be solved.

ROBOSHERLOCK can handle the perception tasks issued by the robot agent in order to accomplish in a robust manner the EASE robot days reported in Section 1.2.2.2. Example perception tasks are shown in Figure 1.14. Current extensions investigate the scaling toward more complex and realistic scenes, including a cluttered fridge, dishwasher, and oven, described below (see Figure 1.18).

An essential robot vision expert employed by ROBOSHERLOCK is ROBOTVQA (Kenfack et al., 2020), which is a scene-graph and deep-learning-based visual question answering system for robot manipulation. At the heart of ROBOTVQA lies a multi-task deep-learning model that infers formal semantic scene graphs from RGB(D) images of the scene at a rate of about 5 fps. The graph is made up of the set of scene objects, their description (category, color, shape, 6D pose, material, mask) and their spatial relations. Moreover, each of the facts in the graph is assigned a probability as a measure of uncertainty. The estimated scene graphs are represented using a probabilistic lightweight description logic.

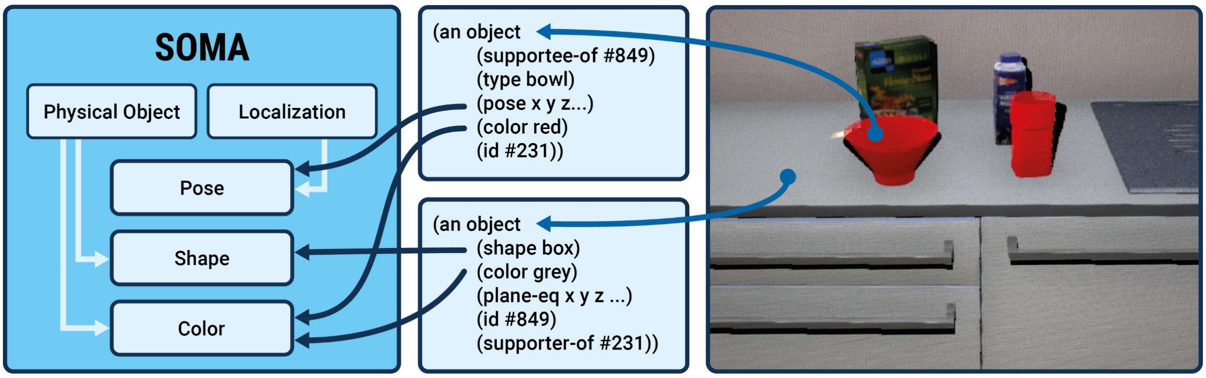


Figure 1.15: ROBOSHERLOCK reconstructs the scenes captured by the robot camera as virtual reality scenes using the models of known objects, which are together with background knowledge represented in the symbolic knowledge base, named SOMA (see Section 1.2.3.5 for more information on SOMA).

ROBOSHERLOCK increasingly uses self-supervised and self-aware learning methods leveraging inner-world models of the robot environments (Mania and Beetz, 2019; Kenfack et al., 2020), which together with the images contained in the episodic memories (NEEMs) (Bálint-Benczédi and Beetz, 2018; Balint-Benczedi et al., 2017) are sufficient to learn robust real-world perception methods.

The overall perception cycle that maintains the almost photo-realistic belief state of the generative model is based on imagination-enabled robot perception[32]. We propose a perception system that maintains its beliefs about its environment as a scene graph with physics simulation and visual rendering. This perception cycle retrieves models of expected objects and places it at the corresponding place in a virtual reality based environment model (see Figure 1.15). The physics simulation ensures that object detections that are physically not possible are rejected and scenes can be rendered to generate expectations at the image level.

In addition, we aim to make the hypothesis generation more robust by discarding object detections that are physically impossible, such as objects penetrating other ones or objects in positions that are physically not stable. There is substantial evidence that such physical reasoning is deeply compiled into the human vision capabilities (Battaglia et al., 2013).

Another direction of research is to make perception more action-aware. To this end, we investigate how ROBOSHERLOCK can be extended to answer perception tasks such as:

- “how could a detected object participate in an action?” which requires the system to infer the roles objects could possibly take in actions,
- “how can a given underdetermined action be executed in the scene that I am seeing?” which requires the system to simulate possible action descriptions in the observed scene, or
- “How could a robot agent replace an intended object with one that is contained in the captured scene?” which might even require a ranking of different objects with regard to their suitability.

Realizing perception capabilities with such functionality becomes possible by leveraging the NEEM collections and the segmented and semantically annotated visual data contained in them.

In summary, ROBOSHERLOCK is a taskable, knowledge-enabled perception framework that uses an extensible ensemble of perception experts to accomplish perception tasks. ROBOSHERLOCK perception experts are special-purpose routines that are employed in the respective perception contexts. Thus, rather than applying a general-purpose plate detector, ROBOSHERLOCK uses context-specific plate detectors. During table setting, it might use one that detects the topmost white horizontal lines in cupboards, while it uses detectors for ovals when cleaning the table. It also might use texture detection

[32] https://www.ease-crc.org/link/video-imagination-enabled-robot-perception

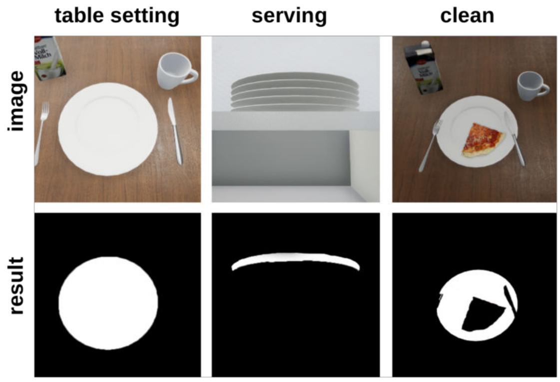


Figure 1.16: Context-specific methods for detecting plates in the EASE household challenge.

when it assumes the plates to not be clean, and so on (see Figure 1.16). Thus, given a perception task and the current context information, RoboSherlock executes the appropriate perception experts to generate candidate solutions for the given perception task and even generates context-specific perception pipelines. The candidate solutions proposed by experts are then tested and ranked to find the best solution. The advantage of RoboSherlock over other perception approaches is that it can combine knowledge with perception and use knowledge-enabled reasoning about objects and scenes to tailor perception capabilities to the respective contexts and thereby make it more effective, robust, and efficient.

Figure 1.17: Expectation for belief state estimation generated through imagistic rasoning.

Within EASE, the capabilities of RoboSherlock were substantially improved through simulation- and rendering-based environment representation (EASE Subproject R03 “Simulation-based reasoning”). Mania et al. (2020) proposed an extension of RoboSherlock that aims at replicating what it sees as an internal belief state implemented through virtual reality technologies (see Figure 1.17). The knowledge base of the robot is populated with object models that consist of CAD models, including the part structure and possible articulation models, a texture model, as well as encyclopedic, common-sense, and intuitive physics knowledge about the object. This imagination-based scene perception approach has the advantage that the robot has perfect knowledge about everything that is contained in the belief state. A second advantage is that the robot can compute very detailed and realistic image-level expectations about what it expects to see. These expectations are used to estimate object poses very accurately and to save computational resources. The approach works as follows: the belief state is rendered from the camera pose as an image. This synthesized image is then compared with the image captured by the robot camera in the real environment.

A second novel capability introduced in RoboSherlock in the first phase of EASE is the facility to learn special-purpose perception routines through a framework for self-training in simulated photorealistic environments (Mania and Beetz, 2019). The framework enables robot agents to use their environment and object representations in order to generate training data for supervised learning for perception tasks. For training, the framework does not only allow the creation of typical scenes in the environment but also the generation of distributions for typical robot behaviors. This way the distribution of training data can be tuned to specific kinds of environments and tasks.

Another ROBOSHERLOCK perception expert is the detector for objects of daily use in cluttered scenes that has been designed, developed and investigated in EASE Subproject R02. This detector is a novel and competitive CNN- and image-based object recognition and pose estimation method that is particularly suited for autonomous object manipulation tasks because it not only returns the detected object and its estimated pose but it also returns a self-estimation of its predicted pose's uncertainty (Richter-Klug and Frese, 2019).

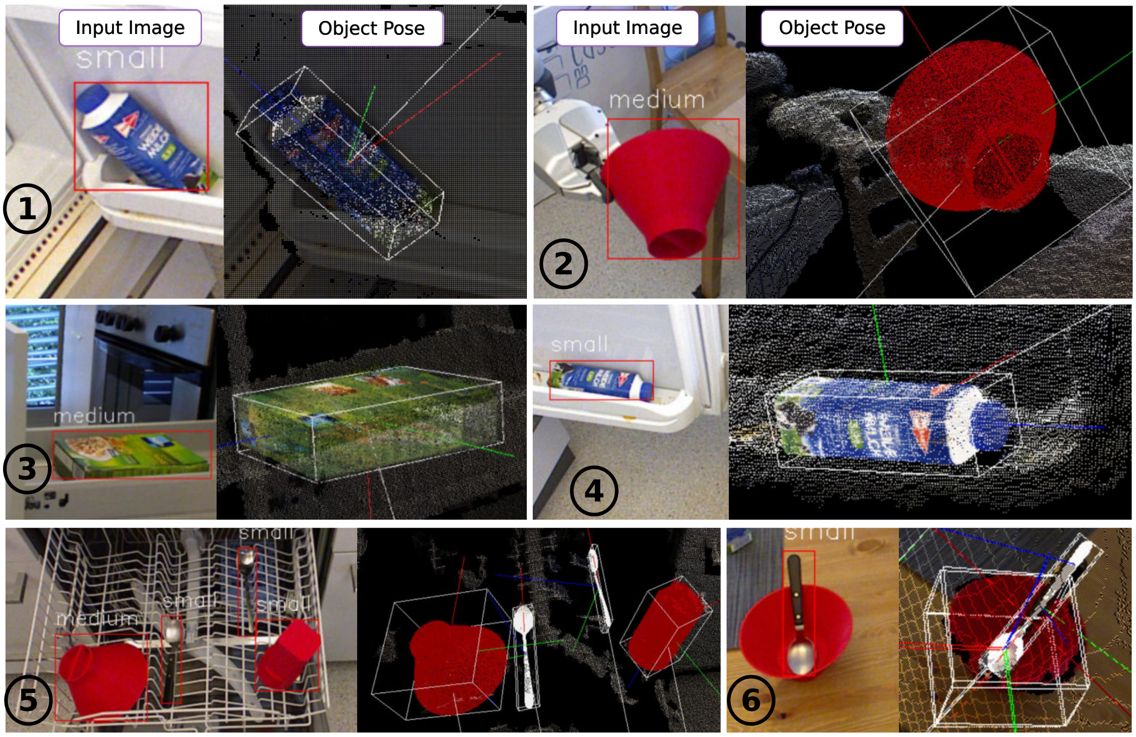


Figure 1.18: Deep network-based object detection and object pose estimation in challenging configurations and contexts in the EASE robot day challenges.

Figure 1.18 shows examples of object detections performed by ROBOSHERLOCK in the EASE robot household challenge. The system is able to detect known objects in a range of relevant configurations, such as spoons in bowls, bowls on plates, milk containers in physically possible but partly occluded poses in drawers, dishwasher, cupboards, and the refrigerator. The object detector was learned in a supervised manner from probability distributions over plausible environment-specific object configurations that were randomly generated in the inner-world model of the knowledge base.

Much of the development of the ROBOSHERLOCK framework itself was not carried out within the EASE project but rather with other funding of EASE researchers mainly coming from EU Horizon 2020 and BMBF projects. In synergy with these projects, ROBOSHERLOCK could provide the perception functionality required for the EASE robot days with sufficient robustness and efficiency (see Section 1.2.2.2). Parts of the results and contributions that have been accomplished for the ROBOSHERLOCK framework are well-documented and explained in the dissertation thesis of Balint-Benczedi (2020).

1.2.3.2.2.4 Action executive

EASE has designed and realized a software framework, GISKARD, to investigate semantic constraint- and optimization-based motion control for manipulation actions. The framework computes adequate body motions for general, underdetermined tasks. GISKARD can be tasked with motion goals and objectives, such as "keep holding a door handle and move the handle according to the articulation model that the handle is part of". These two motion objectives are sufficient to open and close all containers of the kitchen furniture (see Figure 1.19): the oven, the dishwasher, the refridgerator, the drawers, and the cupboards. GISKARD can even execute these motion objectives for different robots. Therefore, the GISKARD action executive is a key component for enabling generalized behavior generation for robot agents. The advantage of this approach to motion generation is that it can compute promising body motion candidates even for novel motion generation problems.

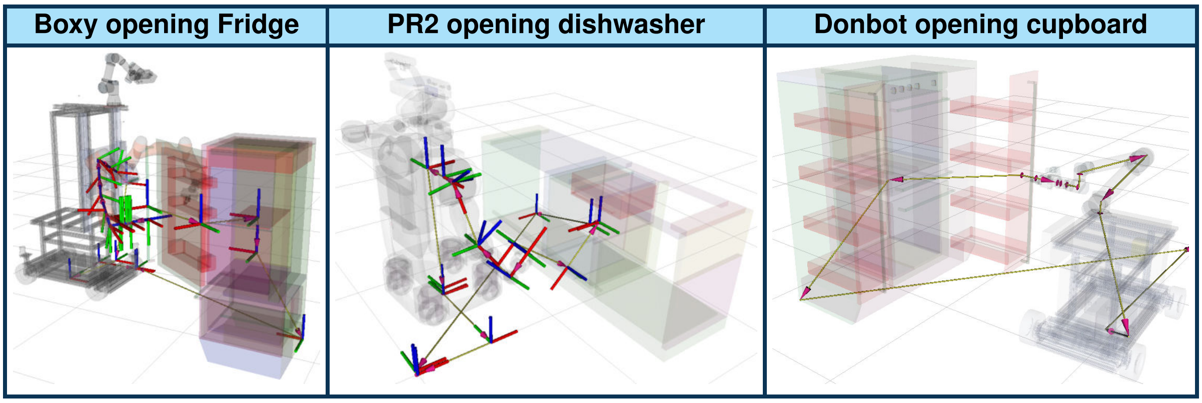


Figure 1.19: Three different robots executing the action request "open the container in which the object to be placed on the table is" for different objects located in different containers.

As constraint- and optimization-based control is a mathematical optimization problem, GISKARD transforms object-based action and motion specifications into mathematically formalized motion tasks.

Tan et al. (2019) have improved the speed of collision detection and penetration computations, which are workhorses in prospection-enabled action execution, by an order of magnitude and achieved faster than realtime performance, so that robot agents can perform realistic plan simulations fast.

While the optimization-based motion generation can compute good body motions, it does so based on idealized and abstract models of the robot capabilities. In many situations the model assumptions are not satisfied. This happens, for example, if the robot moves different body parts at the same time and the composed motion causes inaccuracies in hand motions that are too large for grasping objects successfully, or the motion generation does not take into account the inaccuracies of the estimation of the robot pose.

Such motion generation problems can often be better approached through experience-based learning, for example, by learning manipulation strategies using reinforcement learning. This has been investigated in EASE Subproject R05 for learning hand manipulation strategies for how to open and close containers using tactile-based manipulation procedures, which are linked to the declarative aspects of the developed tactile state detections (Meier et al., 2020). Such processes include movement guidance, tactile servoing, tactile exploration w.r.t. shape or moveability, and different forms of task-related force and touch-based control, e.g. when unscrewing a lid or cutting a piece of bread. At the more abstract semantic level, the procedures need to include concepts, which cover processes that extend over a range of several seconds, while incorporating their declarative and procedural aspects in a low dimensional PEAM representation (see Section 1.2.3.2.2.5 for the definition of PEAMs).

Leidner (2019b); Leidner et al. (2019), throughout their research in EASE, have proposed a powerful cognition-enabled framework for wiping, which allows to realize different wiping tasks such as collecting dirt and brushing off as well as the context-specific refinement of wiping, e.g. through the planning of wiping trajectories. This line of research provides evidence that generalized action plans can also be proposed for other categories of manipulation tasks, in this case wiping.

1.2.3.2.2.5 Metacognition

The generative model developed in the first funding phase also includes fundamental metacognitive capabilities. Metacognition is cognition about cognition and includes aspects such as reasoning about reasoning and knowing about knowing and the use of these capabilities to act better and more robustly. As Brachman (2002) in its seminal article on "systems that know what they are doing" states: "The fact that reflective systems can stop what they are doing and, by stepping back from the situation, possibly get themselves out of a mental box is the reason we believe there is so much promise here."

The metacognition capabilities that are essential for mastering everyday activities by constructing and using generative models include introspection and meta-reasoning, as well as mechanisms for self-programming. Since the latter is a precondition for metacognition because it provides the means by which the cognitive system adapts its own cognitive mechanisms, we will discuss this aspect of metacognition first.

The self-programming precondition of metacognition in the EASE generative model is satisfied because the CRAM plan language is an extension of the Lisp programming language. There are two properties of the Lisp language that facilitate metacognitive capabilities: (1) programs as data and (2) the existence of metacircular interpreters. One powerful idea in Lisp is that Lisp programs are represented as nested lists, that is as Lisp data structures. This means that Lisp programs can inspect and modify themselves: they become plans. The second one is the idea of metacircular interpretation that is that an interpreter for a programming language can be implemented in the language itself. This metacircular interpretation process can be used to make the interpretation of a robot control program explicit and represent it for introspection and metareasoning.

Introspective reasoning enables robot agents to answer questions regarding to *why* the robot made certain decisions, *why* it holds certain beliefs, and *why* it believes that certain physical events occurred. We adopt Feynman's view that answering "why" only makes sense relative to a body of knowledge that is accepted as being true. Without such an asserted truth asking repeatedly "why" is a process that does not terminate and is sometimes even circular. In EASE this accepted body of knowledge is the SOMA ontology: the models of the environment, the body of the robot, the data structures and computational processes, and even the control program itself, along with the body motions and their physical effects are all formalized in the ontology (see Section 1.2.3.5 for more information on the SOMA ontology). The fact that these assertions are logical axioms lets the robot agent automatically infer all the implicit knowledge implied by the ontology and use this knowledge for introspection.

Meta-reasoning leverages an important aspect in recording NEEMs. Specificially, it leverages the fact that the robot agent segments motions into submotions, decomposes the interpretation of plans into the interpretations of subplans, and asserts relations between them. For example, the robot agent might assert that a motion phase of the motion plan has generated an episode of body motion, and that this episode of body motion has caused a change in the environment. This representational structure thus allows the robot agent to identify the subplan that is responsible for the outcome of an action, e.g. opening a drawer. The ability to make such inferences — to map from things that have happened to the process that caused them to happen — is the key to making targeted changes to the plan and allowing sophisticated self-programming to improve the robot agent's cognitive abilities.

Besides laying these essential foundations of metacognition in the CRAM cognitive architecture, we have also proposed the first limited realization of transformational learning in order to revise generalized plans to change the activity structure for specific task and context variations (Kazhoyan et al., 2020b). The long term vision of EASE is to exploit transformational learning and self-programming in two complementary ways: by specialization through PEAMs (pragmatic everyday activity manifolds) and by generalization through metacognitive induction. Both approaches are discussed in the following two paragraphs.

Beneath the familiarity of everyday activities often lies a complexity that can be computationally intractable, especially when you factor in the flexibility that humans exhibit when carrying out these activities, the variety of circumstances in which they carry them out, and the underdetermined manner in which they are described: all key concerns of EASE. This complexity is characterized by the fact that the mapping encapsulated in the EASE generative model is embedded in a very high dimensional space. The mapping is from a vaguely-stated high-level goal to the specific low-level motion parameter values required to accomplish the action successfully. **One of the central ideas of EASE is that, for everyday activities, the generative model does not need to capture all the dimensions of this space: subsets of these dimensions is often sufficient to accomplish the actions successfully.** These subsets are manifolds, specifically PEAMs (pragmatic everyday activity manifolds), and they serve to render tractable the solution of problems that in their full generality are intractable, through specialization. They do this by identifying the constraints that knowledge of everyday activities and the environment bring to bear on the problem.[33] One of the main goals of EASE is to identify and exploit the PEAMs that will result in a robot agent mastering everyday activities.

D

PEAM

PEAM stands for pragmatic everyday activity manifold. This manifold is a subset of the full dimensionality of the mapping from a vaguely-stated high-level goal to the specific low-level motion parameter values required to accomplish the action successfully. Specifically, it is the subset that is sufficient to accomplish certain actions successfully. By operating in a subspace manifold, a PEAM renders tractable the solution of problems that in their full generality are intractable. A PEAM, therefore, represents a form of specialization.

Generalization through metacognitive induction complements the PEAM solution strategy by exploring patterns among generalized actions plans, seeking ways to transform generalized action plans, either by carrying out the action in a more efficient and effective manner, or by accomplishing the outcome of the action in a different way. For example, instead of depending on a generalized action plan to carry dishes one by one to the dishwasher, a more general plan might first stack them if, as in the case of plates, they are stackable, then carry them together, and transfer them from the stack into the dishwasher. Alternatively, if they are not stackable, they might be placed on a tray, carried, and transferred to the dishwasher from the tray. EASE proposes to explore the ways in which this form of induction and transformational learning can be embedded in the CRAM metacognition system.

[33] The idea of exploiting subspace manifolds to render an otherwise intractable problem tractable has parallels in other related domains. For example, in the context of dynamical systems, Schöner (2009) argues that it is possible for a dynamical system model to capture the behaviour of a very high dimensional connectionist system using a small number of variables because the macroscopic states of high-dimensional dynamics and their long-term evolution are captured by the dynamics in that part of the space where instabilities occur, known as the low-dimensional center-manifold.

1.2.3.3 Key result 1: Generalized action plans in the CRAM Plan Language (CPL) Executive

We have proposed, realized, and empirically investigated a novel generative model for mastering everyday activity, in which a robot agent is equipped with a single compact and generalized action plan[34] for fetch&place tasks. This plan generates the variety of behaviors necessary to accomplish complete EASE robot days including setting the table, cleaning the table, loading and unloading the dishwasher for breakfast, lunch, and dinner (Kazhoyan and Beetz, 2019a). A three-minute video[35] demonstrates a robot successfully performing these activities. In this generalized action plan, every context-specific behavior is automatically inferred by the robot based on the object acted on, task context, geometric and physical scene constraints, and robot capabilities. The experiment shows that the generative model detailed in Section 1.2.3.1 is sufficient for generating the appropriate flexible, context-sensitive behavior for transporting a variety of objects from different places to the dining table and cleaning the table afterwards.

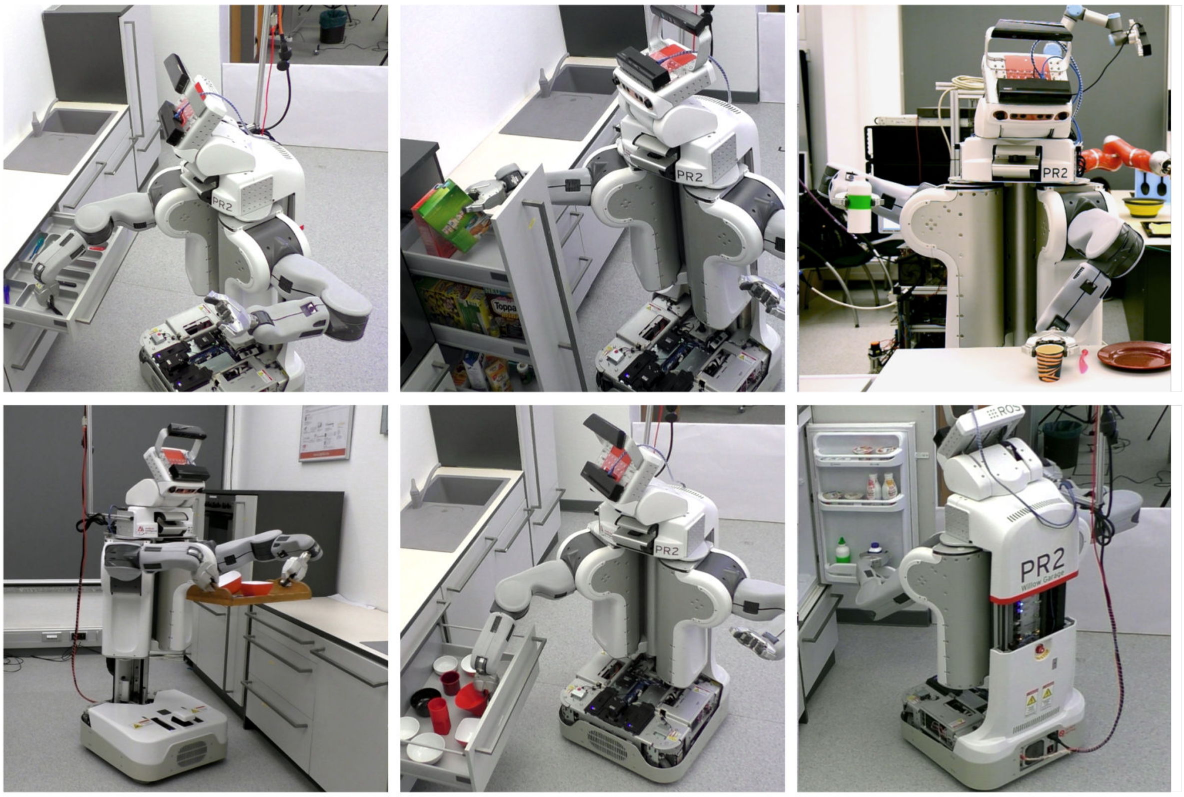


Figure 1.20: Different object grasps selected by the generative model based on the object, task, and context.

Figure 1.20 shows some examples of variations of grasping objects in the context of the EASE robot days, which are all inferred from the robot's knowledge base: the spoon in the drawer is grasped from the top because it is a very flat object; the tray is grasped with two hands because the center of mass would be too far outside the hand for a single-hand grasp; the mug is grasped from the side and not at the rim because the purpose of grasping it is pouring liquid into the mug. The fetch&place plan is sufficiently general that it can be executed with different robot bodies, and so that it works on other objects and in different environments.

[34] In EASE we consider a robot plan to be a robot control program that can not only be executed but also reasoned about and modified at execution time (McDermott, 1992).
[35] `https://www.ease-crc.org/link/video-ease-robot-day`

We have investigated the hypothesis that the performance of robot agents can be substantially improved by adding knowledge to the generative model as well as extending the reasoning mechanisms and learning. Recall that the underdetermined action descriptions in the EASE generative model are implemented as probability distributions over action instances (see Section 1.2.3.1.2). Thus, the expected performance of the generative model depends on the likelihood of the success of an action instance drawn from the distribution for accomplishing the respective request.

Even without knowledge, the generative model has a chance of achieving the respective request because it samples in the space of possible motion parameterizations. However, motion parameterizations that achieve the requests are extremely sparse. In addition, there is a high probability that the request can become unachievable because of unwanted side effects such as objects being broken or pushed out of reach. We call this generative model the uninformed one ($gm_{uninformed}$). As a consequence, this approach is not effective.

Kazhoyan and Beetz (2017) have proposed a knowledge base with hand-coded heuristic rules encoding commonsense, intuitive physics, and other useful knowledge. Examples of heuristic rules are that in order to position yourself to detect an object you have to select a position from where the object is visible or you should first look for objects at places where you believe them to be. This knowledge base has proven to be effective but causes a substantial amount of backtracking.[36] A positive aspect of the heuristic rules included in the knowledge base is that many of the rules are applicable to other kinds of actions such as pouring, wiping, or cutting, too.

We call the generative model using the knowledge base consisting of heuristic rules the *elementary proficiency level* of the model (gm_{epl}) and use it as the baseline performance for other variants of the generative model. We have investigated four extensions of gm_{epl} by augmenting it with (1) prospective capablities $gm_{prospective}$, (2) experience-based learning capabilities $gm_{experience}$, (3) learning from observation capabilities $gm_{imitation}$, and (4) transformational learning and planning capabilities $gm_{transform}$.

The $gm_{prospective}$ uses the temporal projection of the intended action plan as an additional resource for selecting a promising motion parameterization (Kazhoyan and Beetz, 2019b). Using plan projection the robot agent can, for example, predict whether an intended grasp also allows the robot agent to place the object at the intended location. Projection reduces backtracking but can also cause delays in action execution due to the time required to run multiple instances of the simulation before making a decision.

The $gm_{experience}$ generative model records all queries for motion parameterizations, the returned parameterization, and whether the parameterized motion was successful in achieving the action goal (Koralewski et al., 2019). These experience data are then used to learn how to parameterize the motions to maximize the probability of success and the expected performance. Experience-based learning has the effect that the probability distributions that implement underdetermined action descriptions become more peaked and narrower. This means that the information content in the distributions is substantially increased and thereby the performance of $gm_{experience}$ over gm_{epl} is significantly improved.

The $gm_{imitation}$ generative model aims at performing the same learning tasks as $gm_{experience}$ but creating the training data by observing humans operating in a virtual reality environment rather than having the robot collect its own physical experiences (Kazhoyan et al., 2020a). The model learns much faster because humans use their commonsense and intuitive physics knowledge when they generate training data. A complication is that humans are much more dexterous than robots and the solutions have to be transformed from the human body to the robot body to become executable. Again, we could show that $gm_{imitation}$ achieves a significant performance improvement over gm_{epl}.

[36] The backtracking behavior is caused by the robot retrying to perform an action after its execution fails. The most common reason for this is the robot not being able to generate a collision-free trajectory for its arm to reach the goal. In that case the robot repositions its base and retries the action. The more confined is the space, where the object is located, the more challenging it is to find a collision-free trajectory for reaching it. For example, in the case where the action parameters, including the poses for the robot base, are inferred from the heuristics-based generative model, the robot repositions itself on average 3.6 times when grasping a milk box out of the fridge.

The $gm_{transform}$ model is able to transform the motion plan schema by reordering motion phases and replacing specific motion phases with different ones (Kazhoyan et al., 2020b). These plan transformations provide different ways of achieving behavior goals such as closing a drawer by pushing it with the elbow or a door with the foot instead of using the hand. Plan transformation sometimes opens up new optimization possibilities or makes actions achievable by applying a trick.

Bozcuoglu et al. (2018) propose an approach to adapt and optimize CRAM plans based on abstract knowledge provided by other robots.

The knowledge that the robot agent acquires in order to improve its capability in accomplishing everyday manipulation tasks includes (a) the factorization of the possible contexts into categories of contexts that require specific behavior patterns, (b) the optimization of the behaviors by tailoring them to the respective contexts, and (c) the factorization and generalization of knowledge such that it is composable and applicable to novel tasks and contexts.

Model	Comments	Publication & dataset	Media
$gm_{uninformed}$	Theoretically works but not effective.		
gm_{epl}	Works effectively but needs substantial backtracking.	(Kazhoyan and Beetz, 2017)	video[37]
$gm_{prospective}$	Significantly better than gm_{epl}. Might delay execution.	(Kazhoyan and Beetz, 2019b)	video[38]
$gm_{experience}$	Significantly better than gm_{epl}. Major resources required for experience acquisition.	(Koralewski et al., 2019) dataset[39]	video[40]
$gm_{imitation}$	Significantly better than gm_{epl}. Relatively minor resources for experience acquisition but requires adaptation to robot body.	(Kazhoyan et al., 2020a) dataset[41]	video[42]
$gm_{transform}$	Tailors behaviors to specific contexts. Typically little generality but high gain.	(Kazhoyan et al., 2020b)	video[43]

Figure 1.21: Research activities demonstrating that the capability of accomplishing everyday activities can be improved by adding prospection, by acquiring experience, by imitation learning, and by plan transformation. The baseline performance is given by gm_{epl}, the elementary proficiency level.

The results on generative models including references and video demonstrations are summarized in Table 1.21. *We are not aware of any other research initiative that demonstrates such a general, flexible, and context-guided accomplishment of human-scale manipulation tasks under such realistic circumstances.* The respective experiments and their results are accessible as open research, including the open-source plans and the knowledge bases[44] as well as the videos of the experiments and complete recordings of experiment data as NEEMs represented in KnowRob that can be further analyzed interactively through the web-based knowledge service openEASE (see the links in the footnotes of Table 1.21).

[37] http://ease-crc.org/link/video-action-descriptions
[38] http://ease-crc.org/link/video-prospection
[39] https://neemgit.informatik.uni-bremen.de/raw/iros-2019-plan-specialization
[40] http://ease-crc.org/link/video-plan-specialization
[41] https://neemgit.informatik.uni-bremen.de/raw/iros-2020-imitation-learning
[42] http://ease-crc.org/link/video-imitation-from-vr
[43] http://ease-crc.org/link/video-plan-transformations
[44] https://github.com/cram2/cram

1.2.3.4 Key result 2: The symbolic knowledge representation & reasoning framework: KnowRob

A key factor of the competence of robot agents is the knowledge and reasoning capabilities they have in order to contextualize underdetermined action descriptions as well as the introspective capabilities they have to answer queries, including ones about what they do, why they do it, how they do it, and what they expect to happen when they do it.

We structure the knowledge needed for accomplishing manipulation tasks into four categories:

- Knowledge about the environment state, which we call the **belief state** of the robot agent.
- Knowledge about ongoing and previously performed activities, which we call **narrative-enabled episodic memories (NEEMs)**.
- Implicit knowledge that is contained in the data structures and computer programs used by the robot control system, which we call **virtual knowledge bases**. For example, in order to decide whether the robot agent can reach an object, it can call a computation process to determine angles for each body joint such that the robot hand pose is the same as the object pose, that is the application of the inverse kinematics.
- **Generalized knowledge**, including commonsense, intuitive physics, and encyclopedic knowledge that enables robot agents to handle novel tasks and situations. For example, the knowledge chunk that containers that are full should be held such that their content is not spilled applies to all containers, even the ones that the robot does not know yet.

1.2.3.4.1 KnowRob2.0 In the first phase of EASE we have proposed KnowRob2.0, a second generation KR&R framework for robot agents. KnowRob2.0 (Beetz et al., 2018) is an extension and partial redesign of KnowRob (Tenorth and Beetz, 2013, 2015) that provides the KR&R mechanisms needed to make informed decisions about how to parameterize motions in order to accomplish manipulation tasks. The extensions and new capabilities include highly detailed symbolic/subsymbolic models of environments and robot experiences, visual reasoning, and simulation-based reasoning. Aspects of redesign include the provision of an interface layer that unifies heterogeneous representations through a uniform entity-centered logic-based knowledge query and retrieval language.

In addition, KnowRob2.0 is designed to leverage concepts and results from motor cognition and robot control to extend AI reasoning into the motion level and make the robot's reasoning mechanisms more powerful. The new capabilities and functionalities are facilitated through employing modern information processing technologies such as physics simulation and rendering mechanisms of virtual reality engines, big data recording, storage, and retrieval technologies, and machine learning. The use of the above leading-edge information technology enables robot agents to acquire generalized commonsense and intuitive physics knowledge needed for the mastery of human-scale manipulation tasks from experience and observation and to make AI reasoning actionable within the perception-action loops of robots.

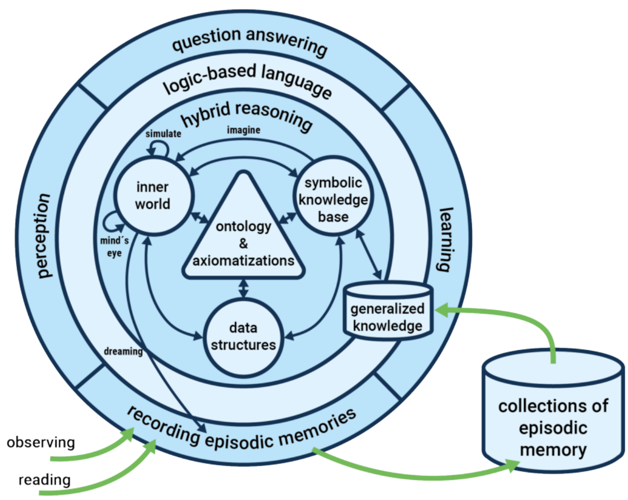


Figure 1.22: Software architecture of KnowRob2.0.

The architecture of KnowRob2.0 itself is depicted in Figure 1.22. A unique feature of it is the central position of the symbolic representation of the ontology, even below the data structures of the control system. This enables the programmer to semantically annotate and lets the control system automatically compute the semantic meaning of data structures.

Around the ontology is the hybrid reasoning shell. Many of the data structures, representations, parameters of computational processes have associated axiomatizations that declare their meaning with respect to the ontology, allowing the robot control system to use this data as if it was a symbolic knowledge base. The hybrid reasoning shell uses multiple methods for knowledge implementation. Key components are the data structures of the control system and robotics algorithms such as inverse kinematics and motion planning, which, for example, allow the programmer to specify that the robot should believe an object to be reachable if the motion planner can find a collision free path to the position of the object.

Another key and novel component of the hybrid reasoning shell is the inner world knowledge base (Billing et al., 2016), which is detailed in Section 1.2.3.4.2. It is a detailed, and photo-realistic reconstruction of the robot's environment in a game engine with physics simulation and vision capabilities, and adds powerful reasoning methods to the KnowRob2.0 knowledge processing framework. First, the robot can geometrically reason about a scene by virtually looking at it using the vision capability provided by the game engine, and predict the effects of actions through semantic annotations of force dynamic events monitored in its physics simulation. As Winston (2012) would formulate it, it allows the robot to reason with its eyes and hands.

The subsequent interface layer exposes reasoning capabilities of control mechanisms integrated below it through a logic-based language. The language exploits control-level data structures for ad-hoc symbol grounding, and ontologies for unifying these heterogeneous representations. To applications above the interface layer, the hybrid reasoning shell appears to be a first-order logic knowledge base, but it is largely constructed on demand from data structures of the control program, and computed through robotics algorithms.

Finally, the interface shell provides the question answering, perception interface, experience acquisition, and knowledge learning interface of KnowRob2.0 that can exploit the rich set of hybrid reasoning mechanisms integrated below the interface layer.

The publication on KnowRob2.0 (Beetz et al., 2018) was included in the list of most important publications of the IEEE RAS Technical Committee on Cognitive Robotics area in the years 2017-2019.

1.2.3.4.2 Inner world/digital twin knowledge representation and reasoning A distinct reasoning capability of KnowRob2.0 is called *digital twin knowledge representation and reasoning (*DTKR&R*)*. DTKR&R is designed to leverage the advantages of machine-understandability and valid reasoning provided by symbolic reasoning frameworks with the level of detail and the groundedness in perception and action that is required for robot control. DTKR&R can enable robot agents to autonomously accomplish underdetermined manipulation tasks because it provides semantic knowledge as well as geometric information and coordinates. It also provides powerful cognitive capabilities such as mental simulation, mental imagery, learning by dreaming, activity interpretation, and imitation learning.

DTKR&R, which is depicted in Figure 1.23, is a hybrid representation and reasoning framework composed of a symbolic knowledge base (KB) and an artificial virtual world (AW) that represents the detailed, geometric and physical model of the world together with its visual appearance. The two knowledge bases are strongly coupled. Each symbolic name of an object or an object part is an identifier of the data structure in the artificial world that implements the respective object and object part. The abstract spatial and physical relations in the knowledge base are abstractions of the respective detailed state of the artificial world.

DTKR&R is made possible by modern game technology, in which huge realistic environments are represented very comprehensively and very detailed and physically simulated and rendered in real time and with currently available computing power. For example, already in 2015 the Unreal game engine[45] showed the technology demonstration “a boy and his kite” featured realistic physics-based rendering of a 100 square miles environment with 15 million pieces of vegetation, realistically moving animals, and physics-enabled animation of a kite flying in changing wind conditions in real time at 30fps.

The key idea of DTKR&R is to take the scene graph data structures that are the implementation basis of the virtual environments also as the implementation basis for the symbolic knowledge representation system. To do so, we extend the data structures in the scene graph that implement an entity, be it an object, object part, an articulation model, or a subscene, that is relevant for the robot agent with a symbolic name. This symbolic name is then axiomatized in the symbolic knowledge base by asserting it as an instance of an object category defined in the ontological knowledge base and providing formally stated background knowledge about the entity. The relationship between the symbolic knowledge base KB and the artificial world AW is then that every relavant physical entity of AW is formalized in KB and every symbolic physical entity is also an entity in AW.

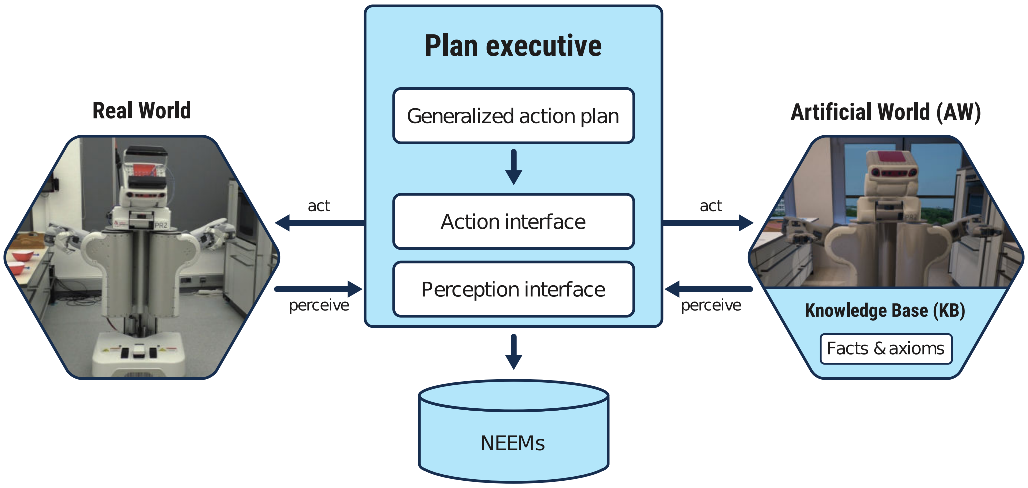


Figure 1.23: The key idea is that the artificial world AW of the knowledge base matches the real world so closely that the robot control program can be executed in AW in the same way as in the real world: the NEEMs generated in the real world cannot be distinguished from those generated in AW.

Because scene graph representations are used by simulation and rendering engines to realistically visualize actions and events in the virtual environment they represent, we can use the simulation and rendering engines as additional reasoning engines.

Ideally, a robot agent employing DTKR&R aims at acquiring and maintaining an artificial world AW that matches the real world so closely that by just looking at the interactions recorded in the respective NEEM of the robot agent with its environment one could not tell whether a robot action was executed in the real or artificial world. If this correspondence holds then we could make the robot agent believe that it is acting in AW and at the same time physically acting in the real world. The advantage of doing so is that for the robot agent acting in AW the robot agent has perfect information about the world, access to all necessary details, and the grounding of symbols is verifiable. Of course, AW will never perfectly match the real world so any robot control program that believes to act in AW must be equipped with a heavy machinery of event and failure detection, diagnosis, and recovery and continously updating the AW to match the real world.

[45] https://www.unrealengine.com

1.2.3.4.3 Dynamic belief states

Our generative model represents the belief state of a robot agent using DTKR&R as an artificial world, which can be visually rendered and physically simulated. Thus to assess the information content of the belief state at a time instant *ti* during an everyday activity episode we can capture an image from the artificial world at time instant *ti* and compare it with an image captured by a real camera in the real environment. Figure 1.24 shows an example of an image generated by this process and the comparison with the real state of the environment at the same time instant.

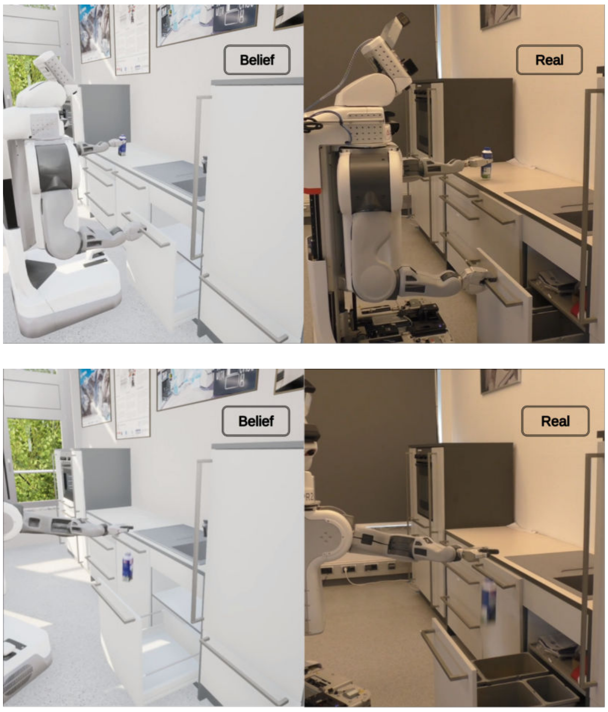


Figure 1.24: Comparison of the belief state and the real state of the environment in table setting episodes. The images on the left are renderings of the symbolic belief state of the robot agent.

This way we can qualitatively judge the comprehensiveness of the belief state and its accuracy. In terms of comprehensiveness, it matters whether the belief state contains all relevant objects and whether the scene representations are sufficiently rich for contextualizing object manipulation. The first criterion for sufficient accuracy is that the parameters for the motions that implement object manipulation tasks can be inferred sufficiently accurately so that the manipulation actions will succeed. As the belief state has a subsymbolic rendering, which is very similar to the images captured by robot cameras, it supports the grounding of symbols in the perception-action loop. The second criterion concerns the accessibility of the knowledge in the belief state. In particular, it concerns the ability of robot agents to infer answers to the body motion queries based on the belief state. For example, in order to fetch a spoon in the context of setting a table, the robot should be able to infer answers from its belief state for questions including the following ones: "Where can I find spoons that are clean and unused?" "How can the container, which the robot believes the spoons to be in, be opened and closed?" "How should the drawer be grasped to open it?" "What is the handle of the spoon by which it should be grasped?"

Knowledge in the belief state can be retrieved through asking queries. For example, the robot could ask queries such as: "which containers open counterclockwise?", "which objects are electrical devices?", or "what is a storage space for perishable items?". The refrigerator is an answer to all of these queries. If the robot then asks for knowledge about the refrigerator, the accessible knowledge would include the part hierarchy of the refrigerator, including the 3D models of the parts, the articulation model of the refrigerator that tells the robot how to open and close it, and so on. Part of this background knowledge is visualized in Figure 1.25.

Note that in order to obtain this question answering capability we make a *weak closed world assumption*; that is that the robots know almost all objects and entities in the environment. The knowledge base of the robot is populated with object models that consist of CAD models, including the part structure and possible articulation models, a texture model, as well as encyclopedic, commonsense, and intuitive physics knowledge about the object. We call the closed-world assumption *weak* because the robot is still required to detect novel objects, for example if the robot unpacks a shopping bag or somebody else put a novel object in the environment.

Under the *weak closed world assumption*, computing the belief state comes down to maintaining a belief about where each object is and which state it has. Domain knowledge can be provided as prior

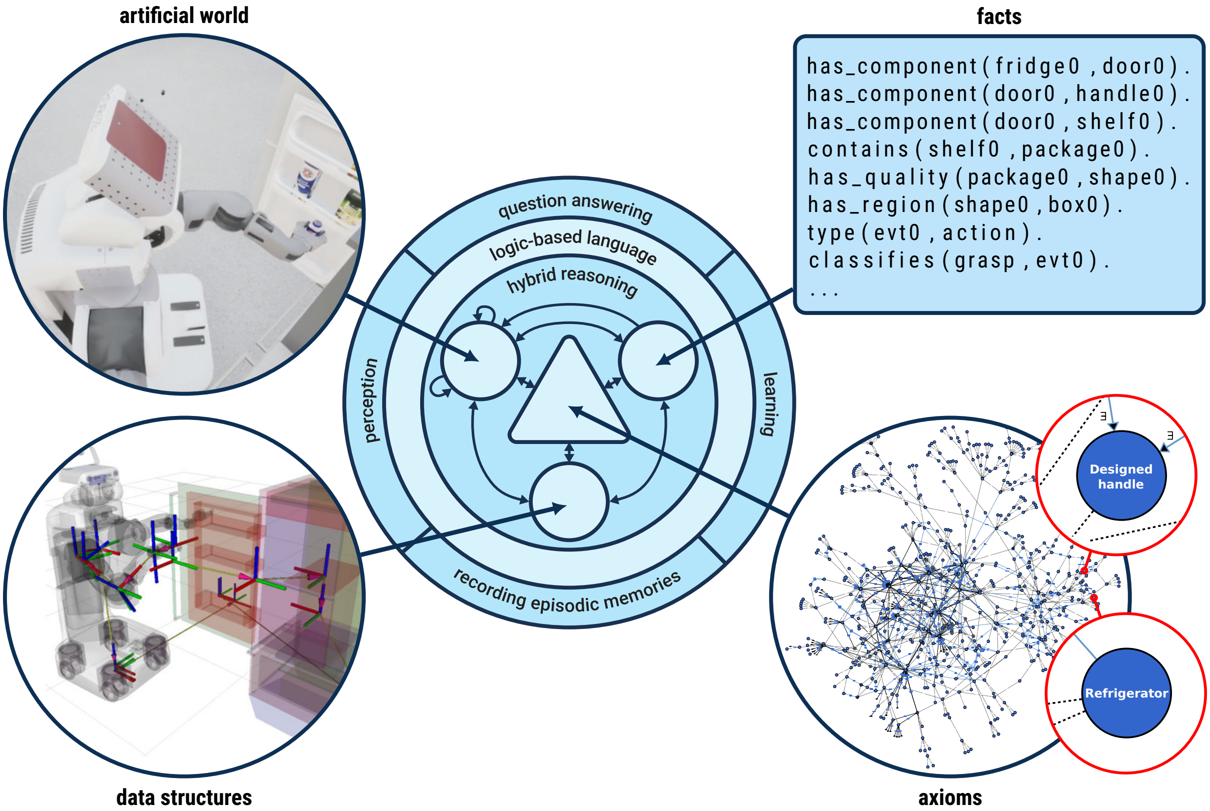


Figure 1.25: Model and knowledge base representing the refrigerator in the belief state of the robot agent.

knowledge through a hand-coded ontology, which contains valuable manipulation knowledge, such as a container can be opened by generating a motion implied by the articulation model of the container (e.g. the knowledge that a screw top cap can be removed by twisting the cap).

Assuming a weak closed world, robot agents in EASE can also answer queries that require prospective capabilities. An example of such a query is: “what do I expect the inside of the refrigerator to look like when I open it?” Answering this query requires the robot agent to visually render the scene inside the refrigerator given its current belief state. Another example is: “what do I expect to happen if I pick up the object in front of me with my right gripper?”

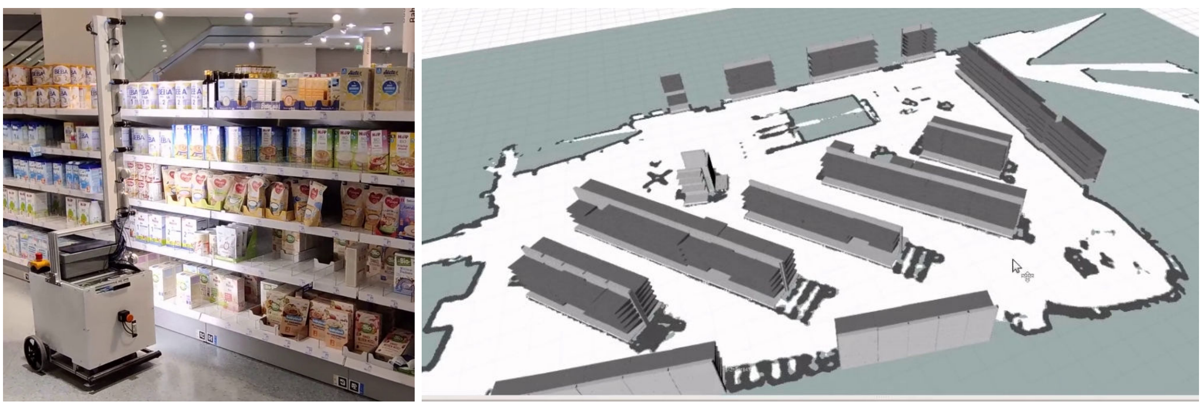


Figure 1.26: Acquiring environment models of retail stores.

A possible barrier for robot agents maintaining such expressive and detailed belief states is the automatic acquisition of the necessary prior models and knowledge when robot agents are deployed in novel environments. In the H2020 Refills project, which exploits the EASE knowledge representation, plan-based control, and perception technology, we have shown that robot agents can autonomously acquire environment models of more structured environments, in this case models of retail stores, which have the comprehensiveness and level of detail of the belief states of EASE robots.[46] These results are evidence that the acquisition of such detailed and knowledge-rich environment models might also become feasible for ordinary human living environments such as kitchens and apartments. This line of research lies, however, outside the scope of EASE.

1.2.3.4.4 Prospection Prospection, the ability to represent what might happen in the future, is an essential cognitive capability that enables robot agents to accomplish tasks effectively by anticipating and taking into account the predicted effects of their actions. Prospective capabilities include simulation, prediction, intention formation, and planning. In the first funding period, we have restricted our focus on individual prospection capabilities, while a key focus of the second phase will be a comprehensive KR&R framework for prospective reasoning.

Within this collection of prospective capabilities, (episodic) simulation plays a central role. Simulation is the process of imagining an action without executing the movements involved. Hesslow (2002) in his simulation theory of cognition argues that a substantial subset of human cognitive capabilities are based on mental simulation. When a human pours pancake mix onto a pancake maker, she can parameterize and adapt the movements such that at the end the pancake mix forms a circular pancake of a certain size. She can also immediately answer questions regarding to what will happen if she holds the pancake mix too high, tilts it too fast, if the pancake is too thick, or too thin. These prospective capabilities enable humans to quickly predict and diagnose the causes of unwanted side effects and adapt the movements to forestall them.

1.2.3.4.4.1 Plan projection One of the prospection methods that we investigated in the first funding phase of EASE is plan projection. Plan projection is an abstract, symbolic method for reasoning about the future execution of generalized action plans. In plan projection mode, the plan interpreter asserts the occurrence of motion phases and the force dynamic events of the Flanagan model as facts in a first-order time interval logic representation (see Section 1.2.3.1.2, specifically Figure 1.9, for more information on the motion phases and the Flanagan model). The concrete parameters of the asserted events and their perceptual and physical events are predicted by context-dependent symbolic rules (Kazhoyan and Beetz, 2019b). An example use of execution time plan projection is shown in Figure 1.27, in which projection is used for finding the best action parameters, whereby the choice of parameters of picking up an object takes into account how the object will be placed at the destination later (see Section 1.2.3.1.2, specifically Figure 1.8, for the overview of the plan syntax).

```
def-plan fetch&place (?object, ?search-location)
  perform (an action
             (type searching)
             (object ?object)
             (location ?search-location))
  with-projected-parameters ?robot-loc-for-fetch,
                            ?arm, ?grasp,
                            ?robot-loc-for-place,
                            ?destination
    with-robot-at-location ?robot-loc-for-fetch
      perform (an action
                 (type fetching)
                 (object ?object)
                 (arm ?arm)
                 (grasp ?grasp))
    with-robot-at-location ?robot-loc-for-place
      perform (an action
                 (type placing)
                 (object ?object)
                 (destination ?destination))
```

Figure 1.27: Example use of plan projection in a fetch&place plan.

[46] A video summarizing this semantic mapping approach can be found at `https://www.ease-crc.org/link/video-semantic-mapping-retail`.

1.2.3.4.4.2 Mental simulation with URoboSim The full power of artificial mental simulation methods for robot agents performing simple manipulation tasks we have investigated by designing, implementing, and empirically evaluating UROBOSIM (see Figure 1.28), a simulation framework for the EASE generative model described in Section 1.2.3.1. Given a hypothetical situation *AW*, specified as a digital twin knowledge base (Section 1.2.3.4.2), a virtual robot also specified as a digital twin knowledge base, and a manipulation task request in the form of an action description, UROBOSIM "embodies" the plan executive into the virtual robot and requests the plan executive to perform the task in the hypothetical situation. Then in each perception-action cycle UROBOSIM renders a camera image for each of the robot's cameras and simulates the physical effects of the generated motions using its physics engine. It also generates and maintains the low-level data structures for the robot control system such as the image data structures and the states of the kinematic tree of the robot such that the perception and action executive can run in simulation mode. The simulation is automatically recorded as a NEEM represented in the KNOWROB representation language.

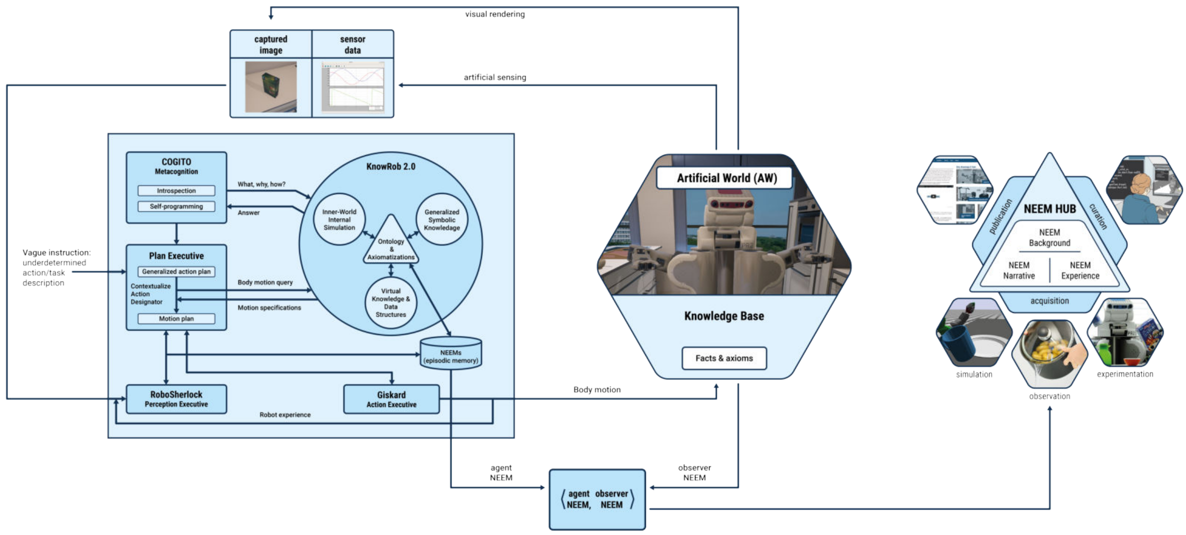


Figure 1.28: UROBOSIM as CRAM component for simulation-based reasoning.

Mental simulation can be used to do execution time motion parameter selection and selecting alternative motion strategies based on simulating action effects and for self-trained experience-based learning.

The level of realism of UROBOSIM was empirically validated by running the simulation passively in parallel with execution and comparing the states. This way simulation can interpolate state parameters that the vision system is not estimating as, for example, the pose of a drawer while pulling it or the position of an object after opening a gripper. The uncertain aspects of simulation of fetch and place actions are (1) how exactly is an object grasped, which is possibly affected by object deformation, weight, object surface properties, (2) how the object moves inside the hand while carrying it, and (3) what physically happens when letting the object go when placing it. As these uncertainties are caused by estimation rather that inaccurate simulation, the most promising fix is to employ manipulation strategies that are robust against the expected uncertainties rather than trying to make the simulation more accurate.

The second criterion is the value of simulation for learning fetch and place. Here, the question is can we learn better motion parameterizations for real-world fetch and place only based on learning data obtained in simulation. The simulation capabilities of UROBOSIM are shown in a video.[47]

Plan simulation and plan projection complement each other well. Plan simulation is based on a very general simulation implementing Newton physics. Because of this generality simulation is appli-

[47] `http://ease-crc.org/link/video-urobosim`

cable to actions that the robot has never done before and still gives reasonable predictions. However, physics simulation requires that all relevant physics parameter are correctly and accurately known to make accurate predictions. Plan projection, on the other hand, requires the set of symbolic prediction rules to be complete with respect to the prospection tasks to be accomplished. As the individual rules are designed to handle specific context categories heuristically, they are much faster and more flexible modeling tools. For example, they can more easily model the interferences between different components of agency such as interferences between the simultaneously executed motions of body parts that might result in not modeled vibrations or the interaction between visual perception and motion control. Another advantage of projection is that robot agents can learn projection rules. For example, they can learn the projection rules from plan execution and thereby learn behavior models that are difficult to implement through analytical computations, or robots can learn projection rules from mental simulations in order to speed up execution time of prospection.

1.2.3.4.4.3 Visual mental imagery Imagistic reasoning or seeing with the mind's eye is another important cognitive capability that can be leveraged to achieve the mastery of everyday manipulation tasks. Visual mental images are representations that produce the experience of seeing in the absence of the usual sensory input (Kosslyn, 2005).

In KnowRob mental imagination is the process of asserting a hypthetical state of the environment as a KnowRob situation and assert a camera position and asking KnowRob the query of how would an image look that I take in this state with this camera. KnowRob then renders the image returns it as an answer to the query. The power of this mental imagination process comes from two properties. First, the imagined image is a rendering of a digital twin knowledge base, which means that one has access to the ground truth data of what is depicted in the image. Second, you can make additional assertions about the situation. For example, one can assert that all objects besides the one of interest has the color black, which reseults in a perfect segmentation of the object in the scene.

One of the main applications of mental imagery is the estimation of the belief state of the robot. In this application we use the current belief state of the robot in order to compute what the robot expects to see in the next perception action cycle. This makes complex scene perception faster, more robust, and more accurate. The robot can use the expectation to focus its visual processing to the regions that do not match its expectations. It can make perception more accurate because it can optimize the pose of a known object that it has a model of typically initializing the optimization processes with poses that are within the convergence radius of the optimization process. It can also improve robustness because the perception process can validate an already existing hypothesis rather than interpreting scenes without rich expectations.

The visual mental imagery capabilities have been used to learn real-world object perception capabilities through self-supervised learning by training them using inner-world models of the world (Mania and Beetz, 2019; Kenfack et al., 2020; Bálint-Benczédi and Beetz, 2018).

1.2.3.4.5 Spatio-temporal reasoning Research in spatio-temporal representation and reasoning within EASE (pursued within Project P03-E) focused on developing methods and tools for computational spatial representation and reasoning about everyday activities from the particular viewpoint of commonsense cognitive robotics (Levesque and Lakemeyer, 2007). In particular, we built on and advanced previous / preliminary work by the participating PIs of Project P03-E in the field of spatial cognition and computation (Bhatt et al., 2013a; Bhatt, 2012; Walega et al., 2015; Bhatt et al., 2011; Bhatt and Loke, 2008; Suchan et al., 2016; Suchan and Bhatt, 2016; Spranger et al., 2016; Bhatt et al., 2013b; Schultheis and Barkowsky, 2011; Schultheis et al., 2014), whilst addressing "space and motion" from a human-centered, cognitive, commonsense formal modelling and computational viewpoint, i.e., space, as it is interpreted within AI (KR, commonsense reasoning), Spatial Cognition and Computation, and more broadly, within Spatial Informatics.

A key highlight of the spatial reasoning research in EASE P03 has been to synergistically address both the "cognitive" and "computational" aspects involved in spatial information processing and interpretation in a range of embodied everyday interaction and control tasks. Here, one of the crucial challenges pursued (from the viewpoint of “spatial cognition and computation”) has been to incorporate cognitive rooted characterisations of spatial information conceptualisation and inference by humans; e.g., a key insight that we build on is the human ability exhibiting strong preferences of how to represent spatial information, efficiently pruning the search space of available options, and more generally, of applying commonsense heuristics in embodied interaction and decision-making through efficient selection and hypothesis building mechanisms (e.g., which object to pick next, what is the most plausible explanation of what happened). It is therefore essential that such heuristics / preferences not only be systematically examined in behavioural settings, but that they also be utilised in the computational cognitive modelling of spatial reasoning for their application in everyday reasoning situations including but not limited to the robotics tasks of core interest to EASE.

From the formal commonsense spatial reasoning viewpoint, our research method categorically marks a departure from previous relational-algebraically rooted “qualitative spatial reasoning” techniques (Ligozat, 2013) that operate purely within a qualitative setup (even when quantitative information is available and / or necessary, as is the case with robotics scenarios). Specifically, commonsense reasoning about space and change in our line of research is pursued in a mixed quantitative-qualitative setting directly within the context of declarative methods such as (constraint) logic programming and answer set programming. For instance, through such declarative formalizations, the grounding of the robot’s dynamic spatial environment within a declarative framework for complex spatio-temporal data abstraction, inference, and query via constraint logic, inductive logic, and answer sets becomes possible.

From the spatial cognition and computation viewpoint our approach rests on the assumption that human mastery of everyday activity is characterized and driven by bounded rationality (Simon, 1955; Jones and Love, 2011; Schurz and Thorn, 2016) mechanisms that aim to minimize physical and mental / computational effort (Hull, 1943; Kool et al., 2010). In particular, representing and reasoning about space plays a crucial role in establishing pragmatic everyday activity manifolds that enable efficient behavior. Against this background, we develop computational cognitive models that realize, exploit, and explain preferences regarding spatial representation (Jeffery et al., 2013; Zwergal et al., 2016; Hinterecker et al., 2018) and strong spatial cognition (Freksa, 2015) to constitute generative models of the processing principles underlying human everyday activity performance.

From the viewpoint of models, algorithms and tools for reasoning about space and motion in the context of everyday activities, one of the core results is a computational framework for: (1). grounding of the robot’s dynamic spatial environment within a declarative framework for complex spatio-temporal data abstraction, inference, and query within constraint logic programming and answer set programming; (2). general methodological foundation for the conceptual, formal, and computational characterization of single and multi-object region & point-based spatio-temporal motion patterns as applicable in the context of everyday activities such as cooking and other household tasks. Cross-domain applicability is demonstrated by grounding and reasoning about other everyday activities such as those involved during autonomous driving, and in digital media domains (Bhatt and Suchan, 2020b; Schultz et al., 2018; Suchan et al., 2019, 2018b; Suchan and Bhatt, 2017b).

Research in the area of spatial cognition and computation has provided several key results regarding the principles that underlie efficient human everyday activity mastery. For one, the efficiency of strong spatial cognition has been found to derive from massive and instantaneous information propagation (Shih et al., 2018; van de Ven et al., 2018; Freksa et al., 2019). Furthermore, preferences of object arrangement and activity organization have been discovered and formalized. A generative cognitive model of preference-based activity organization has been developed and shown to predict human activity organization. This model revealed (a) opportunistic activity planning (Wenzl and Schultheis, 2020a) (b) spatial representation, distance, topology, and strong spatial cognition as important determinants of activity organization (Wenzl and Schultheis, 2020b), and (c) efficient spatial representation of only relevant sub-dimensions (Wenzl and Schultheis, 2020c).

1.2.3.4.6 Interpretation of natural language task requests The interpretation and contextualization of underdetermined task instructions is also a longstanding research problem inestigated in natural language understanding because vagueness allows for efficient communication and less constrained execution. Pomarlan and Bateman (2020) have proposed a pipeline for command understanding employing a semantic parser based on fluid and embodied construction grammar, which analyzes instructional texts, such as commands or recipes, and produces semantic specifications ("semspecs") to represent the constraints on what actions are being requested and what participates in these actions. The output of this processing pipeline are the task requests that the EASE generative model processes.

To this end, Semspecs use concepts from the SOciophysical Model of Activities (SOMA) ontologies as semantic types, which makes available the knowledge of activities encoded in these ontologies to the parsing process. This assists both in disambiguating sentences – though syntactically similar, "pour into the pot" and "cut into pieces" have different meanings – and in identifying which participants are necessary for an action, and therefore which participants have been left unspecificed – in the previous example, "cut into pieces" leaves both instrument and object to be cut unsaid. Some of the underspecification is resolvable by the parser via reference resolution, however other dedicated inference procedures to integrate activity knowledge and contextualize the semantic constraints of the semspec are also necessary.

In addition, Bateman et al. (2019) propose the mental simulation capabilities investigated in EASE as a means to compactly encode knowledge about the physical world. When receiving a command, and selecting an action to perform in response to it, simulation of that action with various parameter values allows judging those parametrizations based on how well the outcomes comport to expectations derived from the command semantics. A sequence of actions can also be simulated, which is useful to understand the physical consequences of a command in the execution context in which it is requested. In particular, the actions of domestic robots should support subsequent actions of humans, which offers a criterion to resolve command ambiguity. For example, if the plates are to be placed on the table so as to eat from them, placing them in a stacked manner is not very supportive of the subsequent eating. Finally, simulation is useful towards a deeper understanding of objects in terms of how they interact with each other and how they can be used, which then enables inferences about why an object might (not) escape containment, or what alternative tool to use when the default is unavailable.

The integration of simulation into a reasoning pipeline is itself challenging: symbolic and qualitative representations, such as semspecs, carry too little information to initialize a simulation; conversely, the numerical data coming from simulation needs qualitative interpretation (Pomarlan and Bateman, 2020). We have tackled this by defining an inference pipeline operating at several layers of abstraction, with appropriate representations for each. The most abstract layers are those of functional relations (relations between objects that constrain behavior) and qualitative spatial relations, represented in defeasible logic. Theories for both these kinds of relations make use of geometric primitive relations, and each such primitive relation is described by generative models to instantiate it and recognize instances of it. Functional relations also make use of expectations, which are formalized as qualitative descriptions of movements conditional on the physical presence of some object. The qualitative movements are themselves described by generative models to recognize instances of such movements coming from simulation. This pipeline architecture allows instantiating an arrangement of objects obeying some qualitative description made in terms of function and relative positioning, simulating the resulting scene, and interpreting the resulting object behaviors based on whether they match expectations or not. Combining the scene instantiation process with exact data about initial object location coming from e.g. robot perception is also possible.

1.2.3.5 Key result 3: The SOMA machine-understandable ontology of all EASE knowledge (and data structures & processes) in KnowRob2.0

The systematic organization and formalization of background knowledge in EASE takes place in the EASE ontology of all EASE knowledge (and data structures & processes), which makes the key concepts, data structures, and processes machine-understandable. By being machine-understandable, we mean that the robot can answer queries that are formulated with the concepts and relations defined in the ontology using symbolic reasoning based on the axiomatizations of relevant EASE concepts in the ontology. EASE employs a collection of ontologies (Bateman et al., 2018a) within a very concise foundational or top-level ontology, called SOMA (Socio-physical model of activities). SOMA is a parsimonious extension of DUL (Dolce Upper Lite) ontology, where additional concepts and relations provide a deeper semantics of autonomous activities, objects, agents, and environments. SOMA has been complemented with various subontologies that provide background knowledge on everyday activity and robot and human activity including axiomatizations of NEEMs,[48] common models of actions, robots, affordances (Beßler et al., 2020b), execution failures (Diab et al., 2019), and so on.

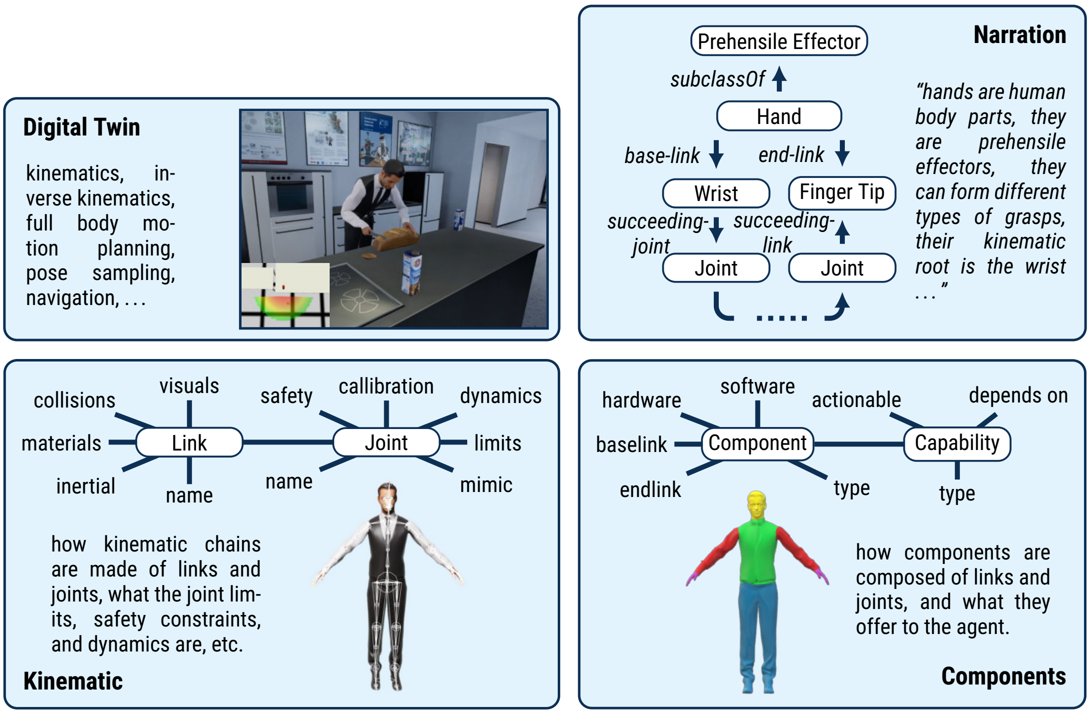


Figure 1.29: Representation of agent kinematics and components in the EASE ontology.

An example is depicted in Figure 1.29, which shows components of the ontology that model physical agents (humans and robots) agents. The ontology defines aspects of physical agents such as their kinematic structure, physical components and their capabilities, beliefs, desires, and intentions, etc. Whenever an instance of the concept physical agent is created the respective background knowledge is inherited, which is a powerful mechanism because every entity in a digital twin knowledge base is asserted to be an instance of a concept in the EASE ontology. Because of this common semantic infrastructure can plan the motions of the human because it can automatically interpret the kinematic structure, more easily compare the actions of humans and robots in NEEMs, build collections of NEEMs that include human as well as robot manipulation actions, and automatically create machine learning problems. The definition and rigorous use of a common ontology opens up a huge spectrum of opportunities for automating learning and reasoning tasks and facilitating metacognitive capabilities.

[48] https://ease-crc.github.io/soma/owl/current/NEEM-Handbook.pdf

If NEEMs are the glue of EASE then the ontology is the superglue.

— David Vernon

1.2.3.5.1 Encyclopedic knowledge bases on robot agency In EASE we attempt to make all relevant data structures, pieces of program code, physical processes, mental processes *machine-understandable*. Machine understandability is achieved by annotating relevant entities with a symbolic name. In the symbolic knowledge base this symbolic name is then asserted to be an instance of the respective concept that is defined in the EASE ontology. The robot agent thereby has access to the background knowledge about this entity. In addition, the entity, be it a piece of a data structure or a computational process, can be accessed through its symbolic name. This implies that for these symbolic names we have direct access to the entities they represent and to the respective subsymbolic representations.

This ontology provides the robot agent and its control program – including such diverse software components as the perception, motion, decision making, and planning components – with a huge common, rigorous encyclopedia for all the data structures that are exchanged and all the computational processes that take place in the control system which is encoded as a first-order logic knowledge base. This is an incredibly powerful idea.

Using this encyclopedia together with a logic interface language for formulating queries and answers to these queries, any component of the system can access any information provided by any other subcomponent. At the same time, each subcomponent can process the data in their own suitable format.

The power of a comprehensive ontology for robot agents and everyday activity cannot be overstated. For example, consider the knowledge preconditions of actions, such as the place from, or the grasp with, which an object can be picked up successfully. Having such a concept the robot agent could automatically refine this concept using the distribution of its experiences. This can be done by extracting from the collection of NEEMs the sub-episodes in which the robot attempted to pick up objects and transform the pair ⟨robot-pose,success/failure⟩ into a data point for learning. Using the data distribution the robot agent can then learn the preimage of poses for which success is predicted. Using this idea we could create action-related robot plan schemata with action-related concepts, such as *the place from which I can pick up an object*, which a robot agent could instantiate using its collected NEEMs when it downloads a new plan.

This is possible because robot agents have a huge advantage over other hybrid symbolic/subsymbolic reasoning systems: they generate and update the subsymbolic data themselves. Therefore they can chunk the data and semantically annotate them in the generation process and thereby already ensure the appropriate structuring of the data as well as their semantic meaning. For example, if the robot reaching for an object to grasp it knows that it can expect a contact force between its grippers and the object, it can also prepare for the increased force needed to lift its hand due to the weight of the object. If this weight increase cannot be observed, then the object was not grasped successfully. Thus, the fact that the sensor stream is interpreted by the pick-up control program makes it possible to automatically infer the force-dynamic events that are needed to model fetch-and-place actions and to segment the action into the respective motion phases.

EASE has substantially contributed to the IEEE-SA P1872.2 Standard for Autonomous Robotics (AuR) Ontology, originated as a sub-group of IEEE WG Ontologies for Robotics and Automation (ORA) (`https://standards.ieee.org/project/1872_2.html`). In this context, Olivares-Alarcos et al. (2019) have compared and reviewed ontology-based approaches to robot autonomy and discussed existing approaches in autonomous robotics that use ontologies to support autonomy. The results of this study are constantly updated and accessible through the web page `https://ease-crc.org/ontology-survey-2019/`.

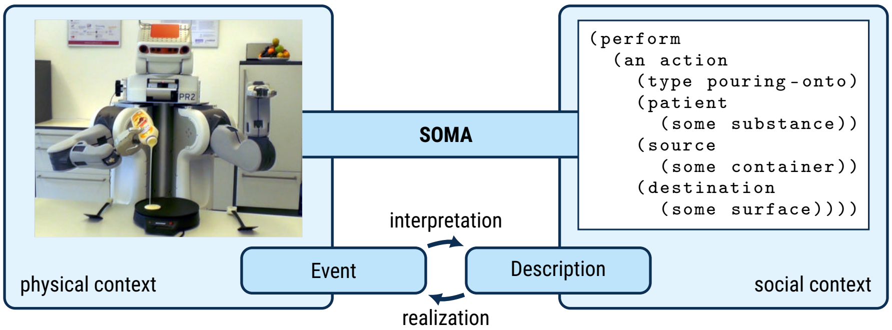


Figure 1.30: Representation of the physical and computational processes and representation involved in the execution of actions in the SOMA ontology.

1.2.3.5.2 SOMA (Socio-physical Model of Activities) ontology One essential achievement is the formalization of the EASE core upper ontology SOMA (Socio-physical model of activities) as an extension of DUL (DOLCE+DnS Ultralite (DUL) foundational framework (Masolo et al., 2003). As depicted in Figure 1.30, SOMA distinguishes representational means for the physical execution of actions, the reasoning about actions, and the relations between both representations. SOMA and its modules for action perspectives and more fine-granular representation can be used to semantically represent actions and their execution from an belief-desire-intention perspective, a constraint- and optimization-based control, perspective, a linguistic one, and several other ones. EASE leverages the discipline specific perspectives on common concepts formalized in a common ontological infrastructure as a means to strengthen the interdisciplinary cooperation within EASE. Figure 1.31 shows the visualization of the SOMA ontology for representing a robot agent performing an everyday activity and the interpretation of an observed human activity.

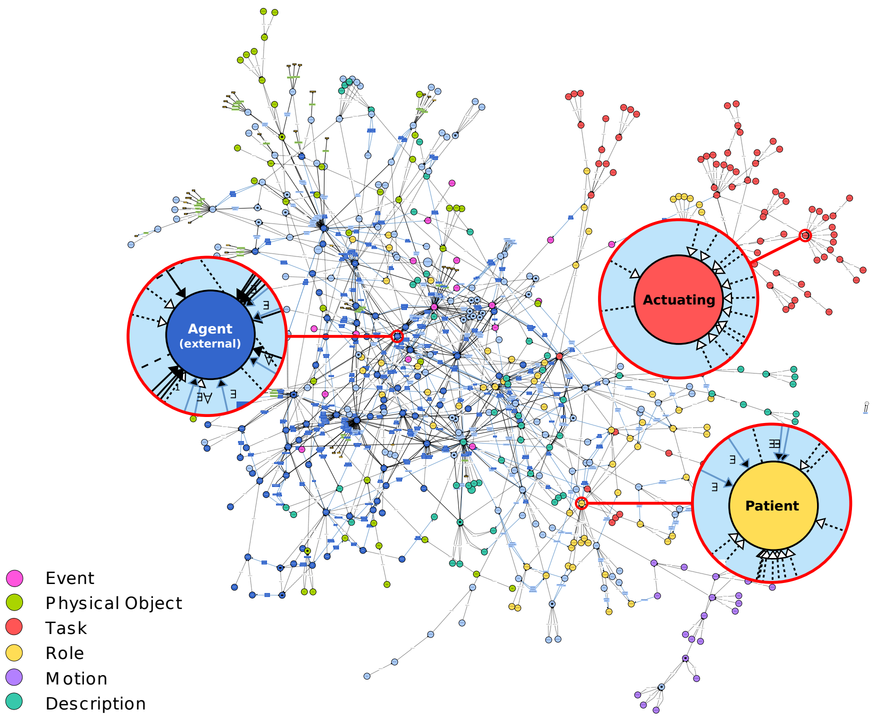


Figure 1.31: Definitions in the SOMA ontology given in VOWL (Visual OWL) notation.

In the first EASE phase various aspects of cognition-enabled robot control were formalized in the ontology. The disposition model proposed by Turvey (1992) to build a theory of affordances was included by Beßler et al. (2020b). Beßler et al. (2019b) formalized the Flanagan et al. (2006) model of actions to enable ontology driven action parsing. Other ontology extensions include Talmy (1988)'s force dynamics model to reason about activities, construction-based parsing (Bergen and Chang, 2003), and theory of image schemas (Johnson, 1987a; Lakoff, 1993).

Further ontology extensions were proposed for different aspects of autonomous robot programming including a taxonomy of execution failures (Diab et al., 2019), robot perception (Balint-Benczedi et al., 2016), and robot motion control. Relevant domain ontologies include the cooking ontology (Krieg-Brückner et al., 2015) and for autonomous robots in the retail business.

1.2.3.5.3 Efficient reasoning in expressive ontologies As ontologies in EASE are becoming very large and the relations between concepts very complex, the simplification of ontologies through approximations to make reasoning tractable is essential. A promising approach to achieve reasoning tractability is to trade-off the level of modeling detail with the reasoning effort needed to solve inference tasks. To advance our understanding of designing tractable, expressive ontologies, EASE studied in the first phase the principled approximation of expressive ontologies in lightweight languages, explored different forms of approximation and their relationships, and advanced our understanding of when optimal approximations exist, how large they are, and how expensive they are to compute. More specifically, we developed principled notions of approximation and studied in the important cases of expressive Horn DLs to inexpressive Horn DLs (Bötcher et al., 2019, 2018) and non-Horn DLs to (inexpressive) Horn DLs (Haga et al., 2020).

1.2.3.6 Key result 4: The NEEM-HUB.

The NEEM-HUB, depicted in Figure 1.32, is the data storage of robot agents that stores and manages NEEMs (Narrative-enabled episodic memories) and provides a software infrastructure for analyzing and learning from NEEMs.

1.2.3.6.1 NEEMs (Narrative-enabled episodic memories)

NEEMs are a way of storing the data generated by robot agents during everyday manipulation in such a way that enables knowledge extraction. More formally, NEEMs are CRAM's generalization of episodic memory — encapsulating sub-symbolic experiential episodic data, motor control procedural data, and descriptive semantic annotation — and the accompanying mechanisms for acquiring them and learning from them in KnowRob.

D

Narrative-enabled episodic memories (NEEMs)

are an agent's memories of activities that it executed, observed, simulated, or read about. A NEEM of an activity consists of the **NEEM experience**, which is a *detailed, low-level, and agent-specific recording of how the activity in the episode evolves*, enriched with the **NEEM narrative**, which is a *story providing information that explains what is happening in the NEEM experience*. Agents collect and store NEEMs in their NEEM system and process them in order to abstract away from specific episode contexts and learn the generally applicable commonsense and naive physics knowledge needed for mastering everyday activities.

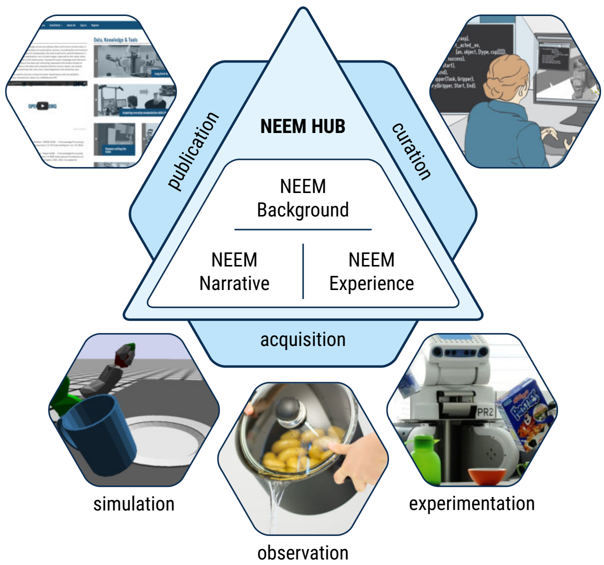


Figure 1.32: The NEEM-HUB provides the means for acquiring, curating, and publishing NEEMs, KnowRob representations of everyday manipulation episodes

NEEMs are inspired by models of the human episodic memory system, which refers to a type of declarative memory that contains autobiographical events. When an episodic memory is recalled, it results in the retrieval of the whole context of the relevant episode, including sensory, affective and cognitive processes. Semantic information such as general facts and concepts are believed to be derived from accumulated episodic memory (Tulving, 1972).

Similarly, robot agents should be able to acquire much of the knowledge needed for mastering everyday activity through NEEMs: While performing an activity, such as setting a table, the robot logs its perception and execution data in great detail. This includes sensory data (images, body poses, etc.) and control signals. These records of external perceptions and the internal semantically annotated control structures enable the robot to look at the low-level data as if they were virtual stories – narratives – about performing the activity in different ways, where robot's intentions, beliefs and behavior, perceived scenes, and effects of actions are related to each other. This story view enables the robot to answer questions regarding to what it did, why, how, how well, etc. The robot can answer queries such as: "Where do I find clean cups?", "Which is the best order to bring items to the table?", "At what times should the table be set?", and "Which perception routines work best for detcting plates in the cupboard?".

NEEMs enable the agent to replay specific experiences with its *mind's eye* and, for example, recall meaningful sub-episodes of successfully picking up a red cup. The agent can use these past episodes to learn new information, even for aspects that were not previously considered for that particular episode.

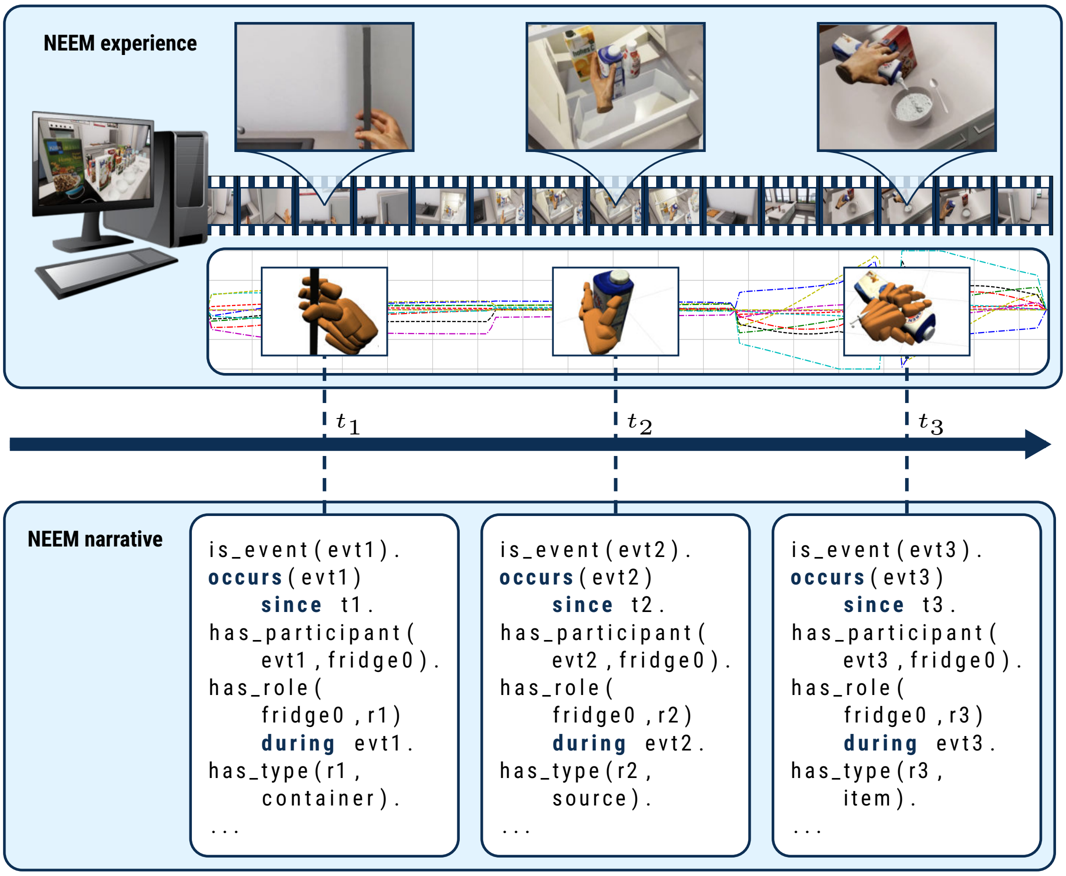


Figure 1.33: Schematic visualization of a NEEM.

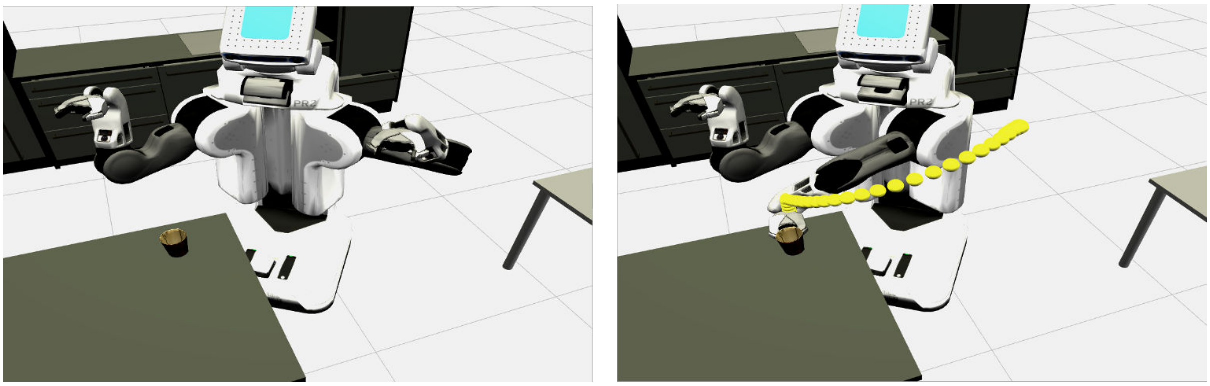
Narrative-enabled episodic memories

As depicted in Figure 1.7 the EASE generative model not only generates body motions that cause physical effects but it also generates an experience which is formally represented as a *narrative-enabled episodic memory*.

Analogously to the belief state, a NEEM can be rendered as a video of the respective activity episode. Again we can take the close matching of the NEEM video with a video captured from the real activity episode as a measure of the comprehensiveness, accuracy, and level of detail represented in NEEMs. In addition, NEEM videos can be enhanced with additional abstract information such as the trajectory of the object being carried in a fetch&place action as well as its original and final pose.

KnowRob provides a query language in order to retrieve information from NEEMs. The expressive power, that is the set of questions that can be asked about a given NEEM, is provided by the KnowRob ontology. One can retrieve all entities of a given entity category in the KnowRob ontology and describe each entity using the attributes defined for the respective entity category. In addition, the relations defined in the ontology can be used to constrain combinations of entities. The assertions about entities and relations are automatically generated from the EASE ontology. Figure 1.34 shows two example queries that are executed on a given NEEM.

The usefulness of NEEMs recorded by robot agents has also been demonstrated in the experience-based learning generative model $gm_{experience}$ discussed in Section 1.2.3.3.



KnowRob (Prolog) queries	

`entity(Tsk, [an, action,`
`  [executed_by, Robot],`
`  [executes_task,[[type,grasping]]],`
`  [has_participant,[[type, cup],`
`    [has_role, [[type,item]],`
`                during, Tsk]]]]),`
`entity(Robot,[has_part,[`
`  [type,base], [name,Base]]]),`
`occurs(Tsk, [TI_Begin,_]),`
`is_at(Base,[map,P,Q])`
`      during TI_Begin.`
	`entity(Tsk, [an, action,`
	`  [executes_task,[[type,grasping]]],`
	`  [execution_status,'Succeeded'],`
	`  [has_participant,[`
	`    [type,gripper], [name,Gr]]]]),`
	`occurs(Tsk, T_Int),`
	`is_at(Gr,[map,P,Q]) during T_Int.`

Reformulation of the queries in natural language	
Let `Tsk` be a task of the robot in which it intended to grasp an object of type cup and let `TI_Begin` be the time instant where this task started. Then infer the pose of the robots base in global map coordinates at time instant `TI_Begin`.	Let `Tsk` be a task of the robot in which it successfully grasped an object, which occurred in the time interval `T_Int`. Sample the trajectory of the gripper that was used to pick up this object throughout the time interval `T_Int`.

Figure 1.34: Visualizations of results of KnowRob queries evaluated on a NEEM in openEASE.

NEEMs are also available at execution time. This means that the robot agent can use the active NEEM to diagnose and recover from execution failures. Note that the last time instant of the active NEEM is the current belief state of the robot agent (Bartels et al., 2019).

1.2.3.6.2 Interpretation and logging of NEEMs

In order to generate NEEMs

- the CRAM plan interpreter logs the interpretation of the generalized action plans, the perception and inference tasks and their results;
- the perception executive logs the images, the action executive the body poses, and their computations;
- KnowRob2.0 logs the evolution of the belief state; and
- the plan interpreter logs the perception and force dynamic events, which constitute the basis of the Flanagan model of the respective action.

The individual logs are time synchronized with a common clock based on the occurrence of the perception and force-dynamic events. The logged data are sufficient to replay the logged activity as a video and visualize abstract models of the logged action.

As a consequence NEEMs represent the low-level data, the *"what"*, of an action episode, the abstract action model, and the internal computations together with the physical causality, that is the *"why"* of the recorded action episode.

Brunner et al. (2018) propose a method for explicitly specifying the dataflow for logging activity data that highly robust longterm operation with varying computational resources. The logging mechanisms facilitates the high-speed (up to 1000Hz) of high-volume activity data.

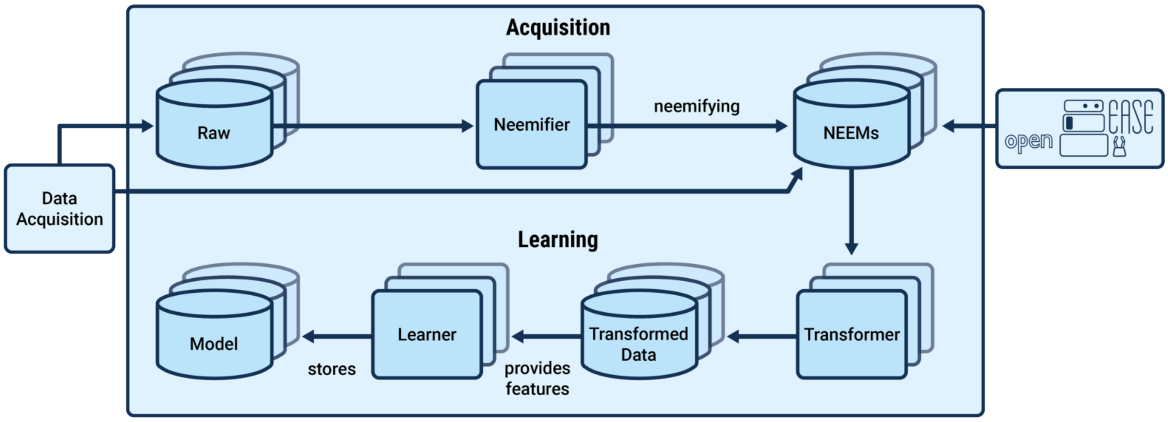


Figure 1.35: Architecture of the NEEM-HUB concept that collects, manages, provides access to NEEMs. In addition, it allows to generate learned models automatically from the collected NEEMs.

1.2.3.6.3 NEEM collection, management, and data sets The NEEM-HUB, shown in Figure 1.35, is a *software service* that collects and manages NEEMs and facilitates the use of NEEMs in information and knowledge retrieval, information analytics, and machine learning. The NEEM-HUB is implemented as a Hadoop Distributed File System (HDFS) that can store large amounts of data and provides fast access and processing for the data. For version control we are using the data version control system DVC[49]. As shown in Figure 1.35 the NEEM-HUB is designed to support two pipelines - a data acquisition and a learning pipeline. In the EASE project, we collect data from different sources. Some data can be acquired directly as NEEMs and some need to be transformed into NEEMs. Data which is not uploaded as NEEM to the NEEM-HUB is identified as raw data. Raw data can be transformed by so-called *neemifiers* into NEEMs. Neemifiers utilize the SOMA ontology to transfer recorded activity data into a valid NEEM format.

Table 1.36 provides an overview of NEEMs collected during the first phase, which are available in the final standardized NEEM format and are available on the NEEM-HUB[50] or are currently uploaded to it. Many more were collected, for example for experiments in the context of paper submissions, but are available only in an outdated NEEM format. The generated NEEMs can be accessed and analyzed by OPENEASE. In addition, all NEEM data can be downloaded in a file format such that they can be processed locally with the user-preferred software tools.

The learning pipeline of the NEEM-HUB facilitates the automatic generation of probabilistic models from NEEM collections. Since NEEMs are collected learning task agnostic, transformers are utilized to extract learning task specific features from the NEEMs. This transformed data can be used by learning models to generate statistical models, which can be used as information resources of the robot agents. A detailed description how data is can be published and accessed in the NEEM-HUB is contained in the NEEM handbook. [51]

[49] https://dvc.org/
[50] `https://neemgit.informatik.uni-bremen.de/neems`
[51] `https://ease-crc.github.io/soma/owl/current/NEEM-Handbook.pdf`

Source	#Episodes	Avg. GB	Content
Robot Bullet World	1600	0.2	Robot setting up table in an environment with naive physics and ground truth for perception.
Robot Unreal Engine	100	4	Robot setting up table in an environment with realistic physics and complete perception system.
Robot Real World	11	6	Robot setting up environment in real world.
Automatic Data Set (VR)	240	0.8	Head+Controllers tracking for "Table Setting","Washing the Dishes" and "Cleaning the Living Room"
Instructions(VR)	60	0.08	trajectories, object decisions and action sequences of table setting / cooking instructions
Neuro Data(FMRI)	600	0.025	Producing labeled segments of brain activity during watching annotated table setting video.
Neuro Data(EEG)	124	0.015	Producing labeled segments of brain activity during watching annotated table setting video.
EASE-MAD	68	0,2	Varying hand-object manipulation data under different environmental settings.
Table Setting Dataset	450	45	Tracking data and video files of people performing table setting scenarios.

Figure 1.36: A listing of collected NEEMs available in the standardized NEEM format.

Figure 1.37 shows the interactive query of OPENEASE. Having loaded a NEEM of a manipulation episode, the user queries for actions in which a door is opened, formulated in the KNOWROB query language. The query returns a bag of results with subepisodes in which a door is opened. The user then inspects one of these episodes, named $\texttt{action}_{125}$. OPENEASE provides different views of $\texttt{action}_{125}$, including a time interval representation of the motion phases, the ontological background knowledge for $\texttt{action}_{125}$, the body motion represented as the joint trajectories of the robot's kinematic chain, and a video generated from the NEEM experience of $\texttt{action}_{125}$.

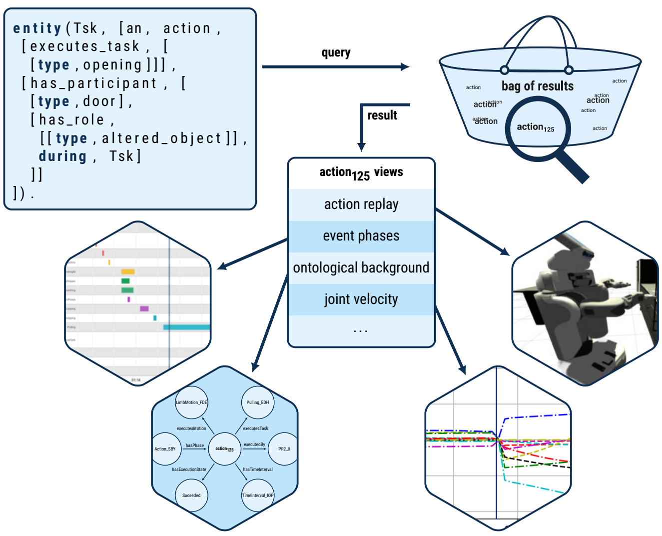


Figure 1.37: OPENEASE interactive query interface for NEEMs enabling users to navigate through the KNOWROB knowledge structures.

1.2.3.7 Key result 5: The integration of the SOMA ontology and NEEM knowledge

The fifth key result is the integration of the SOMA ontology (Section 1.2.3.5) and the knowledge encapsulated in the NEEM-HUB (Section 1.2.3.6) in a hybrid symbolic / sub-symbolic framework for observation and interpretation of activities for reasoning in KNOWROB. Figure 1.38 shows the formalization of NEEMs using the SOMA ontology, which makes the represented NEEMs machine-understandable.

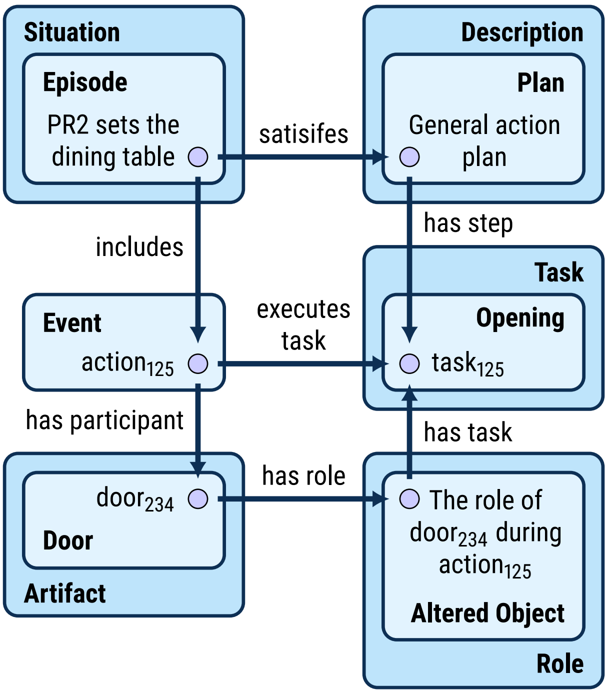


Figure 1.38: Formalization of NEEMs in the EASE ontology.

As NEEMs represent plans, robots, objects, environments, beliefs, intentions, body motions, and the physical effects they cause, the formalization of NEEMs in the EASE ontology makes all these concepts and their relationships machine-understandable. Being able to reason about these concepts and their relationships greatly impacts the the cognitive capabilities of robots. One important aspect of NEEMs that result from robot simulations is that they include the beliefs and intentions of the robot agent as well as the physical state of the environment. The combination of both facilitate powerful meta reasoning capabilities such as reasoning about what the robots do not know, which beliefs are wrong, whether plans are likely to mix up the objects to be manipulated, whether beliefs are too inaccurate to accomplish a given manipulation task, and so on.

Another advantage of formalizing the NEEMs using a comprehensive ontology is that we can define abstract concepts such as the place from which a robot can successfully pick up a large object. Using the definitions in the ontology the robot knows that places can be characterized as clusters of 3d robot poses. It also knows that a classifier for being able to perform an action successfully can be learned by retrieving all action instances in NEEMs, take the robot pose from which the action is executed and whether the action succeeded or failed. The concept of interest is then a predictor function that can be applied to a robot pose relative to the object to be picked up that predicts the success of this action. This way the distributions of motion parameters that are the grounding of action designators can be learned.

1.2.3.8 Key result 6: Automated modelling of human everyday activity

We noted in Section 1.2.3.1.2 that NEEMs are generated from four sources, two originating with the robot and two originating with humans. In the previous section, we focussed on the NEEMs that are generated as the robot acts either in the real world or in its inner world. We turn out attention now to the NEEMs that are generated by humans, both as a result of physical actions carried out by sensorized humans in the real world and as a result of actions carried out by humans operating in a high-fidelity photorealistic virtual reality environment. We begin with the latter mode of operation, which allows NEEM data to be collected more readily and facilitates the generation of large amounts of data which can subsequently be used for training, before moving on to address the former mode.

As noted in the introduction to Section 1.2.3, Result 6, automated modelling of human everyday activities, and the representation of these activities in the NEEM-HUB *and the SOMA ontology represents the second landmark achievement of Phase 1 and was only made possible by the strong collaboration between research areas H, P, and R.*

1.2.3.8.1 Interpreting and modeling human activities in high-fidelity virtual reality environments

We have developed AMEvA (Automated Models of Everyday Activities) (Haidu and Beetz, 2019), a computer system that can observe, interpret, and record fetch-and-place tasks and automatically abstract the observed activities into action models according to (Flanagan et al., 2006), and represent these models as NEEMs. These NEEMs can be used to answer questions about the observed human activities and learn generalized knowledge from collections of NEEMs .[52]

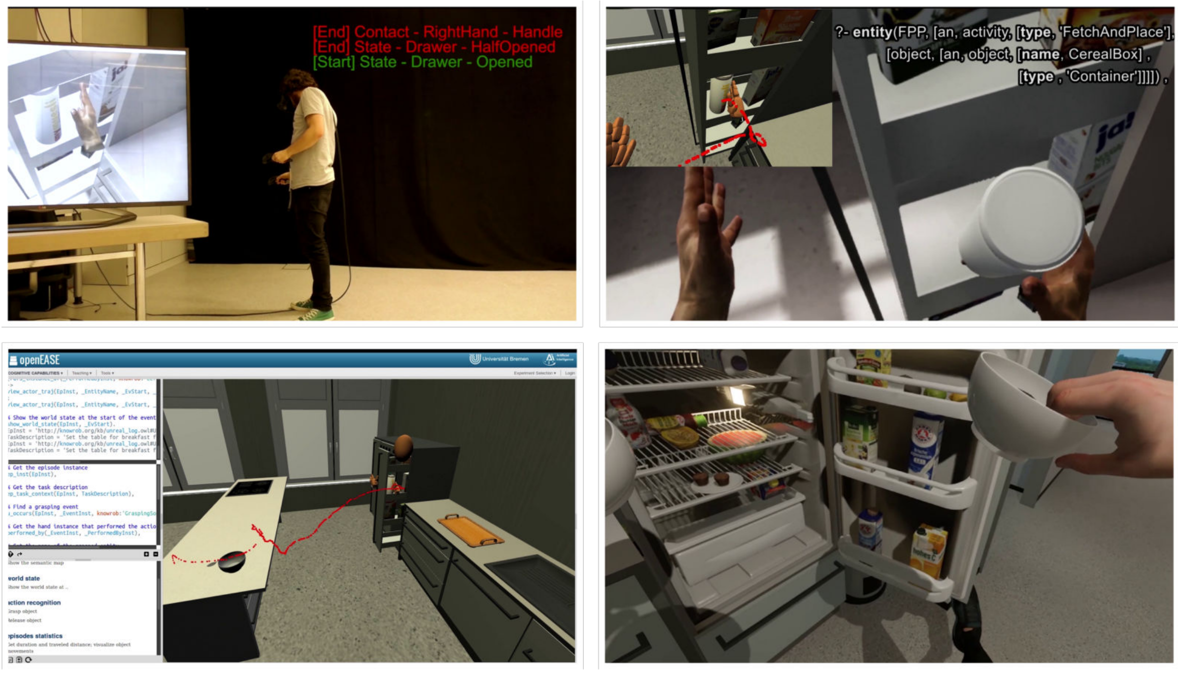


Figure 1.39: Collecting NEEMs from human demonstration.

Figure 1.39 shows the operation of AMEvA. A human is performing fetch&place tasks in a virtual kitchen environment (see Figure 1.39 (top left)). An interpreter tightly interacting with the physics engine of the virtual environment is detecting force dynamic events, such as the hand making contact with the object to be picked up, the object being lifted off the supporting surface, and so on. These events are processed by an activity parser in order to recognize actions and segment the actions into motion

[52] `https://ease-crc.github.io/soma/owl/current/NEEM-Handbook.pdf`

phases according to the action model proposed by Flanagan et al. (2006) (see Figure 1.39 (top right)) (Beßler et al., 2019b, 2020b). The parse trees of activities are then converted into NEEMs, which are compatible with the NEEMs collected by robots and linked to the same ontology (see Figure 1.39 (lower left)). Researchers can interactively work with the collected NEEMs by accessing them through the OPENEASE web service. The current level of realism of the visualization and the physical interaction is shown in Figure 1.39 (bottom right), which shows the physical grasp interaction, the lighting mechanism, and the closing of the door with the foot.

The detailed action model of taking a milkbox out of the refrigerator is depicted in Figure 1.9 (a). AMEvA recognizes the fetch&place action, segments the continuous body motion and interaction with the physical environment into the four motion phases: reach, grasp, transport, and release. The subgoals of the individual motion phases, namely the contact of the hand with the object acted on, the object losing contact with the supporting surface, the object making contact with the supporting surface at the destination, and the hand losing contact to the object being acted on, are all detected by AMEvA as force-dynamic events. Furthermore, the context-sensitive parameterizations of the body motion are extracted and recorded from the physics simulation. These parameters are for the reach phase the pre-grasp pose of the hand, the grasp shape, and the positions of the fingers on the grasped object. The action model allows the replay of the action as a rendered video.

The online parsing of observed activities carried out in the virtual reality EASE kitchen is shown in Figure 1.40. The figure shows four sub-episodes of picking up an object. The KNOWROB query in the right upper part shows the KNOWROB query that extracts the sub-episode, and the abstract visualization in the upper left shows the respective hand and object poses and the object trajectory that are extracted through the query.

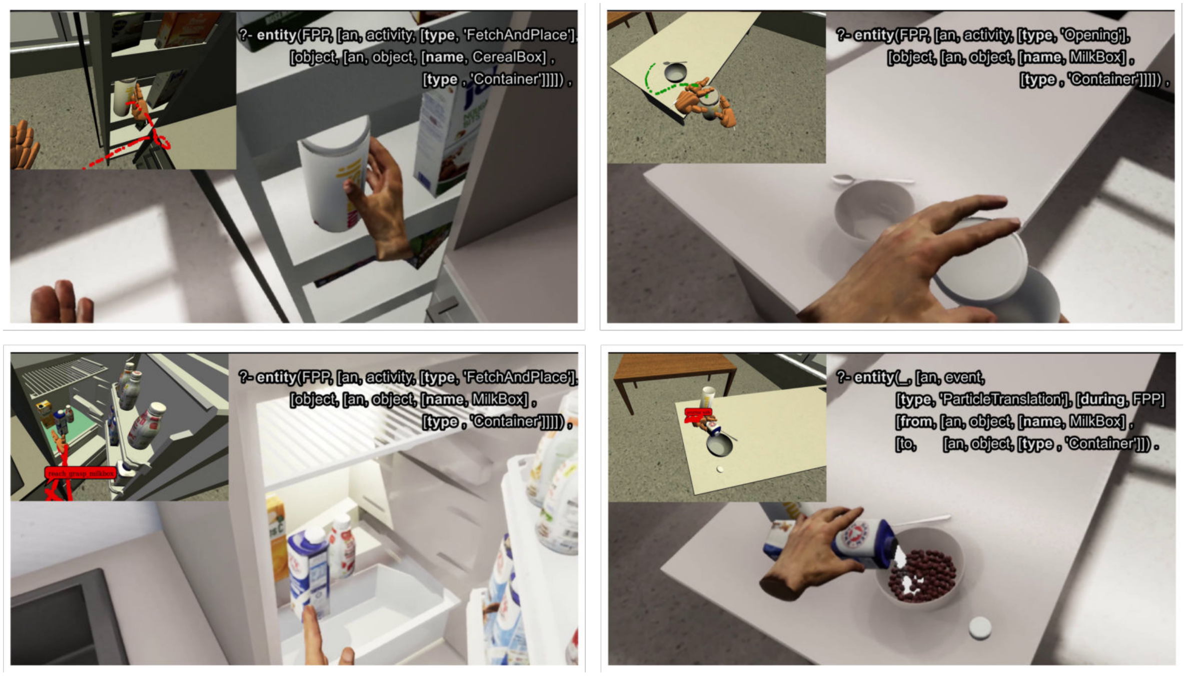


Figure 1.40: AMEVA parsing human virtual reality activities into NEEMs.

The output of the AMEvA action interpreter are NEEMs that are formalized as KNOWROB knowledge bases. They represent episodes using a first-order time interval logic representation. NEEMs are uniformly represented through a common data format and the semantics of the NEEM narratives[53] is formally specified in the EASE ontologies. This semantic groundwork facilitates semantic access

[53] `https://ease-crc.github.io/soma/owl/current/NEEM-Handbook.pdf`

to information and data elements contained in NEEMs as well the automated combination of NEEMs with different modalities by warping their structure using the NEEM narrative. These NEEMs are the data resources that can be used in order to investigate how humans accomplish their everyday manipulation tasks. In the following, we give examples of research questions that can be studied based on automatically acquired collections of NEEMs.

The first example is how humans interpret vague instructions, such as "set the table", as concrete activity goals. Assuming that the human has successfully accomplished the instruction in the respective episode we can take the goal to be the final state of the episode. Thus, we can compute the goal inferred by the human from a given NEEM by collecting each object *obj* that the human has transported in the episode and add the pose of *obj* at the beginning as the initial state and the pose at the end as the goal state. The illustrated KNOWROB query in Figure 1.41 (left) highlights each object that was grasped in the episode in its initial and final state. In Figure 1.41 (right) we show the result of a segmented fetch&place action into its subactions.

Figure 1.41: Visualizing the initial and final state for a "setting the table" episode (left) and the segmentation of a fetch&place action (right).

Given a collection of NEEMs from EASE robot days we can also investigate the flexibility with which manipulation actions can be executed. We can do this, for example, by visualizing the object and the hand pose for every situation in which an object was grasped in the episode, i.e. when the hand made contact with the object to be picked up.

We can further explore the context sensitivity of action strategies. For example, we can investigate how the grasp strategy depends on the object to be grasped, the surrounding scene, and the destination pose of the object. To this end, we can characterize a grasping scene by the object features (including whether the object has handles, size, shape, parts, state, weight), the closest obstacles and location, and the destination. This way we can, for example, learn that a container placed on a hot oven plate is always grasped by its handles. The context-sensitive action strategy can be inferred through decision tree learning on all grasping behaviors in a given collection of NEEMs.

Furthermore, we can study cognitive attention mechanisms of humans during everyday activities. This can be done, for example, by extracting the gaze data of NEEMs (see Figure 1.42 (left)). Here we can exploit the widely accepted principle "gaze leads action" (Land and Hayhoe, 2001), that is that the gaze targets the object to be grasped before the reaching motion starts, and extract the objects targeted by the gaze before starting to reach.

The stereotypicality of hand trajectories in everyday manipulation activities can, for example, be investigated based on the trajectories of objects during the transportation phase of fetch-and-place tasks as visualized in Figure 1.42 (right). Here we can cluster and categorize the trajectory data to get more compact activity models. The most recent results of recording and interpreting human

Figure 1.42: Visual attention during object manipulation (left) and the trajectories of objects from multiple NEEMs during the a fetch&place action (right).

manipulation activities in virtual reality are shown in Figure 1.43 and in the accompanied video.[54]

Figure 1.43: New virtual reality fullbody activity tracking, recording, and interpretation. Fullbody pose and object trajectories (left) and gaze tracking and visual attention (right). The different colors show the respective motion phases.

The examples above give a glimpse of the variations of research questions about human everyday manipulation activities that can be tackled based on collections of NEEMs automatically constructed from observing human activities is virtual living environments.

The machine-understandable semantics of NEEMs facilitates the automated extraction of training data for the learning generalized models of human activity. We can also complement the training data with prior models of human activities which include

- the action model framework proposed by Flanagan et al. (2006),
- the situation model framework for cognitive behavior (Schneider et al., 2020), and
- the generative model of robot activity that was introduced in Section 1.2.3.1.

Machine learning and data mining mechanisms that are suitable for conducting such investigations into the principles of human everyday activities are provided through the interactive web-based knowledge service OPENEASE.

[54] `http://ease-crc.org/link/video-ameva`

Particularly important in this research setting are learned models that are compatible with the EASE generative model and because they can be integrated into the generative models with reasonable effort and then tested in real robot experiments. This allows researchers in area H to test research hypotheses using the generative model and compare different hypothesis using the generative model. This is achieved by providing an application programmers' interface, in which the plan execution system lays the belief state of the robot open as a KnowRob knowledge base and can interpret the answers to the body motion query given by other software systems.

In the context of this method to obtain insights and construct models of human everyday activity, EASE measures the blackbox progress with respect to descriptive and analytical models in terms of

- the amount, variability, and breadth of the segmented, interpreted, and computer-understandable behavior data from everyday activities and the modalities that are included in data streams,
- the generalized knowledge that EASE has accumulated, learned, and abstracted from the activity data,
- the structural constraints on activities that can be exploited to improve the accomplishment of everyday manipulation tasks, and
- synergies that could be obtained through the interaction between generative and analytical models of everyday activity, that is the improvements in generative models that could be achieved through insights obtained from analytical models and insights obtained on human everyday manipulation capability using the generative model as a framework for investigation.

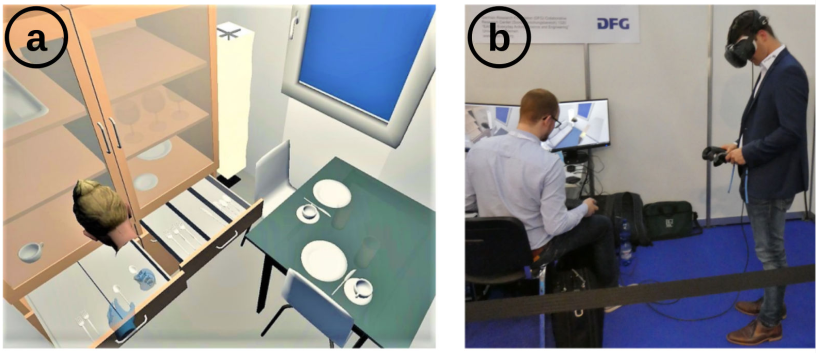


Figure 1.44: The *Household Activities from Virtual Environments (HAVE)* dataset. This figure illustrates how humans leverage virtual reality in a photo-realistic and physics-faithful kitchen (a) at Automatica to record at lower cost activity data (b).

Pfau et al. (2019a) have developed a software pipeline that reads instructions for everyday manipulation tasks from web pages such as WikiHow, processes the langage instructions and identifies knowledge gaps in these instructions. These gaps are then used to create a scene in a virtual reality game (called *"Kitchen Clash"*) that is designed to fill the identified knowledge gaps. For example, the robot agent might want to learn how to "prepare a portion of cucumber salad," for which WikiHow gives the following instruction "Slice the cucumber into thin pieces, place the slices into a bowl and pour dressing over the cucumbers." The instruction lacks many pieces of information that a robot agent needs to accomplish the task successfully. For example, the tool for cutting, the shape of the pieces that result from cutting, the placement of cut pieces in the bowl, and so on. These knowledge gaps are then turned into decision making problems by placing different cutting tools such as scissors, saws, and knives in the scene and observing which one the human game player selects for cutting a cucumber. In order to identify possible reasons for the choice of tools *"Kitchen Clash"* runs physical simulations with the tool and cucumber in order to provide the necessary intuitive physics knowledge. Acquiring manipulation knowledge with *"Kitchen Clash"* has potentially high impact on making robot learning much more efficient. The knowledge which is generated in terms of NEEMs cannot only be used to acquire commonsense knowledge but also to better guide imitation learning and for shaping reinforcement learning problems for more complex manipulation tasks. In imitation learning one of the key problems is to decide which aspects of the demonstration are essential for accomplishing the task, which include the decisons made by the game players. In reinforcement learning the results of the game can facilitate the specification of appropriate subgoals and the definition of appropriate objective functions.

Bates et al. (2017) have, in close cooperation with the Inamura laboratory, developed a VR online activity recognition system, which served as basis for Uhde et al. (2020), who recorded the *Household Activities from Virtual Environments (HAVE)* Dataset (see Figure 1.44). The *HAVE* dataset contains three different household environments: dining room, kitchen and living room, which are used to evaluate different tasks, e.g. setting a table, washing dishes, and cleaning a room. The *HAVE* dataset contains NEEMs from 240 participants recording sessions, which capture a large variety of different household manipulation tasks. This variation is key in generalizing behavior patterns and the conditions under which they are applied.

Uhde et al. (2020) uses the *HAVE* dataset to examine correlations in human activity patterns, in order to extract causal hypotheses of action dependencies. An example for these action dependencies is the requirement of opening a cupboard before being able to grasp objects inside. The causal dependency hypotheses are verified through performing intervention tests in mental simulation.

EASE also generated NEEM data sets from virtual simulation environment at the level of fine-granular hand motions including sound and haptic feedback rendered through a Phantom device. This data set focusses on grasping virtual objects, based on grasp simulation and categorizing grasping episodes using a newly proposed grasp taxonomy. The novel features include a novel physically based Material Point Method for rigid deformable objects, a stable physics-based multicontact manipulation of objects, heatmaps visualize contact points from human grasping in VR, and human grasping with robot hands: adaptive strategies for different physical constraints.

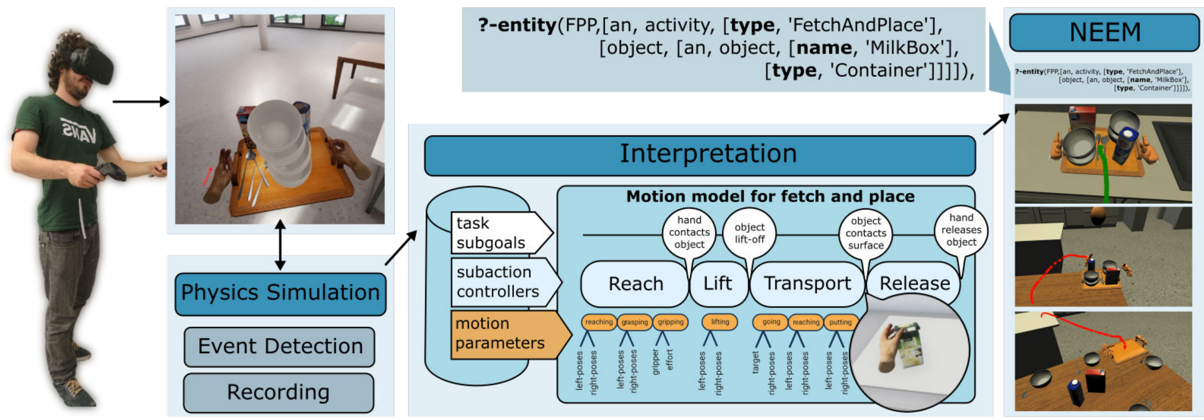


Figure 1.45: Schematic diagram of the AMEvA interpreter for virtual everyday manipulation actions.

Activity recording and interpretation in VR In the first EASE phase we have developed AMEvA (Automated Models of Everyday Activities), a special-purpose knowledge acquisition, interpretation, and processing system for human everyday manipulation activity in virtual environments (see Figure 1.45). It can automatically

- create and simulate virtual human living and working environments (such as kitchens and apartments) with a scope, extent, level of detail and physics that facilitates and promotes the natural and realistic execution of human everyday manipulation activities;

- create a symbolic knowledge base from virtual reality environments that represents all objects in the environment, their parts and their articulation models. This makes the system omniscient with respect to the environment. The knowledge base is extended with naive physics, commonsense, and background knowledge about the objects;
- record human manipulation activities performed in the respective virtual reality environment as well as their effects on the environment;
- detect force dynamic states and events in the recorded activities;
- decompose and segment the recorded activity data into meaningful motions and categorize the motions according to action models used in cognitive science;
- represent the interpreted activities symbolically in KNOWROB using first-order time interval logic formulas linked to subsymbolic data streams.

We apply AMEvA to generalized fetch&place tasks (including organizing the kitchen, setting and cleaning the table, loading and unloading the dishwasher). The challenges include obtaining access to the relevant data structures of the game environment including objects, the functional structure of the objects and their articulation models and the force dynamic events that happen in the physics simulation of the environment. We collect, manage, and provide public access to the observation data, models, and the symbolic representations of the activity episodes through the open and web-based robot knowledge service openEASE.

1.2.3.8.2 Interpreting and modeling human activities in the real world

The interpretation and modeling of human everyday activity is mainly conducted in research area H, which aims to advance our understanding of how humans master everyday activities. This main theme will be driven forward by the following major thrusts, (1) building on the rich and highdimensional multi-modal biosignal databases acquired from humans performing everyday activities along with the knowledge and experience that was accumulated in the first phase of EASE, (2) utilizing these resources to further advance human activity models by both a data-driven bottom-up training of discriminative models based on machine learning and statistical methods and a top-down design and implementation of generative models ranging from low-level fine-grained models of fetch&place actions all the way to sequences of complex activities such as setting a dinner table. The creation of discriminative and generative models of everyday activities is complemented by the development of everyday-like as well as abstract versions of decision tasks entailed in everyday activities to pinpoint the executive control mechanisms of human learning and decision making that enable human mastery of these tasks.

To deliver to the high demands on research area H and to provide both the “push and pull” in EASE (see introduction), extra measures are taken to ensure a tight integration. Here, we build on two major achievements of area H in phase 1: (a) an end-to-end pipeline consisting of a common information processing framework jointly designed, implemented and experimentally evaluated within area H (Mason et al., 2020) and (b) the concept of LabLinking, a technology-based interconnection of experimental laboratories across institutions and disciplines that supports experiments without borders.

Mason et al. (2020) propose the EASE Human Activity Data Analysis Pipeline shown in Figure 1.46. It serves as a common unified framework for data capturing and processing which emerged from the close collaboration within area H. The pipeline outlines the information-flow between humans and robots with the NEEM-HUB as its central component to dynamically store, archive, and retrieve data in all processing stages. The Recording stage encompasses any kind of data acquisition, where a human performs any kind of everyday activity, recorded by any combination of sensors. In the Conversion stage, data is arranged into a NEEM compatible format and automatically uploaded into the NEEM-HUB. Alternatively, third party data can also be integrated in the same way. Once the data is published, it can be accessed by other programs or scientists. The EaseLAN component is located in the center of the framework and comprises the semi-automatized process of NEEM narrative creation for human data. The component is based on the open-source ELAN software toolkit that is very well suited for data visualization and manual annotation. A range of enhancements created within EASE

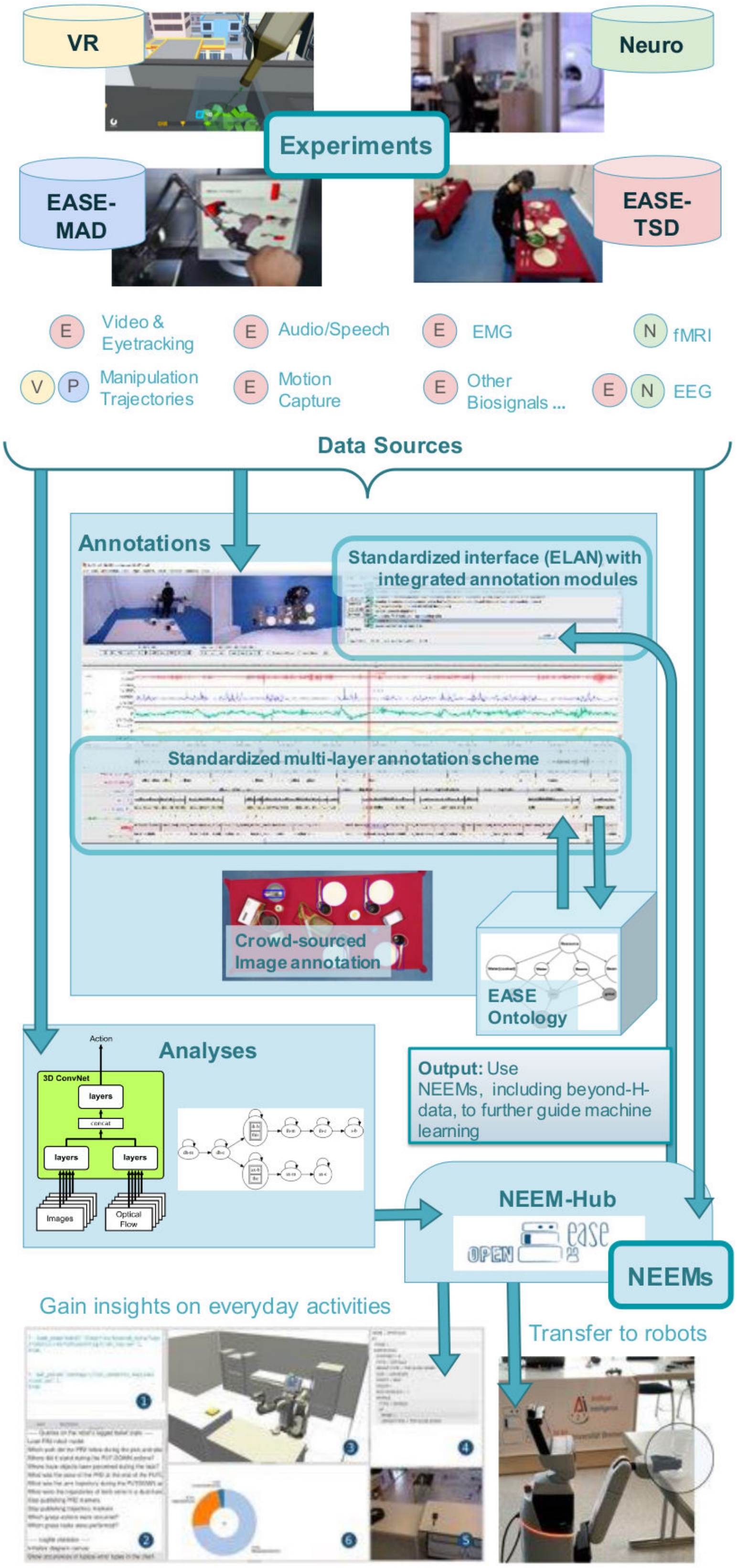


Figure 1.46: The processing pipeline for activity data in human real-world everyday activity experiments.

simplifies the annotation process through automatic annotation of actions, objects, and think-aloud protocols from the multimodal biosignal data. Most importantly, this component integrates the accumulated knowledge and data sources into the EASE infrastructure. For this purpose, the enhancements encompass for example the automatic recognition of think-alouds from audio channels or the recognition of activities from video and motion capture data, which result in automatic annotations that can be iteratively corrected by human annotators in a human-in-the-loop fashion. NEEMs can be downloaded, modified or supplemented and re-uploaded as new versions of an episode. The annotation schema is automatically generated from the SOMA ontology (previously EASE-Ontology, under continuous development in research areas P and R) and is commonly shared within area H. Vice versa, concepts which are unknown to SOMA can be manually added and proposed as extensions to the ontology. Once generated, the NEEMs become available for the robot, either through direct query of individual episodes or indirectly through a generative model which generalizes the available NEEMs.

As mentioned, the Human Activity Data Analysis Pipeline integrates the acquired knowledge and data resources of Human Everyday Activity into the EASE infrastructure, connects with SOMA, generates NEEMs, and leverages the NEEM-HUB to making NEEMs available to the robot and any other agent connected to this infrastructure. It is therefore only a small step to also interconnect humans through this knowledge infrastructure of agents. To establish this link, we use the concept of LabLinking, i.e. a methodology and software stack to conduct decentralized experiments, i.e. in different places, at different times. In EASE the NEEMs and the NEEM-HUB serve as mediator. For example, NEEMs recorded in one place guide an experiment in another place. As a result, more time aligned modalities and biosensor data may add to an episode, either subsequently or in real time. The capabilities of NEEM-HUB are leveraged to conducting those experiments by structured data exchanges as well as providing a framework to express the semantic relations. Such a scenario was implemented and evaluated within EASE. Here, the EASE-TSD data were extended by fMRI recordings of subjects who observed in first-person videos, how another subject performs a table setting task. Next, the setup will be advanced by replacing the previously recorded videos with a video live stream and bi-directional data exchange, thus allowing different protagonists to synchronously interact between labs. A software for distributed experiments developed in subproject H03 is reused in a two-locations synchronous experiment for this purpose. This next level LabLinking approach will provide the stepping stone to integrate decision making models and opens up new avenues for collaboration within and beyond EASE.

1.2.3.8.3 Human activity data sets The EASE-TSD dataset covers about 450 episodes of table setting activities from 71 subjects resulting in more than 10,000 NEEMs annotated on action level and more than 8,000 NEEMs annotated on motion level. Roughly 30,000 NEEMs on action level will be available that the end of the first phase. The EASE-MAD dataset covers more than 2,200 NEEMs of hand-object manipulations on the action level from 32 subjects.

Figure 1.47: Recording table setting episodes including biosignals and "think-aloud" protocols in the Biosignals Acquisition Space and Environment (BASE).

1.2.3.8.3.1 The EASE table setting data set (EASE-TSD) EASE has collected the EASE table setting data set from 100 participants, which includes table setting episodes performed under varying conditions for generalizability. The variations include formal vs. informal meals, meals of 2 vs. 4 persons, and breakfast vs. dinner. activity, eye tracking, motion capture, video, audio (b) concurrent and retroactive think-alouds, and (c) Brain activity from mirror perception (n=67).

EASE has investigated the modeling of human-scale everyday manipulation activities through the creation of hierarchical temporal structures that decompose recorded activities into general units containing semantic descriptors. The aim is to propose human activity models that are person-independent and cover a wide range of EASE robot days. To this end, EASE applied machine learning techniques to find manifold representations of everyday activities of lower dimensionality than the original data. EASE researchers apply deep learning approaches to large volumes of multimodal sensor data captured from many subjects. The use of these learning methods is facilitated by automated data processing that implements the segmentation, annotation, and embedding of the activity data into a semantic representation.

As a result, EASE has acquired a large database of human everyday activities (see Figure 1.47) along with a procedure to automatically structure and label these high-dimensional data into a valuable resource for research in cognitive robotics. The dataset consists of synchronously recorded biosignals from about 100 participants performing everyday activities while describing their task through use of think-aloud protocols (Mason et al., 2020). Biosignals encompass multimodal sensor streams of near and far speech & audio, video, marker-based motion tracking, eye-tracking, as well as muscle and brain readings of humans performing everyday activities. The recorded multi-stream sensor data are automatically segmented and semantically annotated and then transformed into NEEMs.

Activity episodes are represented as NEEMs that include video annotation producing labeled segments for head, body, and arm actions (L2) & low- level motions (L1), and audio transcriptions (L1) of 'Concurrent' and 'Retrospective' think-aloud protocols, with associated utterance codings (L2). The multimodal sensor data recorded and processed in BASE include:

- Brain activity measurements allow evaluation of attentional focus while performing tasks, adaptation to ambiguous or conflicting situations or physical obstacles, decision making processes, and how motor imagery when viewing performance of activities compares with in-situ motor execution.
- Skin and muscle activity sensors, located on hands and forearms, can indicate overall mental state, and information about manual manipulation interactions with objects, such as the force used.
- Full body motion capture provides motion in an environment and object interactions, such as hand trajectories when picking and placing objects.
- Scene video from many perspectives allows tracking of objects and the order of interactions, insight into efficient movement within a space while performing tasks, and the deft object manipulation habits and problem solving strategies humans display in diverse conditions.
- First person video provides understanding of scene aspects people may focus on while planning and executing tasks. Important information for a robot might include attention (internal vs external) and visual search strategies for objects or positions based on contexts such as meal type, formality, or number of diners.
- A stationary scene microphone provides audio of the individual sounds occurring as objects collide, which provide insight into the objects' physical properties - material, shape, weight, surface texture, etc - and the force used when placing items.
- A voice microphone records speech and non-speech vocalizations during table setting activities. Through audio recordings of what a person thinks-aloud while they perform tasks, we gain a rich description of the scene as the performer sees it, obstacles encountered, reasoning and problem solving approaches, frustration or enjoyment, and the task process as a whole.

In order to to get better intuitions about the information processing mechanisms in the human brain that generate the behavior we have also correlated the multimodal data streams with brain activation data. For this purpose we measured the brain activity of 30 participants in an EEG-study, consisting of four 1st-person Videos. The videos were annotated according to EASE-ontology, resulting in

Figure 1.48: Obtaining the MRT data from watching a table setting episode.

312 distinct episodes of various categories and complexity levels. An fMRI-study with 30 participants consisting of ten 1st-person videos was recently finished (see Figure 1.48). For the EASE-TSD experimental recordings, 62 sessions (totaling 16.39 hours) have been recorded, comprised of six or more trials each. Over 16,600 action segments from 72 trials, broken down into between 2 and 12 category sets, have been annotated at varying levels of granularity. Over 28,000 transcribed words from the think-aloud speech during 109 trials have been created, with 62.2 average segments per trial. Think-aloud encoding annotations are underway. Multiple annotations and transcriptions continue to be created as analyses progress. The resulting EASE Table Setting Dataset (EASE-TSD), which is accessible in OPENEASE[55] contains

- 70 recordings of 6+ table setting task trials in the EASE-Base; 69 participants "think-aloud" during the task and/or during a review of video.
- Modalities: Audio mics (scene and voice), video cameras (6), eye-tracker, motion capture, ACC, EDA, EMG, EEG
- brain activity: FMRI- & EEG-recordings of participants who actively imagine themselves to act out presented table setting tasks from a 1st-person perspective
- Modalities: 3-Tesla MRI-Scanner and high-density multi-channel EEG-system

For usage in machine learning we have split the EASE Table Setting Dataset in 75% training set and 25% validation set. The recordings are continuous requiring a sliding window for processing. The annotations are available with various levels of abstraction and multiple annotations can be present, meaning that in a discriminative task a multi-label capable system is preferable. We have evaluated our activity annotators, a multi-modal convolutional neural network Gadzicki et al. (2020), with a sliding window of 0.53s (16 frames with 30fps) and achieved 87.8

1.2.3.8.3.2 The Manipulation adaptivity dataset (EASE-MAD) EASE created the EASE-MAD (Manipulation Adaptivity Dataset) that contains 2260 recordings of hand manipulation episodes under different sensory-motor conditions. The behavior data include semantic annotations, kinematic parameters, hand trajectories, contact points and contact forces (Mason et al., 2020).

Automated classification of intended vs. unintended contact events. By observing the differences of hand kinematics between intended and unintended contacts we developed a model for the identification of motor intention using pre-contact kinematic features (Maldonado Cañon et al., 2020b). By this we are able to trace and label accidental events in NEEMs acquired from recordings of human movements during placing tasks. This model can be applied to automatically separate poor from successful task performances in data used for robot learning from human demonstrations.

[55] `www.open-ease.org`

Abstraction Level	TSD Action	TSD Audio	Neuro EEG	Neuro FMRI	HAVE VR	MAD
L5-Participant Session	71	71	32	30		32
L4-Task Trial	443	443	128	150	240	
L3-Task Phase	1381			300		
L2-Activity	9057	347	18832	14580		2260
L1-Motion/Utterance	10083	13686		29160		

Figure 1.49: Collected NEEMs (decreasing abstraction levels)

Acoustic and tactile feedback in placement tasks. Based on the EASE-MAD data, EASE could analyze the contributions of individual sensor modalities to task execution and action perception. We could also advance our understanding of how humans react when a modality is missing or unreliable. Maldonado Cañon et al. (2019) investigated the effect of sensory feedback on the adaptation rate of the movement kinematics during an object placing task executed in a VR haptic acoustic simulator. At the moment of contact uni-modal (either haptic or acoustic) or multimodal (haptic-acoustic) feedback was provided. Results show that motor adaptation by acoustic feedback alone is possible. Haptic feedback, presumably due to the high sensitivity of human sensors, enables a fast adaptation rate, with acoustic information then providing no further significant improvement.

Parametric model for the transfer of human kinematic skills to robots. Adaptivity analyses were conducted for the improvement of robot motor capabilities and movement generation strategies. Human demonstrations in VR were used to develop a parametric model of task execution, based on smoothness, efficiency and accuracy features. The model parameters can be queried either to retrieve trajectory parameters (e.g. movement time as an efficiency criterion or contact velocity as a safety/accuracy criterion), or for retrieving a generalized trajectory.

NEEMs were generated from the adaptivity experiments conducted in the VR haptic-acoustic simulator as pick and place episodes. These NEEMs contain demonstrations of humans manipulating virtual glasses and spoons under different feedback conditions (uni- and multimodal; visual, haptic and acoustic). The association of these episodes to the concept *pick and place* of the EASE ontology enables to retrieve different parameters of interest (e.g. movement duration or contact force) via queries.

1.2.3.8.4 Summary of human activity datasets

Collected NEEMs EASE has collected a large and diverse set of NEEMs from humans, in different environments (artificial and real), performing different tasks (table setting, cleaning, shelf replenishment, etc). The NEEMs are collected in a standardized EASE data format and semantically annotated with concept definitions of the EASE ontology. The NEEMs are made publicly accessible in OPENEASE. Each data set comes with predefined KNOWROB queries that visualize key information and users can develop additional queries.

EASE researchers and collaborators have acquired a diverse set of NEEM collections, which are made publicly available in OPENEASE and listed in Figure 1.49.

A key result of the first funding phase are the data processing and interpretation pipelines for the acquisition of NEEMs from robot agents, human agents acting in the virtual world, and humans acting in the real world. These processing pipelines are different because in each setting the access to features needed to generate activity models and to ground truth data is different. Thus, different methods are needed to infer the latent components of the models.

1.2.3.8.5 Building activity models and transferring them into the generative model The brain is like a committee of experts. All the parts of the brain work together, but each part has its own special properties, functions, and tasks. In this view the brain can be divided into areas which specialize in different functions. To better understand how aspects of everyday activity are correlated with functions of the brain, EASE investigates the correlation between brain activation and the structure of action episodes. The basic idea is that we identify concepts in the EASE ontology that are believed to be functions of a brain area and use them to label NEEMs. The hypothesis underlying this effort is that one can train deep neural networks to predict, given a NEEM possible brain activation patterns and given brain activation patterns predicting cognitive processes that are active.

The analysis of fMRI-data is based on different methodological approaches. Statistical models such as the General Linear Model (GLM) and Independent Component Analysis (ICA) will allow for the contrastive analysis of differences in spatiotemporal patterns of brain activity related to annotated semantic episodes within a perceived point in time during video presentation (e.g. pick up, place, carry), thus leading to the detection of distinct neuronal networks correlating with ontological categories.

Building on this knowledge of brain areas that are closely correlated in their activity to ontological categories, further research will also aim at developing algorithms that predict stimuli and semantic episodes on different levels of complexity. Based on the present findings derived from a combined use of multi-channel EEG and fMRI and allowing for a detailed examination of spatiotemporal characteristics of NEEMs and ontologically different time periods we will further elaborate the semi-automatic scene recognition approach. Introducing more complex and interactive decisions in table setting scenarios we can feed information into the planned process of automatic activity recognition and its annotation within the EASE Human Activity Data Analysis Pipeline. Data recorded using mobile EEG and EMG sensors placed on the forearms, after undergoing preprocessing, then statistical and spatiotemporal feature extraction, is the basis for classification of pick and place activities. At the lowest level, data from sensors placed at four positions on each arm (e.g., on muscles controlling hand activity of the right forearm) and scalp (e.g., motor regions on the left hemisphere) is used to classify hand movements. EEG data is further filtered to the frequency bands typically corresponding to motor imagery or motor execution–the alpha and beta bands from 8-12 Hz and 12-30 Hz, respectively.

Initially, manually labeled segments such as ‘reach’, ‘grasp’, ‘release’, and ‘retract’ are used for classification in a supervised manner. Methods typically used for motor execution or motor imagery in EEG data, such as support vector machines (SVM), Random Forest/Gradient Boost, and Neural Networks are employed [49], and sequential approaches such as hidden Markov Models (HMMs) and Long Short Term Memory (LSTMs) will likely improve on these results.

As additional data such as eye-gaze or skeletal position coordinates from these experimental recordings is processed, it will be incorporated into the analysis pipeline to classify a broader range of activities. This will provide the basis for additional custom ELAN recognizer plugin development, to generate activity annotations based on multimodal biosignals.

1.2.3.9 Key scientific insights resulting from the 1st phase

The most important insight obtained in the first phase of EASE is that we could propose and realize a generative model that is **sufficient** for (1) for accomplishing underdetermined everday manipulation tasks; (2) infering each body motion through automated reasoning; (3) realizing the introspective capabilities to reason about what the robot does, why, and how; and (4) can substantially improve its capability through self-programming.

The work horses of the generative model are computer programs (algorithms and data structures) that implement core cognitive functions, most notably episodic memories, inner world models, and mental simulation and prospection. All of them were realized through hybrid symbolic/subsymbolic knowledge representation mechanisms to make these representation structures machine understandable.

1.2.3.9.1 Body motion query and digital twin knowledge representation and reasoning One of the key insights of the first phase is that performing everyday manipulation tasks can be framed as answering the body motion query:

how do I have to move my body
    in order to
        accomplish the *given underdetermined action description/request*
        for the current *task context*
        with the *object*s and in the *scene context*
        that I see or believe

and executing the answer as a parameterization of a generalized motion plan. Inferring the answer to the body motion query requires solving many knowledge-intensive reasoning tasks including inferring the goal state, guessing preferences over execution variants, determining the relevant task and scene contexts, and the properties and affordances of the objects and tools relevant for the task for which symbolic reasoning is an appropriate method.

However, KR&R systems abstract away from the motion level and model actions as state transition systems with atomic state transitions (Ghallab et al., 2004). In EASE we have proposed digital twin knowledge representation and reasoning (DTKR&R), which leverages scene graphs of virtual environments as an implementation basis for the representation of robots, physical objects, scenes and environments. DTKR&R represents entities at levels of detail that are sufficient for reasoning about robot motions and in addition, allows for the visual rendering of belief states and the simulation-based reasoning. EASE has shown that DTKR&R reasoning mechanisms are sufficient for generating the contxt-sensitive behaviors that enable robot agents to accomplish tasks such as setting or cleaning the table.

> So far, EASE has only protypically realized part of the envisaged DTKR&R functionality, which is already at this stage a huge step change for the realization of cognitive robot capabilities. We are not aware of any other reasoning system for autonomous robots that comes close to offering the reasoning capabilities of DTKR&R. The investigation of DTKR&R will be a key objective of the second EASE phase.

1.2.3.9.2 Manipulation capability = generalized plans + knowledge A second key insight from the first phase is that manipulation capability can be considered to be the result of combining action category specific high-level- and motion-plan schemata together with knowledge bases consisting of general, modular knowledge chunks.[56] When holding the high-level and motion plan schemata invariant the manipulation capability can be improved by extending and improving the knowledge base. This

[56] This insight is inspired by Kowalski's seminal paper "algorithm = logic + control" (Kowalski, 1979)

is a highly promising insight because the general modular knowledge chunks are applicable even to manipulation tasks that the robot agent has not previously encountered and therefore can help the robot to scale towards open task domains. In addition, the fact that reasoning processes are realized through the composition of knowledge chunks which greatly enhances the introspective capabilities of the robot agents and turn them into robots that "know what they are doing" (Brachman, 2002).

In Section 1.2.3.3 we have demonstrated that tasks such as setting or clearing the table that require a large variety of context-sensitive robot behavior can be handled by providing the appropriate knowledge bases. We have also shown that performance can be further improved by equipping the robot with knowledge that allows for the prediction of action effects. The knowledge base can be further improved by the robot agent itself by learning more task-specific knowledge chunks from the episodic memories that the robot collected performing the everyday activities. One reason that these methods worked was that we could come up with very general motion plan schemata that require a small number of motion phase specific parameters to generate all necessary behavior variations. This finding is consistent with findings in action science Prinz et al. (2013) who propose that human manipulation actions have fixed structures and many invariants.

The discretized reasoning and behavior generated by the motion plans is probably not sufficient to achieve optimality of performance. Again this corresponds well to findings in action science that distinguishes between habital and cognitive behavior (Schneider et al., 2020) and thinking fast and slow (Kahneman, 2011). One promising idea here is to take the optimized general motion plans and use them for shaping reinforcement learning problems that could then achieve near-optimal performance at the cost of losing flexibility and becoming more fragile when acting in open task domains. This suggests a very promising research direction in which we want to explore how to orchestrate cognitive learning and reasoning with end-to-end learning of manipulation behavior in order to obtain the best of the two worlds: fast learning with introspective reasoning capabilities with the unconstrained optimization of behavior performance.

1.2.3.9.3 The role of NEEMs in the mastery of manipulation actions The third key insight is the role that episodic memories, and in particular NEEMs could play, for the capabilities of cognition-enabled robot control systems. NEEMs were already put as one of the key ideas in the original proposal for the first phase. However, this role and the potential of large collections of episodic memories has even strengthened in the first phase. EASE has succeeded in standardizing the NEEM formats across different kinds of agents, environments, and tasks, be it in virtual or real environments. In addition, we have NEEMs fully formalized using the EASE ontology, which provides robot agents with many different ways of accessing information and knowledge from NEEMs in automated ways as the ontology makes the NEEM data structures machine interpretable. It is also important to pinpoint the differences between NEEMs and the episodic memories supposed to be used in human cognition. Unlike episodic memories of humans which are very partial and approximate in nature NEEMs are complete and very detailed. This is an advantage of NEEMs because they allow the learning of many different tasks using the same experiences. But it will also result in episodic memory coctions that are very resource consuming and require deeper investigations of memory compression and forgetting. CRAM— even more so than SOAR (Nuxoll and Laird, 2012), is the cognitive architecture that incorporates the most sophisticated reasoning and learning infrastructure based on episodic memories. A key research focus of the second phase of EASE will be to investigate the acquisition of commonsense and intuitive physics knowledge using a combination of NEEMs together with other knowledge sources.

1.2.3.9.4 The power of ontologies Another key insight regards the power of ontologies. Knowledge representation researchers broadly acknowledged the effect of using ontologies on the impact of reasoning applications: *"For the first time since the advent of KR research, the last decade saw the development, and, more important, wide acceptance of international standards for describing data and ontologies on the Web and for reasoning with them. These advances lead to broad availability of structured data in standard formats for KR researchers to user and consume."*(Noy and McGuinness, 2013)

This is even more so for the reasoning of robot agents and the data driven investigation of everyday activities. One key application of the EASE is making NEEMs machine-understandable. Another application in EASE is to make the different vocabularies in which software components of a robot agent are implement — the perception, the motion, and the reasoning and planning executive. They all use different conceptualizations but have to exchange knowledge and information between them which is facilitated by the common ontology. In addition, addition the ontology supports connecting different cognitive science models with each other by connecting the concepts of the different models to shared data structures.

1.2.3.10 Self-evaluation

The collaborative research center EASE is a joint interdisciplinary research center based at University of Bremen. In EASE researchers from the areas robotics, artificial intelligence, cognitive science, machine learning, computer program verification, human computation, and virtual reality work together to advance our understanding of how to build generative models of mastering human-scale everyday manipulation tasks and realize robot agents that accomplish these tasks autonomously. EASE promotes research cooperation within the framework of interdisciplinary research programmes by having all researchers working on common scenarios, the EASE robot days, by using NEEMs as a commmon data structure for recording everyday activity episodes, using a common ontology for making data machine understandable, by committing to common software frameworks and adapting the frameworks to research needs, and by organizing milestone events in which integrated software systems composed of modules investigated in different subprojects are integrated, tested, and demonstrated. In addition, EASE has established and maintains several physical and virtual research labs including the EASE central research laboratory, the Biosignals Acquisition Space and Environment (BASE) together with open software infrastructure that facilitates the efficient use of these research facilitates. The convergence of research activities towards the EASE vision is promoted through scientific events including symposia (DGR Days 2017/EASE Inaugural Symposium: Everyday Activity Science and Engineering, EASE Symposium @ Humanoids 2020) and conference workshops.

Within the EASE research center interdisciplinary team building is promoted through regular meetings as well as through the yearly EASE fall schools on “Cognition-enabled robot manipulation”. International visibility is promoted through the EASE web presence, which includes a blog for research progress and through focussed cooperations with internationally leading research sites (including University of Tokyo, Seoul National University, Edinburgh Robotics Centre). Taken together these measures enable EASE researchers to pursue ambitious, elaborate and long-term projects by focusing and coordinating the resources of EASE and the University of Bremen.

1.2.3.10.1 Context of research assessment

1.2.3.10.1.1 Positioning in the research field EASE is positioned in the intersection of creating robot agents mastering everyday manipulation tasks and understanding cognitive capabilities of the brain. This is because building a generative model for accomplishing human-scale everyday manipulation tasks is like finding a needle in a haystack and the human brain is the only computational system we know that has managed to meet this challenge (Hassabis et al., 2017). As this research area is the melting pot of a number of scientific and engineering disciplines with very diverse approaches and substantial economic interest we characterize the state of understanding by putting the spotlight on exemplary research activities that push interesting, relevant, and challenging ideas rather than aiming for a complete and fair coverage of activities.

The understanding of the brain has been the recent research focus for a number of large-scale research programmes. The brain is the most sophisticated, complex, and complicated structure under scientific investigation. Because of its complexity research projects have to focus on different aspects and perspectives.

Large-scale research enterprises that have focused on understanding the brain include, among others (Shurtleff et al., 2013; Azrieli, 2011) the *Human Brain Project (HBP)* (Markram, 2006, 2013), which aims at building a very complex and detailed computer simulation of the brain with one of the research focus area being the building of robot bodies for the brain simulation. The *Center of Brain, Minds, and Machines* at the Massachussetts Institute of Technology focusses on a computationally based understanding of human intelligence and establishing an engineering practice based on that understanding. Important links are the concentration on a challenge — the visual Turing test, the evolution of intelligence in early childhood, and methodologically the combination of methods, here machine learning, probabilistic reasoning and learning, and symbolic computation, and the focus of intelligent systems on building models of what they do and can do. Another interesting approach put forward by Eliasmith et al. (2012) is the *building of a fully functional brain*, which combines cognitive capabilities to accomplish tasks that require perception and action. Here, generalized artificial neural structures, called Semantic Pointer Architecture, are investigated that can be leveraged for a variety of cognitive capabilities (Blouw et al., 2016).

Besides the focussed research intiatives whole research fields have advanced our understanding of the brain in a systematic, broad, and sustained manner. One of them is *cognitive neuroscience*, which studies the biological processes and aspects that underlie cognition, with a specific focus on the neural connections in the brain which are involved in mental processes. It addresses the questions of how cognitive activities are affected or controlled by neural circuits in the brain (Tibbetts, 2009). Concepts investigated in cognitive neuroscience include cognitive functions such as perception, decision making, prospection, memory, and attention. Another multi-disciplinary research field that is highly relevant for EASE is *action science* (Prinz et al., 2013) which investigates theoretical and methodological approaches to action and the relationship of action and cognition. The field is based on the hypothesis that evolution has optimized cognitive systems to serve the demands of action. The field studies the relation of action to cognitive functions such as perception, attention, memory, and volition. A research initiative in the intersection of cognitive neuroscience and action science is the *situation model framework of cognitive behavior*. This research direction is initiated and advanced by Schneider et al. (2020) establishing a temporary ZiF Research Group “Cognitive behavior of humans, animals, and machines: Situation model perspectives” with participation from EASE researchers. The goal of the research group is to understand how cognitive behavior of humans, animals, and machines with its key features of flexibility and context-sensitivity are realized at the functional and mechanistic level (Schneider et al., 2020). Findings, theories, models, and implementations of cognitive neuroscience (e.g., humans, rodents, and non-human primates) and artificial intelligence (e.g., autonomous robots) guide this endeavor.

The second pillar of EASE is the design and implementation of rational artificial agents that can accomplish complex, human-scale tasks that require human-level intelligence (Russell and Norvig, 2010; Poole and Mackworth, 2010). To some degree this type of research activity is substantially driven by companies, including DeepMind, openAI, and IBM. Currently, the machine learning approach to solving tasks that require intelligence has gained a lot of momentum. Both DeepMind with learning to play Atari games (Mnih et al., 2015) and creating a computer system that beats the champions in Go (Silver et al., 2016) and similar board games (Schrittwieser et al., 2019) as well as openAI defeating the Dota 2 world champions in a Dota (Berner et al., 2019), a multiplayer online battle arena video game with continuous multi-player action. These are examples of generative action models for cognitively very challenging tasks in which computer systems outperform humans. Another promising research direction are autoregressive language models that use deep learning to produce human-like text, such as the Generative Pre-trained Transformer 3 (GPT-3) (Brown et al., 2020). These models gain their power by learning effectively in unbelievably high dimensional parameter spaces. The potential impact of this direction of research goes far beyond language processing because computer programmer can be formulated as a form of text generation. The key idea is to learn problem solving competence from distributions of training data as input-output functions, where the internal computation processes are opaque and where it is in general not clear how the systems behave outside the distributions of training data. Another category of artificial intelligence systems that are highly relevant for EASE are *open question answering systems*, such as the Watson system (Ferrucci et al., 2010) or the Siri agent on smart phones (Myers et al., 2007). These systems are designed to instantaneously answer a large variety of questions, such as possible questions of a quiz show, or solve digital assistance tasks.

Relevant research fields include *robot control*, which is the engineering discipline that sudies how to move articulated physical bodies in order to achieve motion goals, satisfy motion constraints, and optimize motion objectives (Siciliano and Khatib, 2016). *Artificial intelligence* studies problem-solving methods including heuristic search, knowledge representation and reasoning, probabilistic problem-solving, and machine learning in order to design and realize computer systems that solve problems with optimized performance (Russell and Norvig, 2010; Poole and Mackworth, 2010). The aspect that current AI systems solve indidual intelligence tasks but lack the capabilities of switching instantly between intelligence tasks and use cognitive mechanisms for multiple intelligence tasks is put forward in the field *Artificial General Intelligence* (Fridman, 2020; Lake et al., 2016). Finally, we want to include the field of *artificial cognitive systems* that studies the cognitive capabilities, how the can be realized, and how they can be orchestrated using cognitive architectures in order to build artificial cognitive systems that accomplish complex tasks, such as robot agents (Vernon, 2014).

1.2.3.10.1.2 Why the EASE perspective is necessary In order to understand why the EASE perspective on cognitive agency is necessary for a wholistic understanding of how generative models accomplish manipulation tasks we have to

- identify the essential research questions that are tackled by EASE but not by the other enterprises,
- identify the transformative impact that having answers to these questions would have on robot agents accomplishing human-scale manipulation tasks, and
- explain why the EASE collaborative research center and the scientific approach is a promising approach to meet the research challenges and overcome the barriers towards successfully answering these questions.

To get an intuition and assess what the perspectives, potentials and limitations of individual research approaches one can try to map the different approaches to the rational reconstruction of an intelligent system that we have a better understanding of — the Watson system. We choose the Watson system because the Watson system solved an intelligence tasks that surprised the expert community and because there was a substantial discussion in the research and technology community of why the Watson system succeeded and how easily and robustly the methods can be generalized to transfer them to other intelligence tasks.

The Human Brain Project would aim at rationally reconstructing the Watson system by designing models of chips and their connections as realistically and detailed as possible. MRI would measure the heat on the chips to get more insights about what computations might take place. Deep learning approaches would find ways to characterize categories of question answering, intelligently extract lots of training data from the Watson system and perhaps turn it cleverly into a reinforcement learning problem to optimize performance, and so on.

In contrast, some of the most important reasons why the Watson system worked for answering Jeopardy! questions are of a very different nature. They have to do with having a thorough understanding of the computational problems that the Watson system is to solve. They also have to do with which data and knowledge structures the researchers have chosen and how to structure and constrain computational processes such that computational tasks that are in their general form infeasible become efficiently solvable. Examples of such reasons are that (*) If you type in a jeopardy question in a web search engine than it is likely that some top-ranked web pages returned contain the answers, or (*) answering these questions can be phrased as a hypothesize and test strategy where hundreds of hypotheses are generated and tested, etc.

It is not conceivable that evolution, optimizing the brain for rational agency mastering such complex varieties of tasks, did not have to invent analoguous computation structures to be successful. Indeed, different research fields hypothesize such powerful mechanisms such as episodic memories, imagistic and simulation-based reasoning, the use of schemata as key cognitive capabilities of humans. Answering such research questions is the domain of computer science and artificial intelligence that investigate the design and understand of computational processes in order to solve given computational tasks. These types of questions are not sufficiently addressed by the other research approaches listed before.

EASE conducts research in the areas of AI-based and cognition-enabled robotics and robot (physically embodied) agents inspired by investigations of how humans accomplish and master their everyday activities. The vision of EASE matches one of the ultimate goals of artificial intelligence, namely to realize robots that exhibit the competence to accomplish human-scale tasks in the physical world.

This driver for artificial intelligence research started in the very beginning and was for the first time put into a comprehensive research project through the realization of the Shakey robot (Nilsson, 1984b; Kuipers et al., 2017a) and the investigation of the mechanisms for controlling it. The realization of a complete robot agent that could successfully manipulate objects in the physical world had profound impact on the research landscape for many decades to come. Unfortunately, there were very few attempts to build mobile robot agents with arms and robot hands that could truly autonomously accomplish manipulation tasks in real environments. There were a handful hotspots in Europe, perhaps most prominently the ARMAR robots at University of Karlsruhe, focussing on kitchen manipulation tasks (Asfour et al., 2019), the Justin robots at DLR, focussing on impressive, dexterous manipulation skills (Dietrich et al., 2016; Wimböck et al., 2010; Borst et al., 2009), and the robots at LAAS/CNRS doing manipulation tasks for human-robot interaction (Alami et al., 2010). At the Italian Institute of Technology the evolution of manipulation capabilities inspired by the cognitive development in early child hood is the research target (Bartolozzi et al., 2017; Natale et al., 2016). Internationally, the JSK Laboratory at the University of Tokyo realized generations of autonomous robot agents performing complex manipulation tasks (Murooka et al., 2017). Perhaps most impressively, was the group of humanoid robots collectively performing household chores. Another early and notable development was the Chip robot at the University of Chicago, which performed simple object collection tasks but had a a control machinery leveraging AI technology that gave it an impressive level of flexibility, robustness, and generality (Firby et al., 1996). Besides Shakey, the most impactful activity was the development of the PR2 robot by Willow Garage (Marder-Eppstein et al., 2010). Here, a privately funded company developed a two-armed mobile manipulation platform together with an open-source middleware and open-source libraries for perception, motion planning and simple manipulation. They gave 12 robots to internationally leading research labs investigating component technologies and aimed at building a global research network and community pushing together the state-of-the-art intelligent manipulation robots. In the private sector, Google bought several leading-edge technology start-ups and hired outstanding researchers from universities aiming at personal manipulation robots, an enterprise that was not sustained. Finally, the RoboCup competitions, in particular RoboCup@Home and RoboCup@Work aimed at achieving measurable progress through competitions. Here, the focus of research activities was the performance at competitions rather than the advance of the state of understanding.

The strategical importance of understanding generative models of robot behavior that can achieve human-scale manipulation tasks in open real-world environments has been pushed in many invited talks and is present in almost all roadmaps for artificial intelligence and AI-based technologies. In his AAAI presidential address “Keep the eye on the prize” emphasized Nilsson (1995a) the importance of building a robot agent “that can accomplish the tasks that we can expect it to accomplish given its sensors and effectors” as a means for progressing in our understanding of artificial intelligence. Regularly, challenge papers resulted from conference workshops (Kemp et al., 2007) and literature reviews (Ersen et al., 2017). The realization of robots with autonomous manipulation capabilities in open real-world environments are essential parts of many national and international research roadmaps including the

H2020 strategic research agenda (SRA) and multi-annual roadmap (MAR), the US robotics roadmap promoting the national robotics initiative, the roadmaps of Australia, Great Britain, and others.

There is an apparent mismatch between the strategical importance and magnitude of the research problem and the number of researchers and research teams working on it, if you compare the field research on component problems such as robot learning, in particular deep learning, robot vision, navigation, motion planning, state estimation, reasoning about action — to name only a few. There are a number of reasons for this situation. Already in the Shakey project the researchers pointed out that only a small fraction of the research work was considered to be publishable. For the Shakey project these components of the research work were the seminal works on A*, Strips planning, and triangle tables for achieving robust execution. Other project results that researchers trying to build robot agents desparately needed were published only more than ten years later as a technical report copying together excerpts of the project reports because they couldn't be published otherwise. Another reason is that building robot agents requires truly confluent science and engineering. Researchers conducting their research in diverse fields with different research cultures and vocabularies have to unite and orchestrate their components in one common system. This is also stated in the catalogue of cognitive systems capabilities of the H2020coordinating action EURobotics2: *"the behaviour required of cognitive robots depends on the coherent operation of a network of non-trivial components. This coherence entails that all componentsare mutually compatible. Consequently, designing a cognitive system is not simply a question of assembling off-the-shelf algorithms:* ***the entire system is designed as a whole***."

EASE positions itself uniquely in the research field of robot agents that can accomplish human-scale manipulation tasks.

- complete control system realized through EASE
- AI techniques, which machine learning is only a part of, premeates the whole control system
- all components of the control system above the action and parception libraries are subjects of peer-reviewed publications in conferences and journals.
- the software components are all accessible open-source
- a large repository of research data including NEEMs of robot experiments are available volume and in many cases thoroughly segmented and semantically annotated.

The EASE research agenda is also in line with current research trends that point on the necessity of combining machine learning and data intensive computation methods with knowledge-enabled methods to reach the next level of capability in intelligent systems. Hassabis et al. (2017) discuss the value of using insights from neuro science research for inspiration of artificial intelligence and in particular machine learning and survey progress and prospects in the fields. Lake et al. (2016) argue for the need to go beyond data intensive machine learning methods in order to accomplish learning and reasoning behavior that can progress towards a human competence level. Marcus and Davis (2019a) argue that a combination of symbolic and data intensive learning and reasoning is needed to avoid that reasoning systems work outside the data distributions they are trained with, to go beyond blackbox operation in important decision making problems, and achieve systems where higher competence levels can be reached through functional competition. Ways in which data intensive machine learning and symbolic reasoning can be synergetically combined are systematically surveyed in the constructive analysis by van Harmelen and ten Teije (2019).

These views are consistent with the EASE view where EASE does not investigate intelligent problem solving in its generality but rather focusses on robot agents that accomplish everyday manipulation tasks. This has the advantages that the evolution of intelligence is tightly coupled to the evolution of manipulation capability and that in manipulation episodes the consequences of reasoning and decision making are more directly observable.

1.2.3.10.1.3 The EASE research agenda in the context of community roadmaps Another dimension of evaluation is the potential of the EASE research agenda in comparison to community research roadmaps. Perhaps, one of the most relevant roadmaps is the one currently under development by the Association for the Advancement of Artificial Intelligence (AAAI) and the Computing Community Consortium (Gil and Selman, 2019). This roadmap sets a 20-year research agenda for which a draft version was published in August 2019.

This research roadmap organizes the research activities for the next two decades into three research themes:

- integrated intelligence, which focusses on (*) integration in order to create intelligent systems that have much broader capabilities than today's systems, (*) contextualization, to adapt general methods to specfic tasks, objects, and environments, and (*) knowledge, to provide access to the vast amount of knowledge that is required to operate in the rich world we live in.
- meaningful interactions by integrating different interaction modalities, in the case of EASE perception-action loops and making intelligent systems context-aware, and being able to justify behavior and conclusions.
- self-aware learning, which focusses on learning expressive representations, making learning trustworthy, lifelong learning, and the integration of AI and robotics.

The substantial overlap and synergies between the EASE research activities and plans and the broader community research area demonstrate huge potential that EASE can have a substantial and lasting impact on the field.

1.2.3.10.2 EASE Resources Figure 1.50 shows the logical relationships among the resources that are invested in EASE, the activities that take place, and the benefits or changes that result from these activities. The funding of EASE finances 12 research subprojects with a total of 25 researchers (doctoral students and postdoctoral researchers). For technical support one position for maintaining the laboratory and software management and 1 position for data management. The administrative support is 1 position for managing and controlling the financial affairs and 0.5 position for managing the EASE integrated research training group. In addition the university receives overhead resources for the increased cost of operating a collaborative research center. Additional resources are contributed by the EASE principal investigators, cooperating institutions, the university, and research and innovation projects acquired by EASE principal investigators and contributing to the EASE research vision.

These financial resources facilitate the core EASE activities in basic research, building an institutionalize research organization, teaching and training, and transfer of research results into innovation and outreach into the general public. These activities produce outputs, which we structure into publications, open-source software components, open data sets and knowledge bases, scientific events, awarded doctoral degrees, and innovation activities. After the first phase the outcomes of primary interest are a strengthened position in the research landscape, better training and education of experts in cognition-enabled robot agents, a sustainable impact on the research landscape of the university, and a transfer and innovation infrastructure based on EASE research.

1.2.3.10.2.1 Publications. The output of the first phase of EASE until the end of 2020 is listed in Table 1.1.

[57] top conferences and journals in robotics and AI according to AMiner

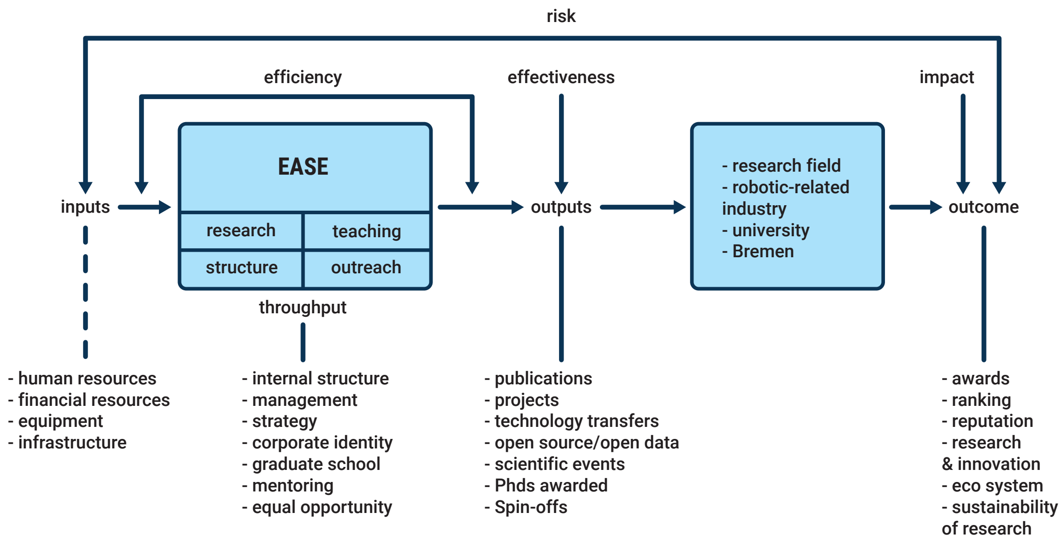


Figure 1.50: EASE framework for self-evaluation.

Category	Number of Publications
# peer reviewed publications with EASE acknowledgement	$>$ 130
# publications at top conferences and journals in robotics and AI[57]	44
# publications with co-authors from more than 1 EASE research group	24
# publications with co-authors from more than 1 EASE subproject	18
# publications with co-authors from different research areas	8
# publications with co-authors not EASE members	45
# publications with international co-authors	24
# open access publications	85

Table 1.1: Number of publications in the first phase of EASE by the end 2020.

1.2.3.10.2.2 Open-source software It is part of the mission of EASE to improve the sustainability of research software, facilitate research cooperation through shared software, and increase the transparency of research results through replicability. Therefore, a substantial part of the EASE research software is made publicly available as open-source software components.

Software components of the generative model/cognitive architecture.

- **CRAM language and plan executive** (*webpage:* `http://cram-system.org`, *main developer:* Gayane Kazhoyan). CRAM is a toolbox for designing, implementing and deploying plan-based control software on autonomous robots.
- **KnowRob KR&R system** (*webpage:* `http://knowrob.org`, *main developer:* Daniel Beßler). KnowRob is a knowledge processing system that combines KR&R methods with techniques for acquiring knowledge and for grounding the knowledge in a physical system and can serve as a common semantic framework for reasoning about robot agency. KnowRob combines static encyclopedic knowledge, commonsense knowledge, task descriptions, environment models, object information and information about observed actions that has been acquired from various sources (manually axiomatized, derived from observations, or imported from the web). It supports different deterministic and probabilistic reasoning mechanisms, clustering, classification and segmentation methods, and includes query interfaces as well as visualization tools.

- **RoboSherlock taskable robot perception system** (*webpage:* `http://robosherlock.org`, *main developers:* Patrick Mania, Franklin Kenfack, Ferenc Balint-Benczedi). RoboSherlock is a software framework for cognitive perception in robot agents.
- **Giskard cognition-enabled motion control framework** (*webpage:* `http://giskard.de`, *main developers:* Georg Bartels, Simon Stelter). Giskard is a software framework for cognitive motion control in robot agents.
- **RobCoG ROBot COmmonsense Games** (*webpage:* `http://www.robcog.org`, *main developer:* Andrei Haidu): RobCoG aims at acquiring commonsense and intuitive physics knowledge for robot agents using games with a purpose. In the games users are asked to execute various tasks in different scenarios. The games are equipped with a semantic logging system which captures and stores symbolic and sub-symbolic data during the gameplay.
- **openEASE open knowledge service of EASE** (*webpage:* `https://open-ease.org`, *main developer:* Daniel Beßler). openEASE is a web-based open knowledge service for storage, analysis, and visualisation of EASE research data acquired from robots and humans performing everyday activities. The main interface of openEASE is a query answering component that provides uniform access to heterogeneous research data through abstraction into a common ontological representation.
- **PRAC Probabilistic Action Cores** (*webpage:* `http://www.actioncores.org/`, *main developer:* Daniel Nyga). PRAC is an interpreter for natural-language instructions for robot applications. The PRAC system aims at making knowledge about everyday activities from websites like wikiHow available for service robots, such that they can autonomously acquire new high-level skills by browsing the Web. PRAC addresses the problem that natural language is inherently vague and unspecific. To this end, PRAC maintains probabilistic first-order knowledge bases over semantic networks represented in Markov logic networks, which is accessible through `http://www.pracmln.org`.
- **NEEM-hub** (*webpage:* `https://neemgit.informatik.uni-bremen.de/explore/groups`, *main developer:* Sebastian Koralewski, Asil Bozcuoglu). The NEEM-hub acts as a central data storage and management system for the EASE project. EASE researchers and cooperating researchers can upload their NEEMs, share them with the community and work with the data on openEASE. The NEEM-hub is already operational but the implementation not yet finished. The NEEM handbook[58] describes how to use the NEEM-hub.

Other open-source software components include:

- **Schemasim**[59]; Mihai Pomarlan
- **EASEHC**[60]; Johannes Pfau
- **RAFCON**[61]; Sebastian Brunner
- **Web-OpenCCG**[62]; Sebastian Höffner
- **IntuitionSimulation**[63]; Sebastian Höffner

Open data and knowledge repositories

- **SOMA ontology** (*web page:* `https://ease-crc.github.io/soma`). The SOMA ontology is an ontological model of everyday activites that has been collaboratively developed in EASE. The main focus of SOMA is the aximatization of physical and social activity context, the motions generated by agents, and the effects caused by them. Currently, SOMA defines 371 general classes, and

[58] `https://ease-crc.github.io/soma/owl/current/NEEM-Handbook.pdf`
[59] `https://github.com/mpomarlan/schemasim`
[60] `https://github.com/JohannesPfau/EASEHC`
[61] `https://dlr-rm.github.io/RAFCON/`
[62] `https://github.com/shoeffner/web-openccg`
[63] `https://gitlab.informatik.uni-bremen.de/ease-ph/dlu/IntuitionSimulation`

2405 axioms. These definitions serve as a foundation for defining more specific models tailored to a specific agent, task or environment.

- **KnowRob robot ontology** (*web page:* `http://knowrob.org`). This ontology is an extension of SOMA including definitions related to the robotics field which are used to characterize robot agents and data structures and computation processes in their control programs. Robot agents can use this ontology to equip themselves with a self representation including how it is composed of hardware and software components, and what capabilities it can derive from them. This ontology currently defines 311 concepts, and 1271 axioms.
- **Failure and Recovery ontology** (*web page:* `https://github.com/ease-crc/failrecont`). FailRecOnt is an ontological model of execution failures and recovery strategies developed in collaboration with the Polytechnics University of Catalunya and the Institute for Cognitive Science and Technology of Trento. It defines 69 general classes, and 288 axioms intended to capture high-level failure categories and how they relate to situations and tasks, along which dimensions failures can be causally explained, etc. This axiomatization of failures enables reasoning to combine information coming from several robot modules into a narrative explaining a failure, and in articulating plausible goals for recovery actions.
- **EASE lexical resources** (*web:* `https://github.com/ease-crc/ease_lexical_resources`). This knowledge resource includes information about morphemes and semantics of English words combining information from various open-access lexical resources into a uniform format. Some of the information was collected in EASE Subproject P01 by studying corpora of cooking instructions. Currently, this resource has information about circa 58000 lemmas. The project includes scripts for parsing various linguistic resources and can generate lexicons for KPML (an NLG production system) and, partially, for construction grammar parsers.
- **Retail store ontology** (*web page:* `https://github.com/refills-project/knowrob_refills`). The retail store ontology extends SOMA with definitions that characterize the structure of retail stores, types of products, and tasks that are commonly performed in retail stores. One of the use cases for this ontology is autonomous semantic mapping where the mapping process is restricted through axioms that constrain the structure of the store. Right now, this ontology comprises of 245 classes, and 1280 axioms.
- **SOMA-Home ontology** (*web page:* `https://github.com/ease-crc/soma`). SOMA-Home is an extension of SOMA specifically designed for everyday activities in household environments. So far, the ontology covers the EASE kitchen environment, which is used in the EASE experimental setups. Extensions for other environments are created, and will be integrated into the SOMA-Home ontology in the near future.
- **EASE robot ontologies** (*web page:* `https://github.com/knowrob/knowrob`). The EASE robot ontology collection is a set of independent ontological modules that each represents a specific category of robots. So far, this collection contains 13 representations of robot models.
- **NEEM collections** The narrative-enabled episodic memories (NEEMs) are an essential knowledge source for developing the manipualtion and cognitive capabilities of robots. Over the first funding period, we have collected around 4 TB of data which consists of about 2500 episodic memories. A large subset of these NEEMs is already available on our NEEM-HUB[64]. The missing data is currently uploaded and will also be made available for the research community. A detailed overview about the individual NEEMs is provided in 1.36.

1.2.3.10.2.3 Scientific events In addition to the EASE Fall schools, which are primarily training events for doctoral students and discussed in Section 1.3.1.3.1.3, EASE has organized several scientific symposia:

- **Cognition-enabled Robotics: Democratising a Disruptive Technology** (29. Bremer Universitäts-Gespräche, November 2016, *weppage:* `https://www.uni-bremen.de/bug/bug-2016/`, *or-*

[64] `https://neemgit.informatik.uni-bremen.de/explore/group`

ganizers: Michael Beetz, Andreas Birk (Jacobs University Bremen).
invited speakers: Yiannis Aloimonos (University of Maryland), David Lane (Heriott-Watt University), Gordon Cheng (Technical University Munich), Tony Belpaeme (Plymouth University), Gregory O'Hare (University College Dublin), Alin Albu-Schaeffer (DLR), Markus Vincze (TU Vienna), Herman Bruyninckx (KU Leuven), Georg von Wichert (Siemens AG), Amos Albert (Bosch Deepfield Robotics)

- **DGR Days 2017/EASE Inaugural Symposium: Everyday Activity Science and Engineering.** *organizers:* Tamim Asfour (KIT), Michael Beetz, Karsten Berns (University of Kaiserslautern), Wolfram Burgard (University Freiburg).
  invited speakers: Sven Behnke (University Bonn), Oskar von Stryk (TU Darmstadt), Jessica Burgner-Kahrs (University Hannover), Torsten Kröger (KIT), Giulio Sandini (Italian Institute of Technology), Markus Vincze (TU Vienna), Frank Kirchner (DFKI), Thomas Schack (CITEC), Kei Okada (University of Tokyo), Alessandro Saffiotti (Orebro University).
- **SCORE workshop in Bremen** The Score 1:0 meeting is the kickoff event of the SCORE project, held at the University of Bremen between July 24-25, 2020. *Invited speakers:* Oliver Kutz (Bolzano), Maria Hedblom (Bolzano), Pietro Galliani (Bolzano), Fabian Neuhaus (Magdeburg), Robert Ross (Dublin), Fumiaki Toyoshima (LOA Trento / JAIST Institute, Japan).
- **EASE 2019 Milestone Symposium,** Symposium of the **Hanse Wissenschaftskolleg Focus Group "Cognition-enabled Robotic Agents"**. (*organizers:* Michael Beetz, Hagen Langer).
- **EASE Symposium at the 2020 IEEE-RAS International Conference on Humanoid Robots.** (*organizers:* Alin Albu-Schäffer, Michael Beetz, Gordon Cheng, Helge Ritter)
  Unfortunately, the conference and therefore the symposium was cancelled due to the Covid situation.

The symposia were complemented with several **co-organized scientific workshops on EASE research topics at international conferences.** These workshops include:

- **ICRA 2018 Workshop on Cognitive Whole-Body Control for Compliant Robot Manipulation (COWB-COMP)** Daniel Leidner (DLR), Alexander Dietrich (DLR), Michael Beetz (UB)
- **Language and Robotics** Takato Horii (Japan), Emre Ugur (Turkey), Tadahiro Taniguchi (Japan), Xavier Hinaut (France), Tetsunari Inamura (Japan), Takayuki Nagai (Japan), Michael Spranger (Japan), Michael Beetz (Germany)
- **Towards Robots that Exhibit Manipulation Intelligence** Michael Beetz, Georg Bartels (Germany), Oussama Khatib (USA), Alin Albu-Schäffer (Germany), Marc Toussaint (Germany)
- **Semantic Policy and Action Representations for Autonomous Robots** Eren Erdal Aksoy (Sweden), Yezhou Yang (USA), Karinne Ramirez-Amaro, Neil Dantam (USA), Gordon Cheng
- **Latest Advances in Big Activity Data Sources for Robotics and New Challenges** Asil Bozcuoglu (Germany), Tamim Asfour (Germany), Karinne Ramirez-Amaro (Germany), Gordon Cheng (Germany)
- **ECAI 2020 Workshop on Artificial and Human Intelligence: On Formal and Cognitive Foundations for Human-Centred Computing** Mehul Bhatt (Sweden)
- **IJCAI-ECAI 2018 Workshop on Cognitive Vision – Integrated Vision and AI for Embodied Perception and Interaction** Mehul Bhatt (Germany), Alessandra Russo (UK), Parisa Kordjamshidi (USA)
- **CogSci 2019 Workshop on Everyday Activities** Holger Schultheis (Germany), Richard P. Cooper (UK)
- **FOIS (Formal Ontology in Information Systems) 2020 RobOntics (Ontologies for Autonomous Robots Workshop)** Stefano Borgo, Aldo Gangemi, Robert Porzel, Mihai Pomarlan, Daniel Beßler, Mohammed Diab, Alberto Olivares-Alarcos

1.2.3.10.3 Assessment of research quality In order to assess the quality of the EASE research conducted in the first phase, we (1) list selected representative publications that we expect to have substantial impact, (2) quote several expert opinions on EASE research, (3) analyze the citations of EASE research and research that EASE is based on, (4) the long-term impact of EASE researchers.

1.2.3.10.3.1 Selected publications 8 EASE publications in the first funding period received awards or were finalists for awards with two additional awards obtained by EASE researchers:

- EASE researcher and designated 2nd phase principal investigator Daniel Leidner (DLR, Institut für Robotik und Mechatronik / Universität Bremen) won the Georges Giralt PhD Award selected from all robotics-related dissertations which have been successfully defended at a European university in 2018 for his dissertation thesis entitled "Cognitive Reasoning for Compliant Robot Manipulation".
- Distinguished Paper Nomination with Honorable Mention, IJCAI 2019. Jakob Suchan, Mehul Bhatt, and Srikrishna Varadarajan (Suchan et al., 2019)
- Four EASE conference papers were finalists for best papers at leading international robotics and AI conferences (IROS 2017 (Stelter et al., 2017), ICAR 2017 (Balint-Benczedi et al., 2017), ICRA 2018 (Bozcuoglu et al., 2018), AAMAS 2018 (Beßler et al., 2018c)).
- Carsten Lutz was Best Paper Awards Runner-Up at KR2020 (Jung et al., 2020b).
- Finalist for best cognitive robotics paper at IROS 2020: Uhde et al. (2020): "The Robot as Scientist: Using Mental Simulation to Test Causal Hypotheses Extracted from Human Activities in Virtual Reality"
- The EASE conference paper "KnowRob2.0 — a 2nd generation knowledge processing framework for cognition-enabled robotic agents" (Beetz et al., 2018) is on the list of most important publications over past 3 years in the area of cognitive robotics selected by the IEEE RAS Technical Committee on Cognitive Robotics.
- Zhen Zeng, Adrian Röfer and Odest Chadwicke Jenkins received the best Cognitive Robotics award at IROS 2020 with the publication "Semantic Linking Maps for Active Visual Object Search" in 2020, where Adrian Röfer is a master student, for whom EASE supported the guest semester at the University of Michigan and Beetz being member of the dissertation committee of Zhen Zeng.
- Best paper award ICANN2020. "A neural network architecture to map cluttered object geometry into contact graphs" (Meier et al., 2020).

EASE researchers have also been invited to contribute to encyclopedias and handbooks in cognitive robotics:

- David Vernon and Michael Beetz are authors of the chapters "Cognitive architectures" and "Knowledge representation and reasoning for cognitive robots" of the book *Cognitive Robotics*, MIT Press, to appear, edited by Angelo Cangelosi and Minuro Asada.
- Michael Beetz and Daniel Nyga are authors of the chapter "Knowledge representation and reasoning for robotic agents" of the book *Robotics Goes MOOC*, which is a comprehensive reference book on robotics with MOOC supplement to be published by Springer.
- David Vernon, Giulio Sandini, and Alessandra Sciutti are authors of the entry "Cognitive Robotics" in the *Encyclopedia of Robotics* to be published by Springer, edited by Marcelo Ang, Oussama Khatib, and Bruno Siciliano, in the section on Personal and Cognitive Developmental Robotics, edited Minoru Asada. Springer describes the *Encyclopedia of Robotics* as a "one-of-a-kind and exhaustive major reference work".

Several publications go beyond individual research contributions and present the larger context of the EASE research. These publications include:

- The EASE article "Purposive learning: Robot reasoning about the meanings of human activities" by Cheng et al. (2019) was published in Science Robotics in January 2019 with three of four authors being EASE researchers (Gordon Cheng, Karinne Ramirez-Amaro, Michael Beetz and Yasuo Kuniyoshi).

There are two comprehensive review articles:

- EASE researchers were among the primary authors of a review on ontologies for autonomous robotics entitled “A Review and Comparison of Ontology-based Approaches to Robot Autonomy” (Olivares-Alarcos et al., 2019). This work has been conducted in the context of the IEEE-SA P1872.2 Standard for Autonomous Robotics (AuR) Ontology, originated as a sub-group of IEEE WG Ontologies for Robotics and Automation (ORA). The website corresponding to the review article is hosted on the EASE website.

- Ramirez-Amaro et al. (2019c): “A Survey on Semantic-based Methods for the Understanding of Human Movements”. Robotics and Autonomous Systems. Vol. 119, pp. 31-50, Sept 2019. Elsevier.

Three special issues in journals have been co-edited by EASE researchers:

- Holger Schultheis, Richard P. Cooper: Everyday Activities. Topics in Cognitive Science (topiCS), which is one of the two premier outlets for innovative research and theory of the Cognitive Science Society. topiCS publishes collections of papers that focus on new and emerging topics or topics which are a bit off the mainstream, but of broad interest.
- Georg Bartels, Michael Beetz, Martin Ruskowski: Special Issue “Smart production”, Künstliche Intelligenz, and
- Section focused on machine learning methods for high-level cognitive capabilities in robotics Tetsunari Inamura, Hiroki Yokoyama, Emre Ugur, Xavier Hinaut, Michael Beetz, Tadahiro Taniguchi Advanced Robotics 33(11) 537 - 538 2019.

Two collections of book chapters were contributed to books published by Springer books were

- Rolf Drechsler and Cornelia Grosse have published a book on “Information storage”, which also features a chapter on NEEMs.
- Kirchner et al. (2020) published a collection of articles on “AI Technologies for underwater robots”, in which 5 EASE PIs are co-authors of book chapters.

The following publications are noteworthy because they required because they were only possible through intense cooperation of different groups and even research areas in the EASE research center:

- Foundations of the Socio-physical Model of Activities (SOMA) for Autonomous Robotic Agents. Beßler, Porzel, Pomarlan, Vyas, Höffner, Beetz, Malaka, and Bateman. *Proposes the core components of the EASE ontology and a cooperation result across the three research areas P, R, and H.*
- A peer-reviewed conference publication of a research area: “From Human to Robot Everyday Activity” by Mason et al. (2020) presents the framework for interpreting and modeling human everyday and resulting from the cooperation of research area H.
- The Robot Household Marathon Experiment[65] integrates research results from the research areas R and P. The research projects from area R realized the control program for the experiment and area P contributed the knowledge processing infrastructure, including the SOMA ontology that allowed to create NEEMs which were used for learning and improving the robot’s performance.

[65] https://www.ease-crc.org/link/video-ease-robot-day

1.2.3.10.3.2 Expert analysis

- Gary Marcus, co-author of the book "Rebooting AI: Building Artificial Intelligence We Can Trust" (Marcus and Davis, 2019a), co-founder of Robust.AI, and Department of Psychology at New York University: *"The framework for cognitive robotics that Michael Beetz is developing is a significant step beyond anything else I have seen — both in terms of what it has accomplished and in terms of the precision of the underlying thinking. It should be ideal for accelerating the deployment of intelligent autonomous robots in an interpretable and effective, human-centered way. As CEO of Robust.AI, a new startup co-founded with robotics legend Rodney Brooks and others, I get to see a lot of interesting, cutting-edge work; Professor Beetz's work stands out as some of the best."*

1.2.3.10.3.3 Presence of EASE research in reviews and special issues Another aspect of impact is the presence of EASE research in review articles on the relevant research topics:

- **3 out of 18** articles in the Artificial Intelligence journal "Special Issue on AI and Robotics" are contributed by EASE researchers (Kunze and Beetz, 2017; Tenorth and Beetz, 2017; Ramirez-Amaro et al., 2017).
- Michael Beetz is the most cited author (**14 out of 134 references**) in the review article "Cognition-Enabled Robot Manipulation in Human Environments: Requirements, Recent Work, and Open Problems" by Ersen et al. (2017).
- EASE researchers are referenced **14 out of 160 references** in the "Survey of Knowledge Representation in Service Robotics" by Paulius and Sun (2019).
- In the "Review and Comparison of Ontology-based Approaches to Robot Autonomy" Olivares-Alarcos et al. (2019) assess KnowRob to be the KR&R framework that covers the ontological topics relevant for autonomous robots.
- CRAM is one of the decision making frameworks discussed in "Deliberation for autonomous robots: A survey." by Felix Ingrand and Malik Ghallab Artificial Intelligence Special Issue on AI and Robotics (Ingrand and Ghallab, 2017).
- Chiatti et al. (2020) refer to KnowRob as, to date, the most comprehensive knowledge base for robots. *KnowRob (Tenorth and Beetz, 2009; Beetz et al., 2018) is, to date, (Paulius and Sun, 2019; Thosar et al., 2018).*
- Thosar et al. (2018) published a review of knowledge bases for service robots in household environments, in which the KnowRob knowledge bases were assessed as the most comprehensive ones and the ones with the highest academic impact.

1.2.3.10.3.4 Long-term impact of principal investigators The **2020 AI 2000 Most Influential Scholars** (`https://www.aminer.cn/ai2000/robotics`) in Robotics are the top 10 most cited scholars from the top venues of this field over the past 10 years (2009–2019). The list is conferred in recognition of outstanding technical achievements with lasting contribution and impact. Inclusion is determined solely based on the Tsinghua AMiner academic data, which indexes more than 133 million expert profiles and 270 million publications.

Five EASE principal investigators are on the AI2000 list (Beetz: rank 4 in robotics, Albu-Schäffer rank 21 in robotics, Lutz rank 30 in Knowledge Engineering & rank 94 AAAI/IJCAI, Drechsler rank 96 in Chip Technology). In addition, three former doctoral students of EASE PIs are on the list (Rusu: rank 2 in robotics, Blodow: rank 22 in robotics, and Tenorth: rank: 87 in robotics).

Guide2Research Ranking for Computer Science in Germany. 5 EASE PIs are among the top 100 researchers in Computer Science in Germany (Helge Ritter (58), Michael Beetz (64), Rolf Drechsler (77), Carsten Lutz (80), Tanja Schultz (81)). The ranking is based on h-index, citations and number of DBLP documents gathered by May 16th 2020. At the level of universities, the Technical University of Munich is ranked as the number 2, University of Bremen as 15, and Bielefeld as 28 of the Computer Science ranking of German universities.

The rankings show the high quality of the EASE research team and its impact over the last ten years, in particular in the section robotics.

1.2.3.10.4 Assessment of research capacity The improvement of research capacity through the EASE collaborative research center will be discussed in Section 1.3.3.

1.2.3.10.5 Effects of the Covid-19 situation

- EASE had planned a big milestone event in which software across different projects and research areas are integrated and we conduct large integrated experiments including physical robots and big collections of human activity data.
  Unfortunately, these activities can only be performed in a limited way using digital cooperation tools. Also more experiments have to be done using simulation infrastructure rather than real systems.
  The milestone event will be realized through digital presentation of EASE research progress in the first phase including research videos, a web presentation, and blog entries for individual research results.
- The recording of NEEMs from robot experiments and human experiments could only be performed in a limited manner. As a consequence, the data sets provided through OPENEASE are smaller than planned.
- EASE has planned to organize an international symposium at the Humanoids 20 international robotics conference (general chair: Gordon Cheng), including sponsoring, and exhibiting research results at the conference site.
  The conference that was planned for December 2020 is postponed and the new date not yet settled. EASE plans to hold the symposium and the exhibition at the conference.
- The EASE fall school 2020 has been cancelled and we intend to continue with the fall schools in Fall 2021.
- Another doctoral training school, the ICAPS-ICRA summer school on robot planning got postponed and will happen as a digital event.
- The implementation of international cooperations is delayed due to travel restrictions. Also research stays had to be cancelled.

1.2.4 Open research, national, and international cooperation

1.2.4.1 Open research infrastructure

An important part of the EASE mission stated in Section 1.2.2.3 is to conduct EASE as an open research project, which means that we make the data acqired, the software implemented, and the formalized knowledge bases accessible for the research community. Inside of EASE this is implemented through the data management project, which has made substantial progress in the homogenization of data, semantically annotating data, storing data in bid data databases, and making them accessible through OPENEASE.

However, making the EASE research more comprehensivle accessible requires a much bigger effort. It requires that the software components have industrial-strength implementations, data structures and application programming interfaces are standardized, systems are downloadable in a ready-to-use form, come with proper and working installation guides, and have accompanying tutorials.

For EASE these additional efforts are partly carried out in the BMWi project Knowledge4Retail, which is to build an open digital innovation platform and an innovation and research ecosystem. Knowledge4Retail proposes to build digital twin knowledge bases of retail stores using leading-edge autonomous robot mapping technologies, and to deploy and operate autonomous service robots in the stores that use the knowledge bases. Thus, Knowledge4Retail uses the KNOWROB knowledge representation and reasoning system, the CRAM plan executive, and the ROBOSHERLOCK perception executive.

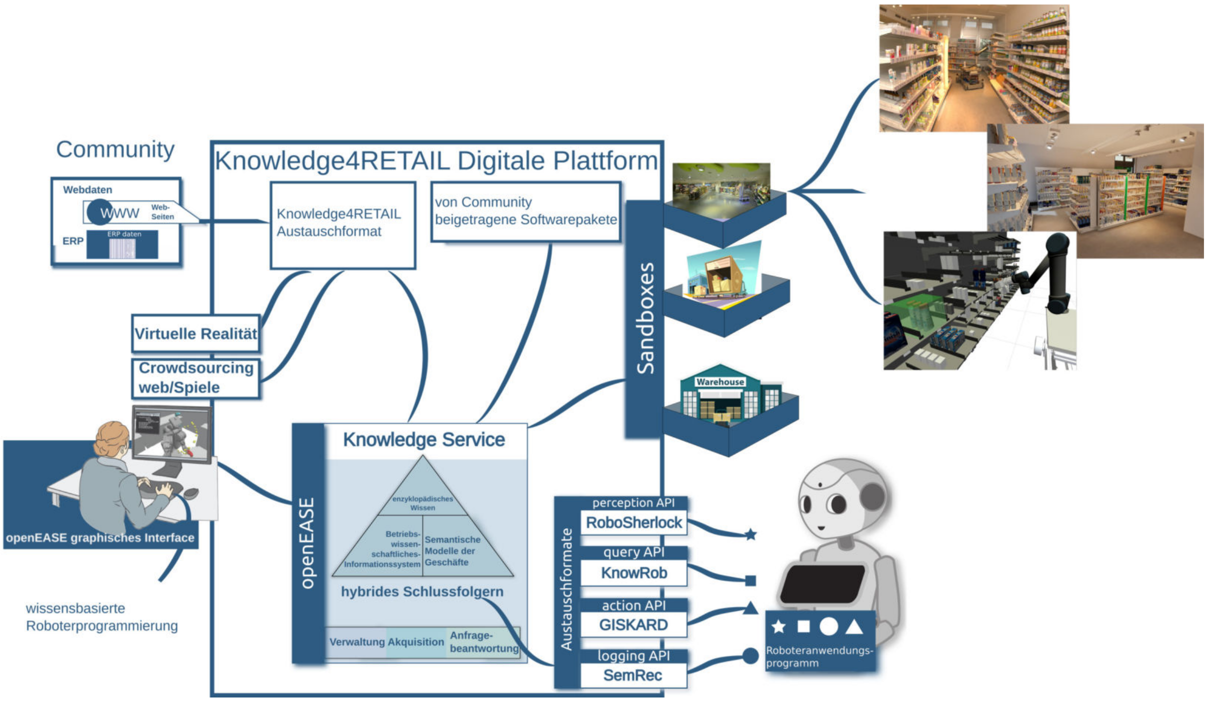


Figure 1.51: The figure depicts part of the KNOWLEDGE4RETAIL digital innovation platform built around the EASE software components OPENEASE, KNOWROB, ROBOSHERLOCK, and GISKARD. We intend to transfer this software infrastructure to build an open research environment for EASE.

Figure 1.51 shows part of the software infrastructure of the Knowledge4Retail digital innovation platform. The software infrastructure runs as a cloud service and allows stakeholders to software experiments in virtual simulation environments of retail stores, the so-called sandboxes.

In the second funding phase of the EASE project we plan to realize the EASE central research laboratory as a sandbox environment using the URoboSim robot simulator. This sandbox environment enables researchers to remotely run experiments in the context of the EASE generative model interfacing their own research software with the CRAM cognitive architecture in a mode that is very close to what is running on the real robots.

1.2.4.2 National Cooperations

Cooperation is already reflected in the composition of the EASE consortium. The principal investigators from the partner institutions: Helge Ritter is the coordinator of the Center for Excellence Cognitive Interactive Technology (CITEC), Alin Albu-Schäffer is the head of the DLR Institute of Robotics and Mechatronics, and Gordon Cheng heads various cognitive neuroscience research initiatives at the Technical University of Munich. In all three cases the cooperation goes far beyond having integrated a single subproject into the EASE collaborative research center.

Researchers at CITEC have established the ZiF Research Group (2019-2020) — a "think tank" for AI and cognitive neuroscience — that brings together an international and interdisciplinary group of researchers from pertinent fields to approach this challenge of cognitive behavior from the conceptual framework of situation models: a situation model details the required processes and the computational space that together connect perception and memory in the service of cognitive behavior. EASE and CITEC intend to further investigate the application of the situation model to cognition-enabled robot agents as an interinstitutional research topic.

The DLR Institute of Robotics and Mechatronics cooperates with EASE on different topics and at different levels. On a software level EASE uses the DLR motion control framework on the Boxy robot and researchers of both institutions work together on robot perception. In the first EASE phase, several doctoral students were jointly supervised at DLR and University of Bremen and received their doctoral degrees from the University of Bremen (Daniel Leidner, Martin Schuster, and Sebastian Brunner). Daniel Leidner has established a DLR junior research group in cooperation with the University of Bremen and will also take the role of an EASE principal investigator in the second funding phase.

DFKI Robotics Innovation Center and EASE cooperate closely in the context of the German AI initiative and aim at establishing the University of Bremen as a leading research institution in the area of cognition-enabled robotics with long-term autonomy. Joint activities include international outreach (TransAIR), the development of digital innovation platform and eco system infrastructure for the retail domain (Knowledge4Retail), and proposals for a competence center for human-robot interaction (CERA4HRI).

The cooperation with TUM strengthens the competence of EASE in cognitive and humanoid robotics and cognitive neuroscience and spans a number of individual cooperations.

1.2.4.3 International cooperations

EASE strives at becoming an international center for cognitively enabled robotics that focusses on hybrid AI technologies for robot agents that are to accomplish human-scale manipulation tasks. To this end, EASE was a consortium member in a proposal of the European robotics community (more than 100 institutional partners in Europe) to establish a "Centers of Excellence Network for Trustworthy Robotics and Intelligent Systems (CENTRIS)," which was unfortunately not accepted for funding. Four institutional partners of EASE were designated partner institutions of CENTRIS (Technical University of Munich (coordinator), DLR, University of Bremen, and University of Bielefeld and two EASE PIs in the designated management board. EASE will continue to look for opportunities for advancing into this direction.

In addition, EASE actively works on the establishment and strengthening of bilateral cooperation with leading international research laboratories with synergistic research agendas. These activities are slowed down through the COVID-19 situation. Partner institutions with which we have and are establishing such cooperations include LAAS/CNRS, Orebro University, Vrije University of Amsterdam, TU Vienna, University of Costa Rica, University of Tokyo, Seoul National University, National University of Singapore, Mahidol University, and Michigan Robotics Institute:

- **LAAS-CNRS (Rachid Alami):** This cooperation has been active for more than 20 years and focusses on knowledge representation and reasoning for robot agents, the combination of task and motion planning. The cooperation included joint supervision of double-degree doctoral students, cooperation in EU cognitive systems projects, exchange of researchers (post docs & doctoral students), and joint publications. In 2021 we will start the trilateral (France, Japan, and Germany) research project AI4HRI (Artificial Intelligence for Human-Robot Interaction), which aims at advancing the cognitive capabilities of robots needed for human-robot interaction.
- **Orebro University (Alessandro Saffiotti, Amy Loutfi, and Achim Lilienthal):** This cooperation is firmly based on Michael Beetz holding a honary doctoral degree from Orebro University. The cooperation spans a broad range of research topics in cognitive robotics and the envisioned modi of cooperation will include cooperation in EU cognitive systems projects, joint organization of research events, and exchange of researchers (post docs & doctoral students). The implementation of the cooperation is delayed through the Covid situation.
- **VU Amsterdam (Frank van Harmelen, Stefan Schlobach, Ilaria Tiddi):** The cooperation targets knowledge representation and reasoning for robot agents and focusses on knowledge graphs, automatic creation of robot knowledge bases from linked open data, and the learning and reasoning with hybrid symbolic/subsymbolic knowledge bases. The cooperation started with a joint master thesis, resulting in a joint paper (Kümpel et al., 2020), but is slowed down through the Covid situation.
- **TU Vienna (Markus Vincze):** The cooperation aims at cognitive robotics with a particular focus on cognitive robot vision. We are cooperating together in joint projects including Knowledge4Retail and the European H2020 project TRACEBOT, which will start in Spring 2021. In addition, Markus Vincze submitted a proposal to FWF (Fonds für wissenschaftliche Forschung, Österreich) for a cognitive robot vision project to be associated with the CRC EASE. Finally, we have established a joint team to participate in the RoboCup@Home competition, which was cancelled due to the Covid situation.
- **Italian Institute of Technology (IIT, Giulio Sandini):** The cooperation targets cognitive architectures, which will be established as a common IIT focal research topic. We plan joint seminars and cooperation in joint research projects. Specifically, the IIT has launched a transdisciplinary initiative — iCog: the iCub cognitive architecture — to establish a working group at IIT to discuss both abstract models of cognition, in natural and artificial agents, and the software implementations of such models to be used as a reference architectures in the field of Robotics and Artificial Intelligence and as shared tools for the iCub's international community. The goal is to expand the IIT-centered iCub community by establishing a network of international labs, with the IAI being the first international lab to join. Our initial contribution is to present a series of three talks on cognitive architectures, including the CRAM cognitive architecture, in order to establish a common base for further discussions.
- **University of Costa Rica (Federico Ruiz):** The joint research interests include bimanual cognitive robot manipulation of objects. The modi of cooperation include longterm hosting of doctoral students and the preparation of a joint German-Costa Rican research project. Federico Ruiz was a visiting professor in the EASE CRC for eight months in 2020 including a master student (Israel Chavez) for two months.
- **University of Tokyo (Masayuki Inaba, Kei Okada):** The cooperation topic include different aspects of intelligent autonomous robot manipulation. The modes of cooperation include mutual research visits of doctoral and postdoctoral researchers and open-source development. In the first EASE funding phase Dr. Asil Bozcuoglu and Dr. Daniel Leidner were visiting the JSK Laboratory at the University of Tokyo for extended research stays. In return, Yuki Furuta visited Bremen in Fall 2016. A main result of the cooperation was a joint ICRA paper about abstract robot to robot transfer learning using OPENEASE, which was a finalist for the best paper on cognitive robotics at this conference.
- **Universitat Politecnica de Catalunya (UPC, Jan Rosell):** The main subject of cooperation is the

development and standardization of ontologies for autonomous robotics. Alberto Olivares-Alarcos and Mohammed Diab have been EASE visiting researchers during September and February 2019, respectively. Both are members of UPC and the IEEE working group on Ontologies for Robotics and Automation (ORA). The coorparation has resulted two journal (Olivares-Alarcos et al., 2019; Diab et al., 2020a) and two conference publications (Beßler et al., 2018b; Diab et al., 2019).

- **Seoul National University (Tak Zhang):** Cooperation in the area of robot imitation learning with deep networks in the joint German-Korean project ILIAS started in April 2019 and Global Frontier Research Program: Human-Level Machine Learning including exchange of postdoctoral and doctoral students. To this end, Dr. Asil Bozcuoglu visited Biointelligence Lab of Seoul National University in Spring 2018. In return, many Korean early-career researchers including Dr. Minsu Lee, Chung-Yeon Lee, Joonho Kim, and Seungjae Jung have visited Bremen in Winter 2017/18, Summer 2018, and Summer 2019.
- **National University of Singapore (Haizhou Li):** Haizhou Li has a research excellence chair at the University of Bremen working with Tanja Schultz and being an associated EASE researcher working on robot listening and data intensive machine learning.
- **Free University of Bozen-Bolzano (Oliver Kutz):** Cooperation in the area of knowledge representation & reasoning focussing on the formalization of image schemas and ontology engineering. The cooperation is supported by the DAAD-MIUR funded exchange programme project SCORE (From Image Schemas to Cognitive Robotics - A formal framework and computational models for embodied simulations, 2018-20) under the lead of Oliver Kutz and John Bateman.
- **National Institute of Informatics (Inamura):** Joint research in the area of imitation learning from VR environments, hybrid symbolic-subsymbolic representation of behavior, and symbol grounding.

EASE has hosted a number of senior and early career researchers for extended research stays. These are listed in Section 1.3.1.2.2 (“EASEopenLab”).

Finally, EASE has sent doctoral students and postdoctoral researchers to international cooperation partners:

- Gayane Kazhoyan: Massachusetts Institute of Technology, Boston, Massachusetts, USA, host: Prof. Leslie Kaelbling. July-October 2019.
- Dr. Daniel Leidner: JSK Lab, University of Tokyo, Tokyo, Japan. March-April 2020.
- Dr. Daniel Nyga: Massachusetts Institute of Technology, Boston, Massachusetts, USA, host: Prof. Nick Roy, January-April 2018
- Johannes Pfau: University of Malta, Republic of Malta. February-March 2020.
- Prof. Rainer Malaka: Chiang Mai University, Thailand College of Arts Media and Technology. December 2019.
- Dr. Asil Kaan Bozcuoğlu: JSK Lab, The University of Tokyo, Tokyo, Japan. July-September 2017.
- Dr. Asil Kaan Bozcuoğlu: Biointelligence Lab, Seoul National University, Seoul, Korea. March-May 2018.

1.2.5 Research goals for the 2nd phase of EASE

The ultimate goal for the 2nd phase of EASE is to investigate cognitive architectures for robot agents accomplishing everyday manipulation tasks by designing, realizing, and studying CRAM2.0, the next generation of the CRAM architecture. We intend CRAM2.0 to be an extension of CRAM as it has been described in the Sections 1.2.3.1 and 1.2.3.2 that adds additional new functionality. It will target the improvement of existing cognitive and manipulation capabilities, add new ones, and refine the mechanisms for orchestrating these capabilities. CRAM2.0 will also consitute an interface layer for cognition-enabled robot manipulation. The next generative models for accomplishing the EASE household challenge to be realized in phase 2 will be rigorously and firmly built based on the CRAM2.0 interface layer. CRAM2.0 will also be designed to provide and support extensibility through standardized programming interfaces and data structures so that it can leverage, in an open-ended and opportunistic manner, the expected rapid progress in research and technology fields including machine learning, physics simulation, virtual reality, optimization, robot control, and computer vision. These technologies are, from the EASE perspective, located under the interface layer for cognition-enabled robot manipulation and constitute powerful computational resources.

The improved capabilities of the cognitive architecture will be necessary because EASE raises the expectations for the robot agents in terms of their cognitive and manipulation capabilities as well as their performance in accomplishing the EASE household challenge. These expectations include:

- For habitual tasks, the robot is able to directly access the knowledge needed for action contextualization and make the necessary reasoned decisions without having to pause its activity: the robot moves fluently and the execution speed is limited mainly by the physical constaints and low-level sensorimotor dexterity but not by the computational resources needed to perceive and reason, or by the need to engage in deliberative online exploration of alternative action strategies.
- We will challenge the generality of the manipulation capablities of robot agents — in simulation environments — by generating new environments, objects, and robot bodies, including ones they might, in the beginning, not have the sufficient knowledge for competent action, such as setting the table in an unknown kitchen.
- The robot has to competently handle much more cluttered scenes, which will require the robot to perceive objects and scenes even in contexts with insufficient sensory evidence: as Aaron Bobick at Georgia Tech noted "Cognitive vision is a lot about being able to assert that something is there, given very little visual evidence, and even perhaps despite evidence to the contrary" Vernon (2008). This requires robots to perceive what they cannot see through imagination, i.e. internal simulation. This will enable them to better forestall possible negative consequences of actions in complex scenes and infer what many not be perceptually evident in the current scene in order to act in an appropriate manner.
- The introspective and cognitive capabilities of CRAM2.0 will be extended to the contextualized execution of manipulation actions: robot agents will be able to detect situations in which it is better to stop the execution and determine how to stop safely, they will be aware of how they have grasped objects and can judge whether the grasp configurations are stable or are likely to deform objects, and they will know where to look for objects if they drop them. So far, these prospective, introspective, cognitive capabilities are limited to the planning of the contextualization of actions.
- At any time a user can open up an OPENEASE web interface with an EASE robot agent that performs everyday manipulation tasks in a real or simulated environment and ask it an open set of questions formulated in the KNOWROB query language that include information about the robot, the objects, the environment, the tasks, its capabilities, its plans, its beliefs, and its intentions. The domain over which the questions can be asked will even include the cognitive architecture itself and how it works in order to generate the robot activity.

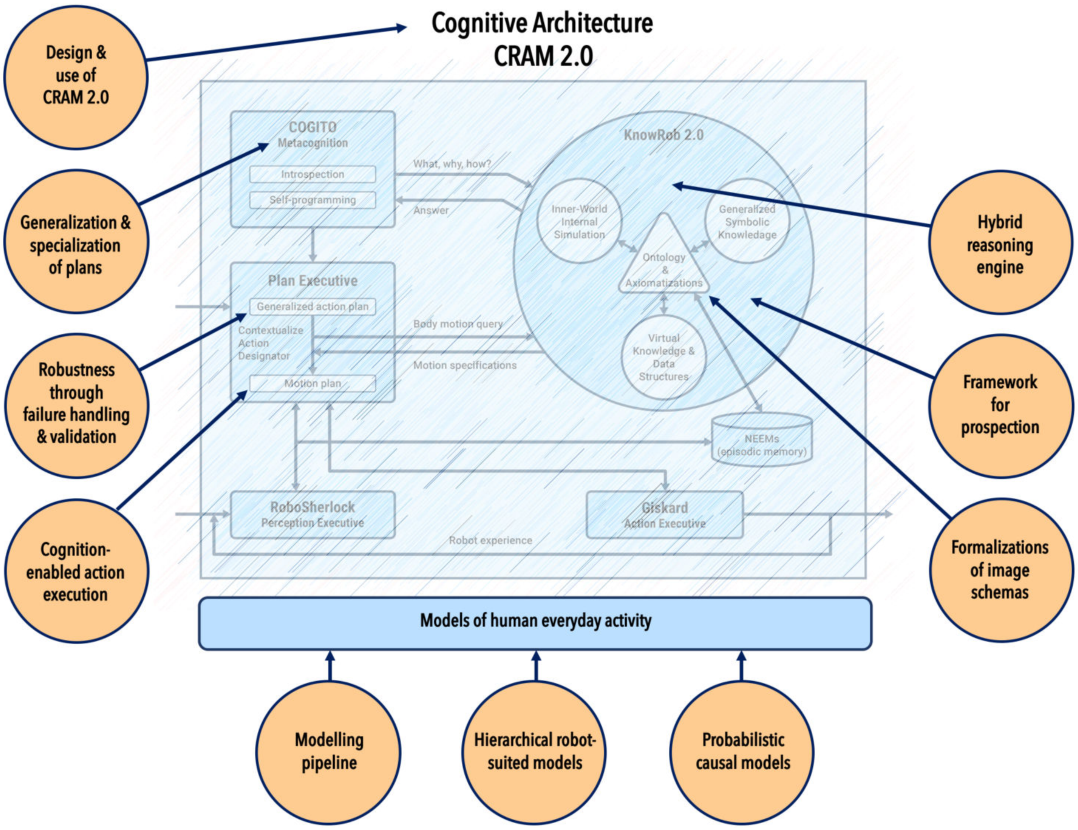


Figure 1.52: Schematic overview of the principal goals of phase 2, all of which combine to create a much-extended version 2.0 of the CRAM cognitive architecture based on the situation model framework. CRAM 2.0 will feature significantly enhanced metacognitive abilities, as well as flexible, context-sensitive cognitive behaviours — both fast habitual behaviors and slower but more adaptive deliberative behaviors — and prospective cognitive motion control during action execution.

- The next generation of cognitive architecture is to improve the support of interdisciplinary research in everyday activity science and engineering by providing a situation model (Schneider et al., 2020) perspective of generative model based on CRAM2.0. This perspective is to help bridge the conceptual gap between the different research areas of EASE that are rooted in different research fields with different conceptualization and terminologies.

To meet these challenges, EASE structures its research into the top-level goals of investigating

- the KR&R capabilities that enable robot agents to accomplish everyday manipulation activities by designing and realizing the next generation of KnowRob in EASE research area P;
- cognition-enabled perception, plan, and action executives and their interplay that leverage the KR&R capabilities in order to generate competent and fluent manipulation behavior in EASE research area R;
- models of human everyday activity, which are formulated from a situation model framework perspective in EASE research area H to facilitate an easier transfer into the generative model.

This top-level structure is refined into four objectives for each EASE research area, which serve as research foci and are aligned with targeted research achievements. This refined structure is shown in Figure 1.52 and will be detailed in the exposition of each research area. The individual subprojects then target a single or a combination of the research objectives of their research area. The goal structure of the individual subprojects are then intended to sketch the research topics (typically the topics of dissertation theses) of the individual EASE researchers.

To implement the dependency structure between EASE research goals and projects more effectively, we shifted the subprojects H02-E in research area H, which investigated the learning of contextualization knowledge from humans, and R03, which investigated mental simulation in research area R, into research area P. Now all research projects contributing to the KR&R framework are part of research area P. We did not enforce that organization when starting EASE because we considered the risk of methodological gaps in embodied reasoning and more formal knowledge representation research. As a result of our close collaboration, we were able to close the methodological gaps early in the first phase of EASE due to joint and common work on the EASE ontology and the standardization of NEEMs across all research areas driven by research teams with the members from different areas. These very positive developments give EASE the opportunity to structure the research areas according to the structure of the CRAM cognitive architecture with research area P now covering knowledge representation & reasoning, and research area R covering the plan executive and metacognition.

In addition, it is a constant and continual goal of EASE to have a complete generative model based on the CRAM2.0 cognitive architecture in operation at any time and demonstrate the research progress in a milestone demonstration every year. This requirement forces the EASE team to maintain and improve a complete, end-to-end robot control system, which means that all necessary software system components are available and function together on demand. Furthermore, we anticipate that EASE researchers in the research areas R and P will invest considerable time and energy integrating their specific research in the cognitive architecture and the control system because it gives them the opportunity to demonstrate the impact of their research in convincing autonomous robot experiments.

1.2.6 Research planning for the 2nd phase

Research Area R: Generative models for mastering everyday activity and their embodiment

Research area R aims at designing, realizing, and studying the CRAM2.0 cognitive architecture, which is to extend the CRAM architecture described in Section 1.2.3.2 to increase the level of flexibility and context-sensitivity of the cognition-enabled robot behavior, to improve the robustness, failure tolerance, and the verifiability of generalized robot plans, and extend the meta cognitive capabilities of the architecture. To this end, research area R will extend the capabilities of the CRAM plan executive, the perception executive ROBOSHERLOCK, the action executive GISKARD, and the meta cognition component COGITO and improve the interplay of these components in order to achieve the next level of manipulation capability. The KR&R component KNOWROB will be investigated in research area P.

Using the CRAM2.0 cognitive architecture as the basis for implementation, research area R will realize the generative models for the EASE household challenge for the yearly milestone events.

The cognitive architecture CRAM2.0 will also operate as an interface layer for programming robot agents to accomplish human-scale manipulation tasks by providing a programming language with data/knowledge and procedural abstractions and composition mechanisms that allow the programming generative models for accomplishing everyday manipulation tasks. The design of CRAM2.0 leverages the insights that we gained in the first EASE phase — "understanding by building".

The focussed research objectives of research area R in the second phase are the following ones:

Objective 1 – Flexible, Context-sensitive Plans *We aim to provide* CRAM*2.0 with the capabilities needed to achieve the flexibility and context-sensitivity required to cognitively create new plans for novel tasks in unexpected circumstances based on old experiences (using inspirations from the situation model framework).* EASE considers the situation model framework to be a promising conceptual framework for studying the generation of cognitive agent behavior because it

- facilitates the characterization of a variety of cognitive behavior patterns;
- targets fluent and optimized behavior for habitual tasks and contexts as well as prospective and reasoning based action for novel and nontrivial tasks;
- is firmly based on concepts in cognitive science and is therefore expected to facilitate the fruitful interaction between the research areas;
- provides a research context in which metacognitive capabilities including the generalization and specialization of generative models can be investigated.

Creating new plans for actions based on behavioral episodes in long-term memory is a key cognitive capability hypothesized in the situation model framework. This action planning capability goes far beyond the capabilities of state-of-the-art planning approaches investigated in the field of AI planning because the AI approaches require the existence of a predefined set of atomic actions with axiomatized models (typically stated as a PDDL problem Ghallab et al. (1998); Fox and Long (2001)). In contrast, the domain of planning in the situation model framework includes all behavior patterns with their associated effects contained in the agent memory, which implies that the robot agent can potentially invent new actions by associating behavior patterns and their effects with goals. In addition, in the situation model framework the robot agent can also adapt and refine its action models based on experience.

Meeting Objective 1 might require the robot agent to (a) infer which aspects of a novel task or situation go beyond the capabilities of its generalized plans, (b) retrieve (sub-)episodes from collected NEEMs which potentially contain behavior patterns that handled these novel aspects. These behavioral episodes could, perhaps, also come from observing humans and models from research area H, (c) generalize the episodes, for example by reformulating them as a partially instantiated subplans with action designators, which makes them adaptable to new situations and force them to be executed through the behavior pattern in the retrieved episode, (d) stitch the self-created plan pieces together, (e) mentally simulate the self-created plan to identify flaws, and even make small bug fixes based on the flaws predicted in the mental simulation.

We will start by looking at appropriate specializations of this open-ended planning challenge and then extend the scope of it.

This objective will be taken on under the lead of Subproject R01 while recruiting the help of Subprojects R03, R04, and R05.

Objective 2 – Automatic Specialization & Generalization *Enabling robots to automatically specialize and generalize cognition-enabled behavior specifications through self-aware learning and self-programming.* One of the key advantages of the CRAM cognitive architecture with respect to other architectures is that it makes the meaning of data structures and processes accessible and machine understandable in the form of digital twin and virtual knowledge bases, it gathers huge amounts of semantically annotated learning data in the form of collections of NEEMs, and it explicitly represents cognition-enabled behavior specifications as generalized action and motion plans. The availability of these machine-understandable and processable representations puts EASE in an excellent position for investigating self-aware learning and self-programming.

An important application of self-aware learning and self-programming will be the automated specialization and generalization of the generative models and generalized action plans. We consider self-aware learning to be the cognitive capability of a robot agent to automatically create relevant learning problems, provide the training data for the learning problems, install the solution, and monitor the effectiveness and the need for adaptation of the learned code pieces.[66] Self-aware learning will take a generalized action plan and a collection of NEEMs to identify habitual tasks, create learning tasks for automating the behavior for these tasks, solve the resulting learning tasks, integrate the learned code into its repertoire of generalized action plans, and update the generative model with (a) the parameter values of the motions derived from this generalized action plan and (b) their associated effects. After having installed the learned code, the self-aware learning capability has to monitor the performance of the code and maintain and adapt the code over time. Given a generalized action plan and a bug (such as, an irrecoverable execution failure, a high cost undesired side effect, or an action that is accomplished more efficiently by humans, the robot agent should revise the generalized plan to competently deal with the bug, either by avoiding it or predicting it. The robot agent then has to decide whether or not to accept this adaptation.

Achieving Objective 2 requires the concerted research effort of Subprojects R01, R02, R03, R04, R05, and P04 under the lead of the Subprojects R01 and R04. Subproject R01 will contribute the software infrastructure for the Two Systems approach of employing a combination of a fast execution mode for habitual tasks and situations and a cognitive and prospective mode for tasks and situations that require weighing alternatives. Subproject R04 will contribute the capabilities for self-programming through transformational learning and planning. Subprojects R04 and P04 will contribute reinforcement learning methods for self-aware learning where R04 focusses on optimizing sequences of behaviors and R05 on combining symbolic and subsymbolic aspects in reinforcement learning. Subproject R02 contributes to this framework by retraining perception CNNs on images with ground truth, rendered by the digital twin to improve perception in these tasks. Subproject R03 will provide the prospective reasoning capabilities needed for behavior specialization and generalization. Finally, Subproject P04 contributes verification techniques which can help in checking the correctness of the applied plan transformation.

Objective 3 – Self-improving Failure Handling *Investigate modular and self-improving capabilities for avoiding and handling action failures.* When inspecting the code of robot control programs that

[66] Self-aware learning, in the broader context of all artificially intelligent systems is one of the three core technical areas of the 20-year community roadmap for artificial intelligence research in the US, starting in 2019 (Gil and Selman, 2019).

autonomously and robustly accomplish long-term manipulation tasks in realistic environments we see that by far most of the code deals with failure detection, analysis, and recovery. Failure handling code is often complicated and typically has to tightly interact with the primary course of action because often failures can occur at any time, in any guise, and this can cause many disturbances of the planned course of action. Even worse, when operating over a lifetime, robot agents will unavoidably learn about new ways in which their actions fail and have to adapt their plans to avoid these failures in the future. Therefore, a powerful and modular framework for handling execution failures competently can be expected to greatly improve the overall performance of robot agents. The modularity of failure handling capabilities is important because plans have to be highly modular and transparent to make self-programming and introspection feasible. In this context, validating the absence of failures under asserted conditions is a promising tool for achieving higher modularity.

Considering the huge impact that competent failure handling has in robot agency, the topic has received surprisingly little attention. It is often being pushed into individual software components that are hidden from cognitive reasoning mechanisms (e.g., in the widely used three-layered control architectures (Bonasso et al., 1997)) or being handled ad-hoc by jumping into a failure state in a state-automata based execution system. Realizing failure handling capabilities in these ways hinders and limits the cognitive capabilities of robot agents because the software code resulting from these failure handling approaches is hard to understand and maintain even for expert human programmers.

In this objective, we will propose and investigate failure handling capabilities that are more comprehensive, capable, and modular, and plan designs that facilitate competent failure handling without making the generalized action plans too complex and interwoven for introspective reasoning and self programming capabilities. We further investigate the formation of failure taxonomies that support the diagnosis of failure causes to initiate better informed failure recovery methods and how these taxonomies can be automatically refined and improved. In addition, we will investigate the learning of prediction models for failures that will enable robots to forestall their occurrence. Finally, plan validation will provide us with the means to verify that under specific asserted conditions execution failures cannot occur providing us with additional means for failure diagnosis and for modularizing failure handling.

This objective will be tackled by the Subprojects R06-N, R01, R04, and P04 with Subproject R06-N taking the lead. Subproject R06-N will primarily investigate taxonomies of failures and prediction models for them. The role of Subprojects R04 and R01 will be the incorporation of failure handling into generalized action plans and habitual behaviors. Subproject R02 realizes perception methods that provide uncertainty estimates for perceived object poses and parameters, and thereby enables the robot agent actively reduce uncertainty when needed. Subproject P04 will investigate methods to assess and improve plan quality and reliability. This includes techniques to detect and avoid execution failures.

Objective 4 – Cognitive Action Execution *Achieve a high level of action, manipulation, and motion awareness by endowing action execution with cognitive capabilities.* In its current state, CRAM has powerful cognitive capabilities to contextualize underdetermined action descriptions and plan their execution through body motions. When it comes to the execution of the behavior specifications, however, CRAM is very short- and narrow-sighted. Its primary windows into execution are the failures it monitors and the sensory feedback it expects to advance into the next motion phase. This limited window into execution manifests itself in flaws in manipulation capability. For example, when picking up an object, the robot agent does not estimate the in-hand pose of the object after grasping. Therefore, it has to place objects very carefully to account for the uncertainty of the object pose while transporting it. Also, the robot does not know where an object has gone if it was dropped while transporting it. Similarly, the robot does not know the pose of a drawer while opening or closing it. Furthermore, the execution of actions is often less flexible that it could be. For example, when filling a glass it matters that the glass is full but the pouring motion does not necessarily matter. Allowing motions to be specified in terms of the effects they are to cause and then backprojecting the effect parameters into the controllable parameters at execution time would further increase the manipulation capabilities of robot agents.

In order to advance the cognitive capacity of robot agents for action execution we intend to leverage the representation and prospection capabilities of KNOWROB2.0. These capabilities allow the robot agent to mentally emulate the execution of actions as a dynamic KNOWROB knowledge base, visually render the emulated state such that the perception system can compare it with the captured image, detect deviations and semantically meaningful execution events. We believe that these extended perception action loops that continually emulate the motions, render the the predicted state, compare the rendered state with observations of the actual state, and explain the differences between the expected and the perceived state will result in a cognitively-aware and therefore more competent action execution.

In order to accomplish these objectives, EASE proposes to establish six subprojects:

- **Subproject R1** The cognitive robot architecture CRAM2.0 for accomplishing human-scale everyday manipulation tasks. (PIs: Cheng, Beetz, Vernon). The subproject targets Objective 1, the 2nd generation of the CRAM cognitive architecture, by analyzing and rationally reconstructing the control programs that resulted from phase 1 and using the resulting insights in order to propose CRAM2.0. CRAM2.0 will be designed more rigorously and in a more modular manner to, accomplish fluent execution of actions. Concepts from the situation model framework will be included in the design, applying the Two Systems approach to thinking fast and slow, and basing the semantics of action descriptions solidly on probabilistic causal models of actions.

- **Subproject R2** Robot perception (PI: Frese). The subproject will continue to investigate computational vision experts for housework perception tasks based on the successful two-stage design of Deep Learning plus geometric Bayesian fusion proposed in the first phase. The new generation of vision expert will go beyond pre-known rigid objects covering also articulated objects (e.g. drawers), objects with parametrized geometry (e.g. plates of different dimensions) and even unknown generic objects (e.g. cups of unknown shape). In addition, the subproject will also investigate how to leverage short term memory to extend and augment the information of a detected object through refinement and examination.

- **Subproject R4** Cognitive capabilities for generalized plan schemata (PIs: Albu-Schäffer, Beetz). Subproject R04(p2) will design, realize, and investigate a **cognition-enabled, plan-based, and context-aware execution component** that can continuously monitor and spontaneously and smoothly intervene in the action execution to maximize success and efficiency. This involves the application of the powerful cognitive capabilities of the CRAM plan executive to the percept-guided execution of actions in order to make it action and motion aware. Novel manipulation capabilities include the estimation of the pose of objects or tools in the hand after they are picked up and the recognition of context-relevant states and events during execution. Achieving these capabilities requires novel methods for transforming symbolic action descriptions into motion optimization problems, metacognitive capabilities for specializing and generalizing motion strategies, and optimization methods for compund motion plans.

- **Subproject R5** NEEM-enabled deep reinforcement learning for accomplishing everyday manipulation tasks. (PIs: Ritter, Beetz). Subproject R05 will investigate the synergistic combination of cognition-enabled robot manipulation and deep reinforcement learning for the acquisition of hand manipulation skills beyond pick and place. To this end, the subproject will (a) extend the CRAM plan language novel with control structures that facilitate reinforcement learning and (b) extend the richness and expressiveness of the representations contained in the deep networks by combining vision and touch and covering long time windows (several minutes) in addition to short ones (very few seconds). The aim is the creation of bottom-up representations that can be aligned with the hybrid automata models.

- **Subproject R6** Fault-Tolerant Planning and Recovery Mechanisms for Everyday Manipulation. Everyday manipulation actions — even when executed by humans — often fail. Competently handling failures in robot plan and action executives is open-ended because failures can happen at any time and under all circumastances and the question of how to continue a partly-executed activity after a failure has occurred is very context-sensitive. Subproject R6 aims at utilizing AI-based planning, probabilistic physics reasoning, and modern machine learning techniques to identify failure situations and recover appropriately. To this end, large bodies of experience will be collected to generalize the observed behavior, probabilistic effects have to be considered, and the data has to be semantically annotated to transfer raw data into interpretable NEEMs.

- **Subproject P4** Formal definition of the CRAM plan language and the validation of plan-guided robot. (PIs: Drechsler, Herdt). Based on the CPL verification framework, Subproject R6 will widen and deepen the investigation of validating complex plan-guided robot behavior behavior towards generalized manipulation tasks that are executed over a long period of time and operate in non-deterministic environments. To tackle these challenges R5 intends to leverage the recently developed Virtual Prototype (VP) based verification techniques for embedded systems that have been very effective in improving verification coverage and finding critical bugs in real-world systems and transfer the underlying techniques and ideas into the robotic domain. We aim to accomplish this transfer by (a) the development of a VP as the simulation backbone with a reasoning engine for complex cognition-enabled plans with extensive environment interactions; (b) the design of a hybrid verification engine that bridges between simulation-based and formal methods in a unified framework; (c) the extension of the environment modeling and integration with the VP-based engine by leveraging fine grained models at different levels of abstraction.

Research Area P: Principles of information processing for everyday activity

The pervasive use of machine-understandable knowledge representation and automated reasoning capabilities, which are realized through the KnowRob knowledge representation & reasoning (KR&R) framework, is a defining characteristic of the CRAM cognitive architecture. Any component of the CRAM architecture can ask queries formulated in the KnowRob query language at any time and automatically process the answers thanks to the SOMA ontology defining the standardized system-wide semantics of the relevant concepts underlying the answers to these queries. The data structures used by the CRAM components that have system-wide relevance are cast as virtual KnowRob knowledge bases and important computational procedures and processes can be queried as if they were KnowRob knowledge bases. The photorealistic visualizations of the robot's beliefs are visual renderings of symbolic belief states represented in KnowRob. Every activity that the robot agent performs, observes, and simulates is automatically recorded as a NEEM, as a conjunction of KnowRob facts. Because the symbolic representations of NEEMs are time-synchronized with the subsymbolic data streams of the NEEMs, the robot agent can automatically generate the training data for learning problems through KnowRob queries. The contextualization of underdetermined task requests is implemented through KnowRob reasoning processes.

The role of research area P is to provide, improve, and extend the KnowRob knowledge representation and reasoning capability. To this end, research area P will aim at extending the SOMA ontology by introducing formalizations of image schemata for more adequate concept definitions, while also providing stronger support for modelling concepts across multiple levels of abstraction, a key requirement in the first phase of EASE. A second focus will be the provision of a comprehensive KR&R framework for prospection. This will provide different reasoning capabilities, including mental simulation, symbolic plan projection, prediction, and imagistic reasoning. This KR&R framework will maintain the consistency between these capabilities and adapt the realism of representations to real-world experience. Finally, research area P will investigate meta-reasoning capabilities that enable KnowRob to select and orchestrate the required reasoning techniques for answering KnowRob queries.

More specifically, the research objectives of research area P for the second phase will be:

Objective 1 – Formalized Image Schemas *The design, implementation, and evaluation of a library of fully formalised image schemas to deliver an abstraction layer for contextualizing underdetermined action descriptions more effectively.* Image schemas are hypothesized as recurring structures within our cognitive processes, which establish patterns of understanding and reasoning. Image schemas activate the use of mental images of objects, containers, paths, and the like, to support the understanding and contextualization of vague expressions. Therefore, they constitute promising concepts to bridge from bodily interactions to higher-level cognition.

For this reason, EASE aims at extending concept definitions in the EASE ontology with formalizations of the respective image schemas underlying the concepts. We believe that the formalization of image schemas will facilitate the acquisition of commonsense knowledge as well as the contextualization of underdetermined action descriptions. The provision of image schemas therefore represents a major manifold (PEAM) for understanding activities and the natural language instructions to perform them in a manner that is consistent with how humans understand and perform directed, situated, and appropriate actions. This objective contributes to the essential principles of the broader EASE framework in the second phase which aims at establishing a fully generic representation and reasoning framework.

This objective will be primarily tackled by subproject R01 in strong cooperation with the EASE ontology working group.

Objective 2 – Ontologies with Multiple Levels of Abstraction *The design, formalization, and investigation of an infrastructure for ontologies for robot and human everyday activity that supports modeling across multiple levels of abstraction in a principled way, focussing particularly on aspects of* *knowledge representation and reasoning for robotics.* This objective is expected to provide thorough, theoretical foundations for the ontology-based modelling of knowledge and its representation in the EASE ontology. This effort will be lead by subproject P02 and be applied in the EASE ontology modelling of the respective working group.

Objective 3 – Prospection *The design, realization, and investigation of a representation and reasoning framework for prospection, which provides the repertoire of prospective reasoning methods needed for competently accomplishing everyday manipulation tasks.* Given KNOWROB queries about the future, the framework will select the appropriate mental representations of possible futures through anticipation, simulation, projection, and imagistic reasoning. The framework will also learn and adapt models and representations for prospection that are consistent with each other and realistic with respect to real-world execution. As prospection is to be tightly integrated into the perception-action loops embedded in the EASE approach to robot agency, an investifation of prospection requires close cooperation with the subprojects in research area R.

Objective 3 is tackled under the lead of Subproject R03 with substantial contributions of Subproject P01 to the representation of the results of prospection and Subproject P05-N to the meta-reasoning problems of selecting the appropriate prospection methods.

Objective 4 – Question Answering Capability *The investigation of a question answering capability for* KNOWROB *queries that employs the hybrid representation and reasoning techniques available in* KNOWROB *through meta-reasoning mechanisms.* Currently, the sequence of subproblems, as well as the reasoning methods with which KNOWROB queries are answered, are specified by programmers through Prolog rules. Objective 4 aims at (a) automating the process of reasoning about how to solve which reasoning problem and (b) the learning-based optimization of reasoning.

Objective 4 is the target of Subproject P05-N. To this end, Subproject P05-N cooperates closely with Subproject R03, which investigates a framework for prospection, Subproject R03 investigating representations at different levels of abstraction, and Subprojects R01 and R04 that conduct research on fast execution time reasoning methods.

In order to accomplish these objectives EASE proposes to establish five subprojects:

- **Subproject P1** Embodied semantics for everyday activities. (PIs: Bateman, Malaka). Subproject P1 investigates the research hypothesis that hybrid representations involving both formal logical theories and simulations could constitute a powerful representation infrastructure for everyday manipulation tasks, and in particular that the cognitive linguistic notion of image schemas could serve as a level of abstraction capturing similarities across very different physical situations. In the second phase, explicit formalization of image schemas as hybrid theories will be pursued, involving logical descriptions at an abstract level directly building on, and feeding into, the EASE ontology framework and construction grammars as well as subsymbolic descriptions in terms of generative models for particular situations. The envisioned result will be the design, implementation and evaluation of a library of fully formalised image schemas to deliver an abstraction layer generalising across all domains of activity relevant for EASE so as to support and evaluate the transfer of methods across reasoning tasks.

- **Subproject P2** Ontologies with Abstraction (PIs: Lutz, Bateman). The development and use of ontologies that cover a wide variety of aspects of robot and human agents and everyday manipulation tasks was one of the big successes of the first phase of EASE. This pervasive use of ontological knowledge and the complexity of axiomatizations validated the usefulness of the ontology approximation techniques investigated by P2 for cognition-enabled robotics. In Phase II of EASE, the P02 project focusses on aspects of ontological representation and reasoning that are of particular importance for robot agency, namely the support of multiple levels of abstraction. We will design ontology languages based on description logics (DLs) that support explicit reference to levels of abstractions and also provide a suitable formalism for representing data. P2 will also investigate different modes of reasoning (including subsumption and query evaluation), design reasoning algorithms and analyze their complexity, and also study the expressive power of the emerging languages.

- **Subproject R3** Anticipate: a KR&R framework for robot anticipation (PIs: Zachmann, Beetz). Prospection — the ability to represent what might happen in the future — is one of the most essential cognitive capabilities of robotic agents. It enables robot agents to prepare for future actions, learn manipulation skills, maintain a promising course of action, and detect and forestall challenges and threads for task accomplishment. The EASE subproject R03 is to design, develop, and investigate a knowledge representation and reasoning framework for prospection that includes different forms and modes of prospection. It will investigate how representations needed for different forms and modes can be automatically generated from NEEMs (narrative enabled episodic memories, i.e. the multi-modal experiences of the robot as it carries out its everyday activities). It will also research the task-specific choice of forms and modes and the roles they have in improving the performance of the robot control system. The proposed prospection framework will be realized based on the digital twin knowledge representation and reasoning (dtKR&R) capabilities of KNOWROB2.0.

- **Subproject P5** (new project) Hybrid representation and reasoning framework for robotic agents (PIs: Bateman/Beßler,Malaka). KNOWROB2.0 (Section 1.2.3.4 and Figure 1.22) is a KR&R framework that consists of a hybrid representation and reasoning core and a logic interface layer that casts the hybrid knowledge base as if it was a logic-based knowledge base. So far, the high-level reasoning tasks formulated in logic are translated into reasoning tasks that can be accomplished through the hybrid reasoning core by KNOWROB rules that are specified by programmers. Subproject P5 will investigate how hybrid representation and reasoning methods can be automatically organized, orchestrated, and optimized in order to solve complex, declaratively formulated reasoning tasks. We will do this by designing, realizing, and studying HYRES (Hybrid Reasoning System) that is to combine and orchestrate a suite of reasoning mechanisms operating on different representations at different levels of abstraction and granularity.

Research Area H: Descriptive models of human everyday activity

The overarching goal of the three subprojects in Area H is to integrate models of human everyday activities into a joint cognitive architecture following state-of-the-art approaches (Vernon, 2020; Schneider et al., 2020; Kotseruba and Tsotsos 2020). In particular, we propose to further advance the top-down and bottom-up components utilizing the data, knowledge, and experience built up in the first phase of EASE, and to tightly integrate these models into the cognitive architecture. To address the purposefulness of human activities, we plan to integrate a decision making component that informs the cognitive system (Vernon, 2014) about WHEN humans make decisions during everyday activities, WHICH decisions they make, and HOW they learn to make decisions.

Starting from the established H-pipeline (Figure 1.46) and the concept of LabLinking, area H has both a structure with a common vision and the integrating components necessary to advance our understanding how humans accomplish their everyday activities. In particular, research area H has the following objectives for the second phase.

Objective 1 – H-Pipeline To advance and extend the previously-developed multi-stage H-pipeline which ultimately establishes the envisioned push-pull cycle relating human behavioral data via models to robot activities. A particular focus will be on the development of causal models (H1), discriminative and generative models (H3), as well as models of decision making (H4), all to be tightly integrated into the H-pipeline.

Objective 2 – Activity Hierarchy To identify a human activity hierarchy using both deep network and constructural grammar models, allowing for assessment of the relative advantages of each approach. The models will target representations at a level of granularity that is compatible with the implementations requirement of CRAM2.0, while also faithfully capturing the nuances of human activity, they will be i.e. distinctive enough to achieve high-quality temporal segmentation but also well suited for generalization.

Objective 3 – Model Granularity To develop generalized models that are well-suited to robot implementation and that allow for generation on all levels of human everyday activities, ranging from low-level sensorimotor to high-level complex everyday activities. In particular, they will be high-level physical and cognitive behavior models that provide insights from human decision making that can be assimilated into the EASE generative model of robot agency and cognitive architectures for robots.

Objective 4 – Uncertainty To understand how humans flexibly adapt everyday activities to new and uncertain environments by studying human behavior, ranging from sensorimotor and complex activity to high-level decision making, and exploring causal representations of invariances between environments.

In order to accomplish these objectives, EASE proposes to establish three subprojects in area H, as follows.

- **Subproject H.1:** Sensory-motor and Causal Human Activity Models for Cognitive Architectures (PIs: Schill, Didelez, Zetzsche). Subproject H1 focusses on the causal sensory-motor models that are encapsulated in long-term memory. This causal modelling, by its predictive power, its inherent generalization capabilities, and its potential for explanation, impacts all three levels of operation. First, it facilitates flexibility in metacognitive expansion of existing capabilities, effectively generating new knowledge and new action capabilities, and allowing the robot to operate in unexpected or novel situations, adapting both the plan language and the generative model. In other words, the concept of metacognition allows us to expand capabilities by generating new action policies rather than by adapting existing ones. Second, it provides for flexibility through more effective sampling of the joint distribution in the generative model. At the same level of operation, causal modelling can also benefit attentional processes, both internal and external. Third, it will provide the flexibility necessary during action execution for adaptive movement generation, as provided by the causal sensory-motor models obtained using the enforced adaptivity paradigm and a causal structure analysis of the data.

This project will contribute to causal models to objective 1; investigate low-level units for objective 2; develop robot-suited low-level sensory-motor activities for objective 3; implement adaptation schemes of sensory-motor activities and explore causal representations for objective 4.

- **Subproject H.3:** Discriminative and Generative Human Activity Models for Cognitive Architectures (PIs: Schultz, Schill). Subproject H3 focusses on the development of hybrid discriminative and generative models of human activity (exploiting context-free grammars and probabilistic action units and deep multi-modal networks) to identify new ways of describing the temporally-extended hierarchical organization of motion primitives that comprise complex actions. The discriminative models will contribute to the metacognition level that seeks to expand action capabilities in the cognitive robot and also to the second contextualization level that seeks to leverage system 2 compositionality of behavioral episodes. To complement this research, and with the same goal, subproject H3 also focusses on generative models, again exploiting both deep learning and context-free grammars, while also incorporating findings on causal modelling in subproject H1. By covering both discriminative and generative modelling, subproject H3 leverages the respective strengths of both, (a) directly learning the posterior distribution that characterizes the space of action sequences comprising everyday activities and (b) inferring the posterior distribution by learning the joint distribution over motion values and complex actions (Ng and Jordan, 2001). Furthermore, components like the grammars will be shared to benefit both discrimination and generation.

  This subproject will contribute to discriminative and generative models to objective 1; create the hierarchical grammar as core contributor to objective 2; developing generation for high-level complex activities for objective 3; implementing adaptation schemes of high-level activities and exploring causal representations for objective 4.

- **Subproject H.4:** Decision Making for Cognitive Architectures (PIs: Herrmann, von Helversen, Schultz) Subproject H4 addresses the purposeful and goal-directed nature of human activities. It plans to model human learning and decision-making processes to inform a cognitive robot about what decisions are necessary and when they are necessary to master complex everyday activities, drawing on human abilities to generate flexible context-sensitive behaviour. The ability to flexibly adapt is important when generalizing behavior to new decision situations and in particular when human and robots are challenged with ambiguous situations, uncertainty about action plans, and the processing of interfering information. The objective is to find the optimal trade-off between exploiting object- and situation-specific knowledge and abstract knowledge by understanding how humans acquire knowledge that allows generalizing beyond the distribution of the data which characterizes the situation in which they committed errors and learned, even with very sparse experience.

  This subproject will contribute to the pipeline of information processing by analyzing the temporal structure of decisions both with respect to the identification of NEEM-related decisions (objective 1) and the analysis of the neural representation/recognition for temporal segmentation of decision behavior (objective 2). Further, H4 will investigate how humans learn to master the decision tasks entailed in action plans for everyday activities (such as table setting) with the goal to identify and to model the learning mechanisms and conditions that enable them to flexible adapt to new environments (objectives 1 and 4). This subproject then aims at comparing human learning and behavioral adaptation with robotic learning in the same tasks to highlight entry points for improving the robot cognitive architecture (objective 3).

1.3 Research profile of the University of Bremen

The University of Bremen is a medium-sized research university. Together with the local research institutes and cooperation partners, it constitutes the leading research hub in northwest Germany. The University is a top performer in many areas of national and international excellence. Its international research profile is shaped by six large, interdisciplinary high-profile areas, which are also the key innovation areas of Bremen.

EASE is part of the high-profile area Minds, Media, Machines (MMM) that aims at advancing our understanding of intelligence, cognition in autonomous agents and teams of agents in the context of mediatized worlds.

EASE plays an essential role for the realization of the research strategy and the AI strategy of Bremen. The Bremen research strategy (Wissenschaftsplan 2025) states:

- "AI-based robotics – with the SFB EASE as its nucleus – is to be strategically developed into one of the core topics with which Bremen can position itself in the AI strategy of the federal government and the European Community."
- "The goal is to maintain an internationally leading position for Bremen's research in cognitive and AI-based robotics. To this end, the networking and formation of research alliances with leading national and international research centers will be further strengthened."
- "The SFB EASE is expected to (...) develop into an international beacon in AI-based robotics and establish a leading position in AI-based robotics for autonomous robots that can perform complicated manipulation tasks. For this purpose, suitable structures in research (openEASE) and teaching (graduate school and, e.g., MOOCs) as well as structures in the field of innovation/transfer must be established."

The role of EASE for Bremen's AI strategy is emphasized in the following statements included in the strategy:

- "Bremen's AI strategy focuses on four pillars: (1) strong AI-based robotics research, (2) strengthening the AI economy, (3) opening up to society, and (4) qualification and securing of skilled workers to expand Bremen's position as an AI hub."
- "With AI-based robotics, the University of Bremen is already making a nationally visible research contribution to so-called strong AI through the Collaborative Research Center EASE. The Open Science strategy pursued by EASE – with its components Open Source, Open Data and Open Research – has also secured significant international visibility."
- "In the field of AI-based robotics, Bremen currently holds a unique position because it has an excellent infrastructure of industrial companies as well as many young IT companies with diverse specializations. A wide range of applications for AI-based robotics can be found in Bremen: logistics, aerospace, underwater robotics, navigation and medical technology as well as the know-how of local IT companies are indispensable for the future development of the technology."

In the context of these strategies, Bremen aims at becoming a national center for cognition-enabled robot agents with long-term autonomy and applying this technology to application fields including ocean science, material science, retail and logistics, space exploration, and robot assistants for aging societies. EASE provides an essential component of the basic research foundations for this strategy. Complementary projects for implementing this strategy comprise KNOWLEDGE4RETAIL (funded through the BMBF innovation competition) that develops a digital innovation platform and ecosystem for stationary retail based on the EASE software components KNOWROB, OPENEASE, and ROBOSHERLOCK. In addition, EASE researchers participate in the digital innovation platform KI-SIGS, which aims at establishing an AI space for health appications in northern Germany. The TRANSAIR (Transatlantic AI-based Robotics) project is aimed at developing a transatlantic dialogue on Artificial Intelligence and Robotics in order to learn from the complementary approaches of AI research, commercialisation and public debate in Germany and the USA. In this context, the central goal is to identify bridges for cooperation between the two countries. In the competition for innovative digital teaching material (KI Campus) we are developing an interactive course on cognition-enabled robotics.

In addition, EASE is also at the core of proposal initiatives for a competence center for human robot interaction in assistive robotics technologies and an innovation ecosystem based on cognition-enabled robotics technology.

1.3.1 Role of EASE in the research profile of the University Bremen

1.3.1.1 Long-term research strategy of the University of Bremen

The long-term strategy of the research focus area Minds, Media, Machines of the University of Bremen aims at establishing a convergent engineering research center to tackle societal challenges of our aging societies. The goal is to design, implement and investigate cognition-enabled robot agents that will enable people to live longer independently in their homes and to improve their quality of life. Quality-of-life technology (Kanade, 2012) aims to develop the technologies that bridge the gap between what people are able to do and what they want to do. An example of a quality-of-life technology system is a cognition-enabled robot that can be used by a user with disabilities to get the item that they need whenever they want it, without asking somebody else for help and exerting stress on caregiving family members. The cognition-enabled robot agents that accomplish everyday manipulation tasks and are investigated in the EASE research center and build the core engineering technology in this enterprise. The engineering technology of cognition-enabled robotics is to be combined with a novel convergent science field, which we call "Living Technologies 2.0." It will bring an interdisciplinary research team together to investigate how the independence and quality of life of users can be most effectively improved.

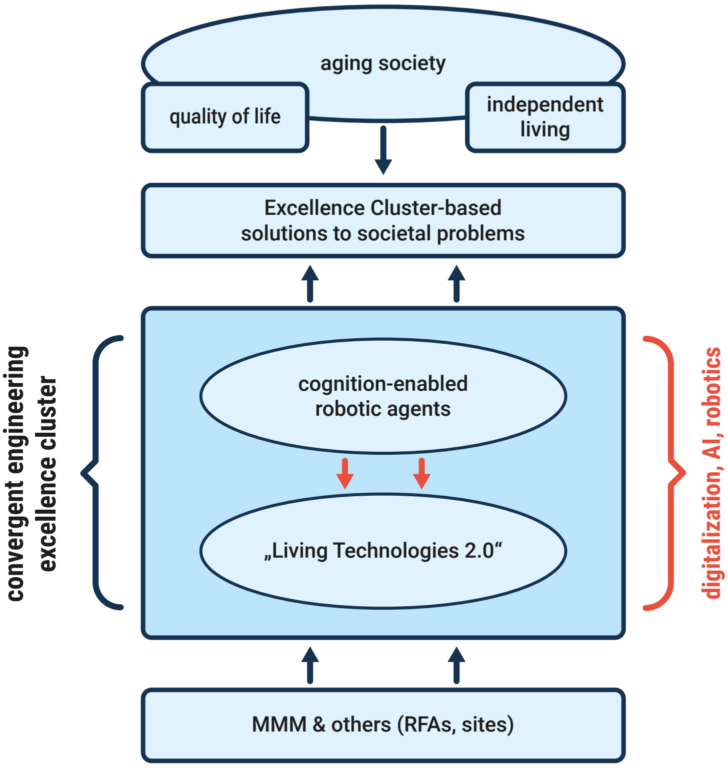


Figure 1.53: Long-term strategy of the research focus area Minds, Media, machines.

1.3.1.2 The EASE-based research and innovation ecosystem

Besides being an essential component of the long-term research strategy of the University of Bremen, EASE also has an immediate impact on the research and innovation landscape of the University of Bremen. Figure 1.54 shows the spectrum of research, innovation, teaching, and training activties that have been built around the EASE collaborative research center.

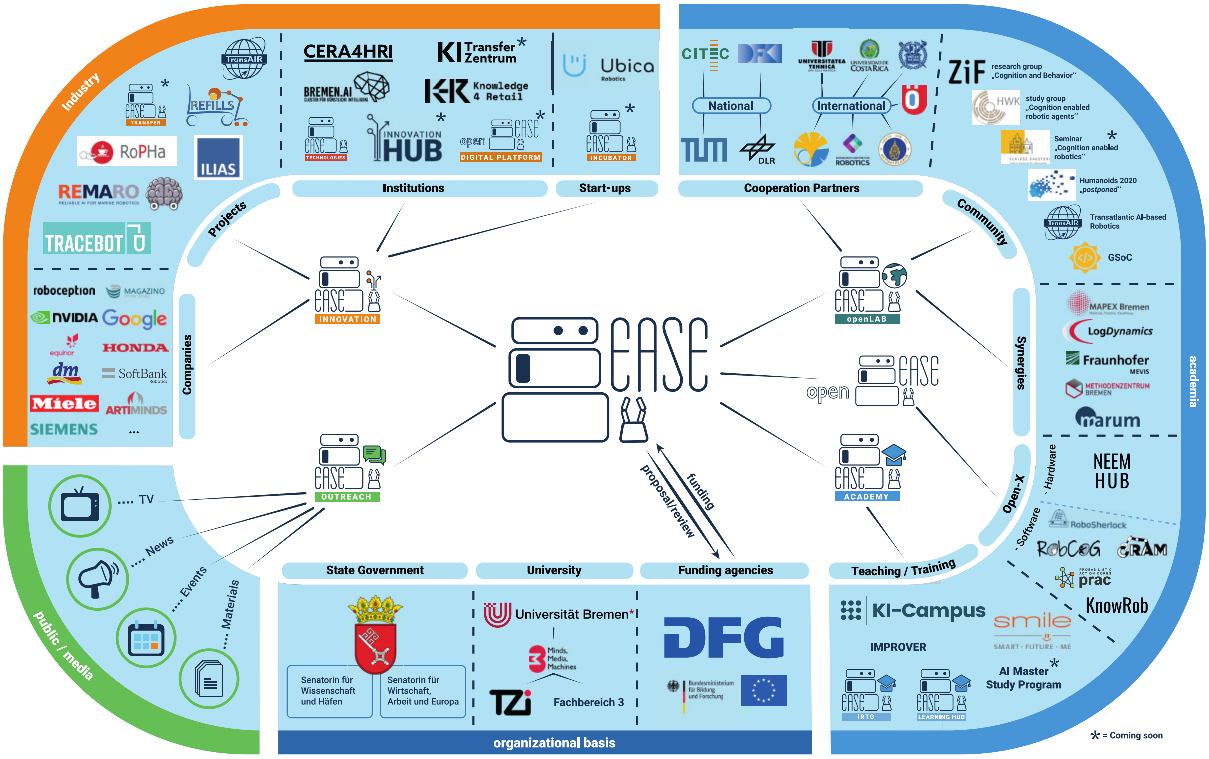


Figure 1.54: The EASE ecosystem.

Main components of the research and innovation ecosystem built around the core research center include

- **EASEINNOVATION**, which aims at pushing insights and results of EASE to higher technology readiness levels and building digital innovation platforms and innovation ecosystems around it;
- **EASEOPENLAB** bundles the efforts of EASE to promote research cooperations with the university's other research focus areas as well as national and international partners. It also includes activities for community-building and researcher exchange mechanisms;
- **OPENEASE** are the EASE activities that promote open research, open data, and open software to implement the mission of EASE.
- **EASEACADEMY** organizes and manages the teaching and training activities of EASE, including the web-based EASE teaching and training hub; and
- **EASEOUTREACH** promotes and organizes the interaction with the general public as well as the activities to attract secondary school students to seek university education in the cognition-enabled robotics field.

1.3.1.2.1 EASEINNOVATION The purpose of EASEINNOVATION is to take the expertise and appropriate outputs of the basic research in EASE and make them available for businesses, governmental and non-governmental organizations, start-ups, and junior entrepreneurs.

To this end, EASEINNOVATION aims at developing industrial-strength implementations of the EASE open-source software components and providing software components running as virtual sandbox laboratories in cloud-based services. The goal is to build a digital innovation and research platform and an ecosystem based on EASE knowledge representation and reasoning and plan-based control methods investigated in EASE. The results will be applicable in various business domains.

EASEINNOVATION also operates living research laboratories. The first one is a living laboratory for retail. The EASE central robotics laboratory will be made available part-time as a second living research laboratory. In addition, we are in the competition phase for establishing a competence center for human-robot interaction.

Organizationally EASEINNOVATION is a department of the Technologiezentrum Informatik (TZI), which primarily obtains funding through joint projects funded by the German government (BMBF, BMWi) or the EC (H2020).

Representative innovation and transfer projects are REFILLS (EC H2020), TRACEBOT (EC H2020), REMARO (Marie Curie Innovative Training Network), ILIAS (BMBF + Corea) and ROPHA (BMBF). REFILLS aims at improving logistics in a supermarket thanks to mobile robotic systems in close and smart collaboration with humans, addressing the main in-store logistics processes for retail shops: in particular, robots will allow a smarter shelf refilling. Information on the supermarket articles is exploited to create powerful knowledge bases, used by the robots to identify shelves, recognize missing or misplaced articles, handle them and navigate the shop. Reasoning allows robots to cope with changing task requirements and contexts, and perception-guided reactive control makes them robust to execution errors and uncertainty. ROPHA (Robuste Perzeption für die interaktive Unterstützung älterer Nutzer bei Handhabungsaufgaben im häuslichen Umfeld) investigates human-robot interaction technologies for meal preparation. TRACEBOT (starting Spring 2021) aims at the realization of laboratory robots that operate in sterile environments and applications with a high demand on flexibility using AI and cognitive methods for creating traceable assembly actions. ILIAS (Imitation Learning from Human Demonstrations in Virtual Reality for Physical Human-Robot-Interaction in Assistance Tasks) strengthens the international cooperation with Byoung-Tak Zhang's Biointelligence laboratory, one of the leading international research groups in South Korea. The ultimate goal of the Biointelligence laboratory center is to discover a large-scale, neurocognitive computational model of the brain that autonomously develops or evolves towards human-level machine intelligence in lifelong interactions with the environment. To get there their research focusses around deep, recurrent, and sparse hypernetwork architectures and learning algorithms that self-organize their structures instantly, incrementally, and continuously in a self-supervised way by perception-action cycle. A high level of synergies exists with their StarLab project that investigates cognitive agents which learn everyday life.

The Marie Skłodowska-Curie Innovative Training Network REMARO targets the development of reliable and trustworthy AI for underwater robotics. It aims at developing the first ever submarine robotics AI methods with quantified reliability, correctness specifications, models, tests, and analysis & verification methods. It generates synergies with EASE in that it will advance two founding principles: (1) The submarine robot autonomy requires a comprehensive hybrid deliberative architecture, a robotic brain. (2) Safety and reliability must be co-designed simultaneously with cognition, not separately as an afterthought.

CERA4HRI is a proposal (competition phase) for establishing a competence center for human-robot interaction at the University of Bremen. The methodological basis of the competence center is planned to be strongly based on the generative model and cognitive architecture investigated in EASE but the application domain will be human-robot interaction rather than autonomous object manipulation. If successful, the CERA4HRI will strongly support the third funding phase of EASE, which is to focus on multi-agent everyday activity.

In addition, EASEINNOVATION manages the EASEINCUBATOR, which is the start-up facilitator for cognition-enabled robotics technology. Currently, the incubator supports one start-up, Ubica, which develops and operates autonomous mobile robots that perform retail store inventory. Figure 1.55 shows Ubica's timeline for transferring from a research project to a company.

Finally, EASEINNOVATION organizes the service robotics interest group of Bremen.AI. Bremen.AI

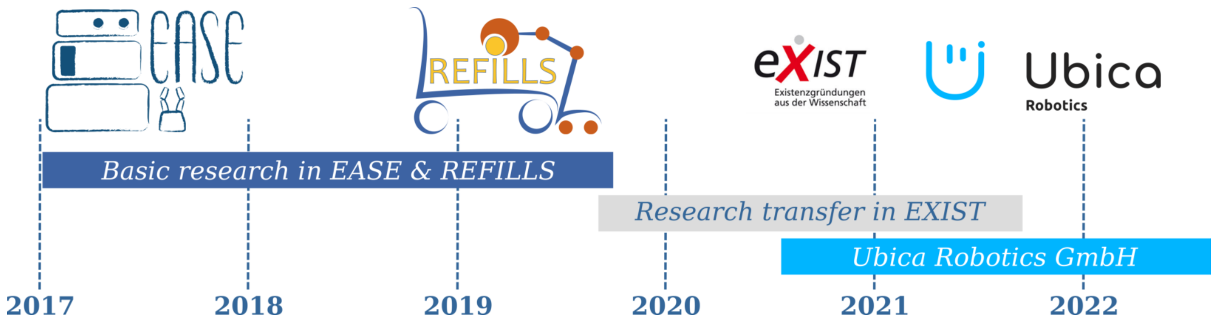


Figure 1.55: Robotics start-up company "Ubica" is based on EASE research results

is the forum in which researchers, developers, and users of AI technologies in Bremen meet, which builds the network between industry and research in Bremen, and which provides the local job forum for AI professionals.

The government of the federal state of Bremen and Bremerhaven has announced theestablishment of an AI transfer center to support small and medium-sized enterprises and increase the visibility of Bremen as an excellent location for AI research. EASEINNOVATION will participate and be a key partner of the transfer center. This participation will substantially strengthen innovation and transfer activities initiated by EASE basic research.

1.3.1.2.2 EASEOPENLAB Encouraging open research and facilitating research cooperation is a key part of the EASE mission statement (see Section 1.2.2.3). EASEOPENLAB organizes and manages the activities along this dimension. As shown in Figure 1.54 the EASEOPENLAB mainly includes the national and the international cooperations as well as temporary research communities. The national and international cooperations were already described in Sections 1.2.4.2 and 1.2.4.3.

EASE manages and participates in temporary topical research communities. The most important one is the focus group "Cognition-enabled robot agents" at the Hanse Wissenschaftskolleg in Delmenhorst (duration: 1.1.2019 - 31.12.2021). The ultimate goal of the study group is to support the establishment of the research area cognition-enabled robot agency and to position the CRC EASE as a key player in this field. In order to accomplish this goal, the focus group discusses recent developments in the areas of cognitive architectures. On that basis it sets an agenda for research on architectures for robot agents. An important secondary goal is to work towards building an international community of experts in Everyday Activity Science and Engineering, who investigate the problem of intelligent physical agency from a holistic perspective. Major outcomes of the focus group will be a special journal issue on "Cognition-enabled Robots: Mastering Everyday Activities" and a combined textbook and Massive Open Online Course (MOOC) on the topic of the focus group.

A second temporary research community that is influential for the research agenda of EASE is the ZiF Research Group "Cognitive behavior of humans, animals, and machines: Situation model perspectives" at the University of Bielefeld with Helge Ritter being one of the co-coordinators. This interdisciplinary research group with internationally leading researchers in cognitive psychology, cognitive neuroscience, AI, and robotics tries to establish a framework, called situation model of cognitive behavior, with a homogenized vocabulary that enables convergent research activities in the understanding of flexible and context-sensitive behavior for accomplishing agent goals. EASE researchers Helge Ritter, Kerstin Schill, David Vernon, and Michael Beetz are fellows of this research group. This research activity is invaluable for the EASE research agenda because the EASE generative model for accomplishing everyday manipulation tasks can be viewed as a manifestation of the situation model.

EASE also participates in several national and European research and innovation networks. EASE is a member of euRobotics AISBL, an international non-profit association that collaborates with the European Commission (EC) to develop and implement a strategy and a roadmap for research, technological development and innovation in robotics. EASE participates in the Confederation of Laboratories for AI Research in Europe (Claire.AI[67]) and the European research network for AI (AI4EU[68]).

In order to promote the connection and cooperation with other researchers we are also executing international research outreach projects, including the Transatlantic AI-based Robotics project (TRANSAIR[69]), which organized the virtual conference on the democratization of AI robotics research [70].

Strong research cooperations on EASE-related topics were realized with a number of universities, e.g. with the Seoul National University (Prof. Tak Zhang) on symbolic/subsymbolic robot imitation learning, with the TU Wien (Prof. Marcus Vincze) on robot vision and joint RoboCup@Home competition team, with the University of Costa Rica (Prof. Federico Ruiz) on autonomous robot manipulation, University of Cluj Napoca (Prof. Sorin Herle) with respect to research student exchange, and with Orebro University (Prof. Alessandro Saffiotti) on cognitive robotics (in preparation).

International cooperation is also promoted through funded projects such as ILIAS, which was introduced as part of EASEinnovation (Section 1.3.1.2.1), and AI4HRI. AI4HRI (Artificial Intelligence for Human-Robot Interaction) is a trilateral French-Japanese-German basic research project conducted together with internationally leading research groups in human-robot interaction. The Robotics and InteractionS (RIS) group at LAAS led by Rachid Alami and HRI Lab at Kyoto University (led by Takayuki Kanda) will investigate software architectures and knowledge representation and reasoning for human-robot interaction. AI4HRI is intended as preparatory research for EASE phase 3, which is planned to focus on multi-agent everyday activity (see Section 1.2.2.4).

EASE also repeatedly participated in the Google Summer of Code programme, where EASE researchers advised international doctoral students on EASE-related programming projects funded by the "Google summer of code" program.

EASE is also promoting inner-university cooperation with other research focus areas, most notably the focus areas of material science (MAPEX[71]), logistics (LogDynamics[72]), and ocean science (Marum[73]). In all cases the cooperation aims at the application of cognition-enabled robots in the respective science and application fields.

- Senior researchers that cooperated with EASE through longer research stays included:
  - Prof. Dr. Haizhou Li, National University Singapore, 2019 - 2022, Research Excellence Chair
  - Prof. Dr. David Vernon (Carnegie Mellon University, Kigali, Rwanda): 03.07. - 17.07.2017 and 24.06. - 13.07.2019
  - Prof. Solomon Teferra Abate (Addis Ababa University, Ethopia)
  - Prof. Byoung Tak-Zhang (Seoul National University) 10.01. - 04.02.2018
  - Prof. Martha Yifiru Tachbelie, Alexander Von Humboldt Foundation with George Forster Research Fellowship for Experienced Researchers (Addis Ababa University, Ethopia)
  - Prof. Tetsunari Inamura and Mizuchi Yoshiaki, PhD. National Institute of Informatics Tokyo, Japan, 11.06. - 28.07.2019
  - Prof. Federico Ruiz Ugalde, Universidad de Costa Rica, Costa Rica, 02.01. - 14.07.2020
  - Dr. Jae Hee Lee (Cardiff University, United Kingdom), 25.11. - 06.12.2019

Early career researchers who were sent to EASE from research partners to increase the level of research cooperation included:

- Jose L. Part, 17.07. - 28.07.2017, Heriot Watt University, Edinburgh, United Kingdom
- Pawel Gajewski und Paulo Albeha, 16.10. - 22.10.2017, University of Aberdeen, Scotland, UK

[67] `claire-ai.org`
[68] `ai4eu.eu`
[69] `transair-bridge.org`
[70] `transair-bridge.org/conference-2`
[71] `uni-bremen.de/mapex`
[72] `logdynamics.de/`
[73] `marum.de`

- Christopher Paxton, 04.12. - 14.12.2017, Johns Hopkins University, USA
- Nestor Garcia Hidalgo, 14.04. - 14.07.2018, Universitat Politecnica de Catalunya, Spain
- Myat Sun Yin, Pochara Sangtunchai, Amonnat Tengputtipong, Kitiphong Duwa, 31.05. - 16.08.2018, Mahidol University, Thailand
- Jin-Young Lee, Chung-Yeon Lee, Kibeom Kim, Sung-Jae Cho, Seung-Jae Jung, Joon-Ho Kim, Beom-Jin Lee, 04.08. - 29.08.2018, Seoul National University, South Korea
- Mizuchi Yoshiaki, PhD (National Institute of Informatics, 09.01. -15.02.2019, Tokyo, Japan
- Mohammed Diab Elsayed Sharafeldeen, 21.01. - 06.03.2019, Universitat Politecnica de Catalunya, Spain
- Marco Costanzo, 16.05. - 09.08.2019, Universita degli Studi della Campania “Luigi Vanvitelli”, Napoli, Italy
- Michail Theofanidis, 16.05. - 29.08.2019, Texas University in Arlington, Texas, USA
- Razvan Gambutan, 15.07. - 30.09.2019, TU Cluj-Napoca, Romania
- Israel Chaves Arbaiza, 02.01. - 31.03.2020, Universidad de Costa Rica, Costa Rica
- Shingo Kitagawa, 03.02. - 03.03.2020, JSK Robotics Laboratory, University of Tokyo, Japan
- Derrick Odonkor, 17.02. - 24.04.2020, Carnegie Mellon University, Kigali, Rwanda
- Anna de Groot, 01.03. - 03.03.2020, remote cooperation due to COVID situation: 29.04. - 27.06.2020, University of Amsterdam, The Netherlands

1.3.1.2.3 openEASE The activities of openEASE will be described in Section 1.3.3.4 (Web-based knowledge service infrastructure openEASE).

1.3.1.2.4 EASEoutreach The activities of EASEoutreach will be described in Section 1.4.4 (Knowledge transfer and public relations).

1.3.1.2.5 EASEacademy and EASE learning hub The activities of EASEacademy and the EASE learning hub will be described in Section 1.3.1.3 (Impact on teaching and training).

1.3.1.3 Impact on teaching and training

1.3.1.3.1 EASEACADEMY The EASEACADEMY is the institution that manages and runs the collection of teaching and training activities offered by EASE.

1.3.1.3.1.1 EASE integrated research training group EASE has established the EASE integrated research training group (EASE IRTG). The EASE IRTG supports EASE in qualifying the doctoral students for the research programme and for their later careers in academia, industry, and administration. Its aim is to integrate the interdisciplinary doctoral students through graduate and postgraduate education rooted in a strong research foundation. It offers doctoral students a comprehensive training programme in the foundations of everyday activity science and cognitive robotics and key skills. EASEACADEMY and in particular EASE IRTG is directed by Prof. Michael Beetz and managed by Dr. Jörn Syrbe. EASE IRTG promotes the development of young scientists in EASE and affiliated research projects at the University of Bremen and the EASE partner institutions in terms of both research and personality. The research and training programs are created for individual qualification and development. The main focus is on the scientific excellence of every graduate's individual doctoral research and thesis.
The EASE IRTG aims

- to promote and strengthen the creation of a convergent, interdisciplinary team of young researchers from disciplines including computer science, artificial intelligence, robotics, linguistics, cognitive psychology, and neuroscience in order to overcome scientific scientific and cultural barriers and create the best possible synergies between the individual disciplines;
- to provide a framework for a well-balanced, structured doctoral education; and
- to foster scientific independence of the young researchers.

Situated in a highly interdisciplinary research environment, training in the IRTG aims to equip young researchers not only with in-depth expertise in a researcher's main focus area but also with disciplinary and methodological breadth, which opens up many avenues for the young researchers' future academic career. Furthermore, given that the research environment in which the young scientists receive their trainig is focused on the development of integrated AI systems, trainees are also particularly well-equipped to pursue a career in industry as, for example, a CTO. Consequently, though training of young researchers in the context of EASE may be considered very challenging and demanding, the acquired knowledge and skill profile is worth the effort, because it provides the young researchers with unique opportunities to continue their career in both science and industry.

To support young researchers in the training process, the EASE Graduate School provides the following measures. The supervisors closely guide and counsel the students' research, while at the same time encouraging and supporting the students in growing into the resposibilities of an independent researcher. In addition to the support from the supervisor's methodological assistance, the EASE Graduate School offers a training program according to the students' needs. One important component of this program are the annual EASE Fall schools (1.3.1.3.1.3). The Fall schools provide the students with the necessary basic methodological skills through lectures by international guests as well as state-of-the-art workshops and tutorials. The Fall schools are complemented by EASE-specific training events such as expert workshops, gender workshops, topic group, and regular meetings.

All these events are part of the strategy to enable the EASE graduate students to become proficient and responsible researchers. During personal meetings, the students plan their professional development supported by the graduate school management and supervisors. Modeled on the Vitae Researcher Development Framework (RDF), the researchers' training covers four different areas of professional development.

The first area contains the knowledge and intellectual abilities needed to be able to carry out excellent research. Area two involves the personal qualities, career, and self-management skills required

Figure 1.56: Article reviewing at the EASE doctoral students retreat 2019

to take ownership and control of professional development. The knowledge of the standards, requirements, and professional conduct that are needed for the effective management of research are part of area three. Area four covers the knowledge, understanding, and skills required to engage with, influence, and impact the academic, social, cultural, and economic context. The RDF helps to identify the personal strength of the EASE graduate school members. During regular meetings, the students are asked to determine which of their skills they would like to improve.

The University of Bremen provides access to the *Researcher Development Framework Overview* digital platform for the doctoral students of the EASE integrated research training group. The Researcher Development Framework (RDF) is a major new approach to researcher development, to develop world-class researchers and build a research base. The RDF is a professional development framework for planning, promoting and supporting the personal, professional and career development of researchers in higher education. It articulates the knowledge, behaviors and attributes of successful researchers and encourages them to realise their potential.

Parts of these efforts are workshops hosted by partners of the University of Bremen and self-hosted workshops. These self-hosted workshops are provided during regular meetings, special lunch-to-lunch-events, doctoral retreats, and doctoral schools, or the EASE Fall Schools.

1.3.1.3.1.2 Structures in the doctoral training In the EASEAcademy, the following structures have been established in order to support the doctoral students in organizing the scientific work, train their transferrable and soft skills and assist them in planning their carrers.

Brown-bag Meetings: The Brown-bag Meetings are the meeting point for all doctoral students of the EASEAcademy. The Brown-Bag meetings exclusively address members of the EASEAcademy and provide the opportunity to share experiences, discuss insights, and receive feedback from peers. On every three weeks base, the meetings are used for formal and informal interaction. Each meeting has two slots for presentations, such as rehearsal talks, work-in-progress demonstrations, and research challenges. Every graduate is asked for feedback, and as requested for possible solutions.

Doctoral agreements: In addition to the doctoral registration at the University of Bremen and partner universities, the EASE Academy provides a doctoral agreement, including the title of a student's thesis and a set of objectives the student wants to achieve. These objectives are mainly focused on transferable skills, like soft skills, self-perceptions, and career planning. This doctoral agreement is annually updated and supports doctoral students to gain competences apart from their technical skill and knowledge.

Doctoral students retreat: The annual EASE Doctoral Retreat is organized to allow the EASEAcademy members to step back from their everyday work and concentrate on their professional devel-

Figure 1.57: Pictures from the second EASE fall school on cognition-enabled robot agents.

opment. The professional development includes setting members' effectiveness goals, team building, and networking across the research area borders. The primary tool of the retreats is the RDF planner, its usage, and benefits. The planner provides the opportunity for self-reflection from different perspectives, e.g., the intellectual abilities and the personal qualities to be an effective researcher. The doctoral retreat is also used to discuss and review the latest research of the field. In small groups, students examine suggested articles and document the benefits of the articles.

Transferable skills training: The members of the graduate school are encouraged to visit the Bremen Young Researcher Development Courses (BYRD). BYRD is the center for early-career researchers at the University of Bremen. Its focus is on transferable skills, like presenting skills, scientific writing, and voice and body coaching. In addition to the courses offered by BYRD, the EASEACADEMY has conducted targeted soft-skill seminars like two-day workshops for female researchers in presentation training and personal development, a seminar on patent law and one on science communication with social media.

1.3.1.3.1.3 EASE Fall Schools The main components for teaching and training the subject knowledge needed by the EASE doctoral students is taught in the EASE Fall Schools, which the students are requested to attend. The schools offer a combination of introductory and specialized tutorials through internationally leading experts. Their purpose is to teach robot manipulation for everyday activity research and technologies combined with hands-on courses and interactive programming exercises in the afternoons. The lectures are made publicly available in the EASE learning hub (see Section 1.3.1.3.5). The EASE Fall Schools are planned as annual events:

- **2018 EASE Fall School on Cognition-enabled Robot Manipulation**[74] (organized by Jörn Syrbe and Michael Beetz). In 2018, the first EASE Fall School for cognition-enabled robot manipulation welcomed 35 EASEACADEMY members and five guests from Germany and Romania. The subjects of the fall fchool ranged from knowledge representation, the usage of episodic memories for robots' tasks, machine learning to virtual robotics. The topics were provided by EASE experts, but also partners from academia and industry. Prof. David Vernon and Prof. Markus Vincze introduced the challenges of cognition-enabled robotics and robot Vision. Practical applications of these topics was

[74] ease-crc.org/ease-academy/fall-school-2018

part of the hands-on courses organized by experienced EASEACADEMY members. *Speakers/Lecturers: Michael Beetz, Jörn Syrbe, David Vernon, Michael Suppa, Markus Vincze, Moritz Tenorth, Daniel Nyga, Jürgen Sturm, Christoph Schütte.*

- **2019 EASE Fall School on Cognition-enabled Robot Manipulation**[75] (organized by Jörn Syrbe and Michael Beetz). The EASE Fall School 2019 opened the lectures on cognition-enabled robot manipulation to over 60 students, among those were 20 international students and 40 from the University of Bremen. In 2019, the fall school again provided lectures and hands-on experience held by leading experts from the University of Bremen and invited international partner researchers. Prof. Kei Okada (University of Tokyo), Prof. Rachid Alami (LAAS CNRS), Prof. Frank Guerin (University of Aberdeen), Dr. Timothy Patten (TU Wien), and Dr. Ilaria Tiddi (VU Amsterdam) provided insights into humanoid robots, human-robot joint actions, the development of knowledge graphs for everyday activities, decisional issues in human-robot interaction, and robot manipulation in open environments. A special business topic event concluded the program of the EASE Fall School. *Speakers/Lecturers: Michael Beetz, Jörn Syrbe, Kai Okada, Rachid Alami, Ilaria Tiddi, Holger Schultheis, Animesh Garg, Gayane Kazhoyan, Sebastian Koralewski, Andrew Melnik, Asil Bozcuolu, Georg Bartels, Alexis Maldonado, M. Tabachnyk, Andrei Haidu, Timothy Patten, Frank Guerin, Ferenc Balint-Benczedi.*

Two more doctoral schools were planned for 2020: The ICAPS-ICRA Summer School on Plan-Based Control for Robotic Agents Paris, June 8-12, 2020 is postponed and the 2020 EASE Fall School on Cognition-enabled robot manipulation and AI-based Robot Planning 2020 is cancelled due to the COVID-19 situation.

1.3.1.3.1.4 Open Source Software Packages In addition to publications and networking activities, providing open-source software packages can considerably strengthen a young researcher's profile and, thus, increase her visibility in the scientific community. Accordingly, the EASE CRC and IRTG provide the infrastructure, context, and encouragement for the development and hosting of such software packages. A complete list of open-source software developed in EASE can be found in Section 1.2.3.10.2.2.

1.3.1.3.2 Academic programmes With its course offerings, EASE substantially contributes to the bachelor program in informatics. It also plays a key role in establishing a new master program in AI systems at the University of Bremen.

1.3.1.3.2.1 Contributions to the bachelor program in Informatics

- **RoboCup@Home** In the RoboCup@Home challenge, the contestants develop robotic solutions for typical everyday activites like cleaning up or storing groceries in a competitive environment. The IAI Bremen has buil a team of undergrad as well as PhD students to participate in the RoboCup@Home German Open in 2019 in Magdeburg, Germany. To tackle the challenges, the students utilized the EASE software framework to implement solutions for perception, manipulation, knowledge and high-level planning problems. All components were integrated into a complete system on a Toyota HSR platform to fulfill the "storing groceries" task. We planned to participate in the RoboCup@Home 2020 as well, but due to the COVID-19 pandemic the event was cancelled by the organizers. However, we are still actively extending the system and training new students how to work with it, planning to participate in RoboCup@Home 2021.
- **Bremen Big Data Challenge** The Bremen Big Data Challenge aims at sparking interest in data research among students in Bremen. The Challenge 2017 focused on the load of the Studentenwerk Bremen's university cafeteria. Participants were supplied with the cafeteria's load in five-minute-slots from 2009 to 2015 and supplementary data such as the cafeteria menu, semester times and

[75] `ease-crc.org/fall-school-2019`

weather. Their task was to predict the cafeteria load in the five-minute-slots of the year 2016. 24 teams participated in that year's challenge and achieved a range of good results with the best team's prediction missing the true number of receipts issued in a five-minute-slot by just 8.6 receipts. Combining the top 5 results, the number of receipts issued was missed by only 8.28 receipts.

- **Game Engine in AI** The IAI Bremen established a new bachelor course “Game Engines in AI”. The course is designed to give students an introduction to game engines as well as the application of them in the context of AI research. The course focuses on introducing Unreal Engine 4 and AI topics research related to EASE. This includes, among other topics, the representation and control of robots in simulations, human representation in simulations, and robot vision. Students are provided with hands on exercises that give an in depth experience how to use the covered topics and prepare them for working in a scientific environment. Course participants will be well positioned for a Bachelor thesis or position as working student in EASE.
- **SUTURO** (joint course with linguistics department) is a 1-2 year bachelor/master practical programming course established by the IAI, in which teams of students develop all components that are necessary for programming a robot to perform simple manipulation tasks in human environments, such as knowledge-based visual perception, plan-based behavior and motion control. The student teams use EASE software components and receive coaching from doctoral students who are experts in the respective fields. The SUTURO projects aim at preparing students for the participation in the RobotCup@Home challenge. The projects are designed to attract high-potential PhD students.
- **IMPROVER** The goal of the course IMPROVER is to train the next generation of software designers and engineers who are capable of creating software components for the robot agents of tomorrow. The course covers the whole development process of a cognitive robotic agent and enables students to understand and apply (1) AI-based methods of robot perception and control, (2) Artificial intelligence (knowledge representation and reasoning, statistics-based learning and reasoning, task planning, decision-based control and machine learning), and (3) the principles of cognition (cognitive architectures, learning and development, prospection, memory, internal simulation and metacognition). The course was designed for students in the last semester of their bachelor's degree or in the first semester of their master's degree. It builds on the course "Integrated Intelligent Systems" of the Master of Computer Science, which it replaced. The course will be combined with expert tutorials from EASE Fall Schools for structured doctoral training. In addition, the course is part of the KI-Campus Project which is funded by the BMBF.

1.3.1.3.2.2 Master program for AI systems The Department of Informatics at the University of Bremen proposes a new master of science program in artificial intelligence and intelligent systems (AIIS). The proposed program is designed to give students a comprehensive framework for artificial intelligence with specialization in one of three areas: cognition-enabled robotics, machine learning and data science, or cognitive assistants. Students will engage in an intensive core curriculum intended to develop depth in all core concepts that build the foundations for artificial intelligence theory and practice. Students will also be given the opportunity to build on the core knowledge of AI by taking specializing and elective courses, including selected ones from other departments, to explore key contextual areas or more complex technical applications. Program graduates will be well positioned to attain research and development positions in a rapidly growing field or to pursue doctoral degrees in related fields.

1.3.1.3.3 Tutorials at scientific events EASE researchers have also held tutorials at international conferences, other scientific events, and doctoral training schools:

- Mehul Bhatt, Carl Schultz. “Declarative Spatial Reasoning – Theory. Methods. Applications,” IJCAI 2018.
- Mehul Bhatt, Jakob Suchan. “Cognitive Vision: On Deep Semantics in Visuo-Spatial Computing,” AAAI 2018.
- Mehul Bhatt, Jakob Suchan. “Spatial Cognition in the Wild: Methods for Large-Scale Behavioural Research in Visuo-Locomotive Perception,” Tutorial, ETRA 2018.

- Mehul Bhatt, Carl Schultz, Przemyslaw Walega. “Declarative Spatial Reasoning – Theory. Methods. Applications,” IJCAI 2017.
- Michael Beetz. “Automated Models of Everyday Activity,” CITEC Summer School “Cognitive Architectures,” Excellence Centre CITEC (Cognitive Interaction Technology), University of Bielefeld, Germany, 2017.
- Michael Beetz, Daniel Beßler. “Reintegrating Robotics & AI,” Summer School on Foundations of Robotics and Autonomous Learning, Berlin, Germany, 2017.
- Mehul Bhatt, Jakob Suchan. “Spatial Cognition in the Wild: Methods for Large-Scale Behavioural Research in Visuo-Locomotive Perception Tutorial,” ACM Symposium on eye tracking reserahc & applications, Warsaw, Poland. 14.06. - 17.06.2018.
- Daniel Beßler. “KRR for robots,“ SemWeb outje 2019: “Robots, Language and Theory of Mind,“ 13.05, Hotel Zuiderduin, Egmond aan Zee.
- Mehul Bhatt, Jakob Suchan. “On Deep Semantics for Explainable Visuospatial Computing Tutorial,” Cognitive Vision, 29.08. - 08.09.2020, ECAI 2020.
- Michael Beetz. “Knowledge Representation and Reasoning for Cognition-enabled Robot Manipulation,” October 19 - 20, ICAPS 2020.

1.3.1.3.4 EASE seminar series

Date	Speaker	Title
13.07.2017	Prof. Dr. David Vernon (Carnegie Mellon University, Africa)	Cognitive Architectures – Roles, Requirements, and Realization
18.07.2017	Jose L. Part (Heriot Watt University, Edinburgh, UK)	Teaching Robots through Situated Interaction
18.12.2017	Prof. em. Keith Clark (Imperial College London, UK)	Rule Control of Teleo-Reactive, Multi-tasking, Communicating Robotic Agents
23.01.2018	Prof. Byoung-Tak Zhang (Seoul National University, South Korea)	Teaching Robots to See, Hear, Talk & Act Like Humans Using Videos
23.01.2018	Dr. Daniel Kohlsdorf (XING SE, Germany)	Minimizing the Rage: One Step At a Time
06.02.2018	Prof. Dr. David Lane (Heriot-Watt University, UK)	ORCA Hub: Offshore Robotics For Certification of Assets
11.04.2018	Prof. Bruno Siciliano (University of Naples, Italy)	Robotic Dynamic Manipulation
29.05.2018	Dr. Moritz Tenorth (Magazino GmbH, Germany)	Robotics Startup Magazino: Recent developments and lessons learned from an academic perspective
05.06.2018	Dr. Huan Lin (Bangkok University, Thailand)	Explanation Generation in an ITS for Dental Surgical Skill Training
25.09.2018	Dr. Karinne Ramirez (Chalmers University, Sweden)	A Semantic Reasoning Method for the Recognition of Human Activities
08.10.2018	Dr. Markus Funk (TU Darmstadt)	Cognitive Learning Support using Augmented and Virtual Reality
06.11.2018	Dr. Sandra Buchmüller (TU Braunschweig)	Geschlechtergerechte Technikforschung
06.11.2018	Prof. Luc Steels (Vrije Universiteit Brussel, Belgium)	Emergent communication on real robots

7.11.2018	Prof. Dr. Herbert Jaeger (Jacobs University, Bremen)	Controlling and shaping neural dynamics with conceptors
13.11.2018	Prof. Dr. Volker Krüger (Lund University, Sweden)	Using Robot Skills for Industrial Tasks
15.11.2018	Prof. Dr. Gerhard Kraetzschmar (Hochschule Bonn-Rhein-Sieg)	Software Development for Robotics
27.11.2018	Dr. Ingmar Posner (Oxford University, UK)	Robots Thinking Fast and Slow
05.12.2018	Dr. Robert Ross (Technological University Dublin, UK)	Squeezing Shared Meaning from Behaviour
16.01.2019	Stefano Borgo (Laboratory of Applied Ontology Trento, Italy)	How to build a culture-aware robot
17.01.2019	Dr. Patrick Courtney (tec-connection, Konstanz)	The Analytical Laboratory as a Place for Robotics and AI
24.01.2019	Mohammed Diab (Universitat Politècnica de Catalunya, Spain)	A knowledge-based planning framework for smart and autonomous manipulation robots
11.03.2019	Carlos Corbato (TU Delft, The Netherlands)	ROS and robot control architectures from open source to metacontrol
29.04.2019	Remi van Trijp (Sony Lab Paris, France)	Why is Construction Grammar Important for Robotics?
07.05.2019	Prof. Kei Okada (University of Tokyo, Japan)	Task instantiation based on long-term experience memory
14.08.2019	Prof. Krishna Rajan (SUNY Buffalo, USA)	Materials Informatics for Analyzing Chemical Complexity
29.08.2019	Michail Theofanidis (University of Texas, USA)	Deep reinforcement learning methodology based on dynamic movement promitives
11.09.2019	Alberto Olivares Alarcos (Institut de Robotica i Informatica Industrial, Spain)	Knowledge Representation for Industrial Collaborative Robotic Tasks
25.09.2019	Elizabeth Croft (Monash University, Australia)	Social Work: Collaborative behaviours that measurably improve human-robot interaction
12.02.2020	Prof. Dr. Markus Vincze (TU Wien, Austria)	Perception of Robots
06.03.2020	Prof. Joachim Hertzberg (Osnabrück University)	Applications of plan-based robot control in agriculture and other domains

1.3.1.3.4.1 On-site workshops for secondary school students In a joint effort with the SMILE project, EASE has aimed at increasing the proportion of female bachelor students in Informatics by informing secondary school students about the computer science course programs and professional opportunities as well as motivating them to obtain degrees in informatics. To this end, EASE female master and doctoral students visited secondary schools in Bremen and Bremerhaven as role models to conduct information events. This way informatics and artificial intelligence was introduced to 526 female secondary school students (grade 8 to 12) and their teachers.

Figure 1.58: EASE Workshops for secondary school students in cooperation with the SMILE project

In addition, we developed two onsite workshop formats with EASE teachers that gave insights into the EASE research field:

- “Mein smarter Roboter” (*My smart robot*) is a 2-day workshop for female secondary school students of grades 5 and 6, in which the students interactively learn what service robots in human environments can do for people and how they have to be programmed. This is done using a graphical programming language for the Pepper robot and its simulation environment. EASE conducted the workshop 5 times in cooperation with the SMILE project. In total, 76 secondary school students attended the workshops.

- the SMILE/EASE workshop “Bring Pepper into the game” for female students of the grades 7 to 9, in which the students programmed Pepper with a game app (7 participants).

Based on demands from parents, EASE offered a workshop for male secondary school students (grades 5 and 6) with the title “Was macht Roboter smart? (what makes robots smart?)” in the context of the exhibition “Einfach wissenswert 2019: Robotik und KI”10. It was attended by 15 students.

1.3.1.3.4.2 Outreach to general public The activities of EASEoutreach will be described in Section 1.4.4 (Knowledge transfer and public relations).

1.3.1.3.5 EASE learning hub EASE considers the provision of open-access teaching and training materials for cognitive robotics to be an essential tool for outreach in the research community and for recruiting talented students. The EASE learning hub provides a variety of resources which were created by members of the EASE collaborative research center to support education and research in the interdisciplinary field of everyday activity science and engineering. Video lectures, tutorials, learning materials from online and classroom courses, software tools, datasets, and hands-on activities allow users to explore current research challenges, learn about the computational and empirical methods used to study human and machine intelligence, and experience the excitement of the latest discoveries in the field.

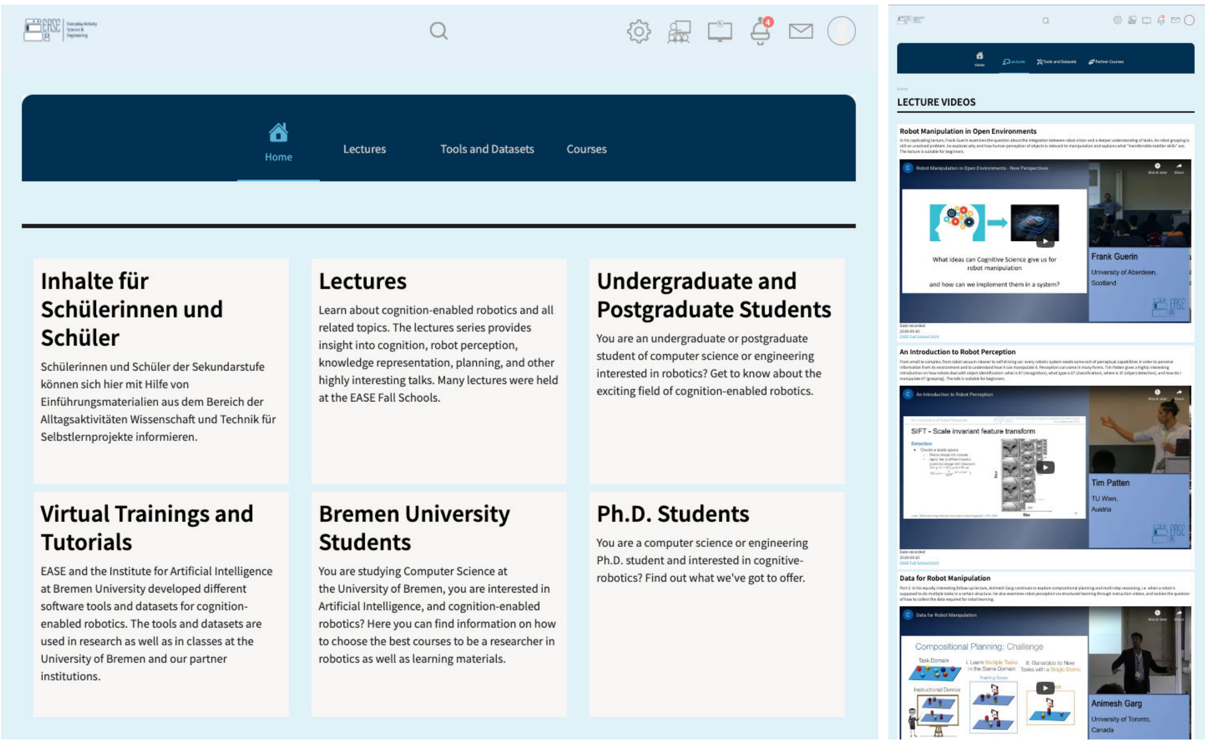


Figure 1.59: EASE learning hub website.

1.3.1.3.5.1 Complete lecture courses and lecture units

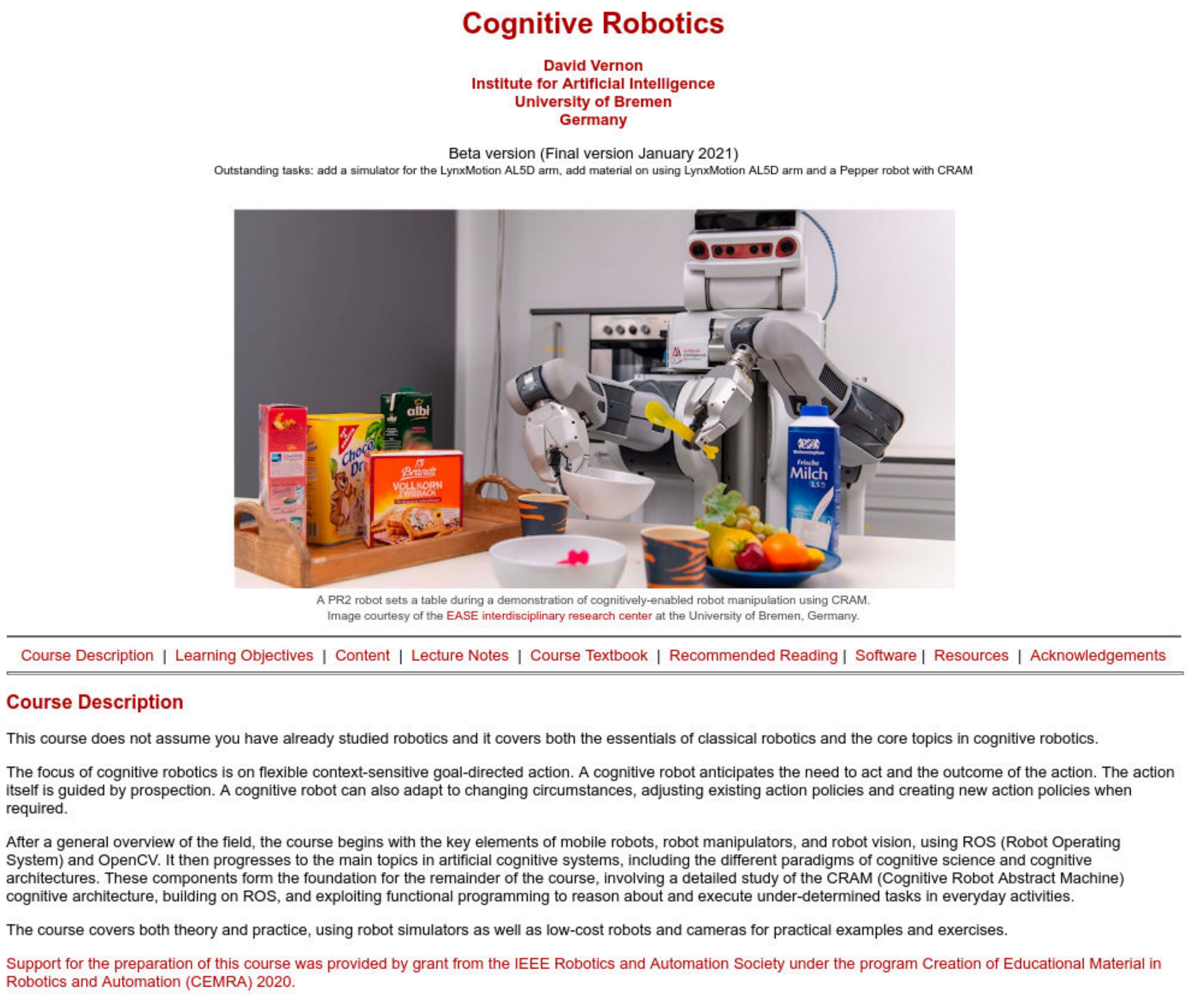
EASE contributes substantially to the generation of teaching material for two complete bachelor-/master-level lecture courses in the academic field of cognitive robotics, which EASE is positioned in. The teaching material includes videos, textbooks/scripts, slides, and interactive teaching material. The material covers the basic module "Integrated intelligent systems" and a course in the specialization "cognition-enabled robotics" of the master's program "AI systems".

- The first of the two mentioned courses is an **"Introductory Course on Cognitive Robotics"**[76] developed by David Vernon in cooperation with EASE and supported through an IEEE RAS-funded project for the creation of educational material in robotics and automation (CEMRA Projects).

  This course covers both the essentials of classical robotics (mobile robots, robot arms for manipulation, and robot vision) and the principles of cognition (cognitive architectures, learning and development, prospection, memory, knowledge representation, internal simulation, and meta-cognition). It brings these components together by working through some recent advances in robotics for everyday activities. It also includes practical and detailed material based on the CRAM (Cognitive Robot Abstract Machine) cognitive architecture, incorporats the KnowRob knowledge base, utilizes ROS (Robot Operating System) and exploits functional, object-oriented, and logic programming to reason about and execute under-specified tasks in everyday activities. The course emphasizes both theory and practice and makes use of physical robots as well as robot simulators for visual sensing and actuation. *A course textbook is in preparation.*

- The second course contributed by EASE is **"Design and implementation of cognition-enabled robot agents"**, which is under development for the digital MOOC learning platform for the AI Campus ("KI-Campus"). The course is a winning entry (14 winners from 137 entries) for the German AI Campus competition and is implemented as a project executed together with ZMML (Zentrum für Multimedia in der Lehre) in the period from winter term 2020 to 2021.

[76] `cognitiverobotics.net`



Cognitive Robotics

David Vernon
Institute for Artificial Intelligence
University of Bremen
Germany

Beta version (Final version January 2021)
Outstanding tasks: add a simulator for the LynxMotion AL5D arm, add material on using LynxMotion AL5D arm and a Pepper robot with CRAM

A PR2 robot sets a table during a demonstration of cognitively-enabled robot manipulation using CRAM.
Image courtesy of the EASE interdisciplinary research center at the University of Bremen, Germany.

Course Description | Learning Objectives | Content | Lecture Notes | Course Textbook | Recommended Reading | Software | Resources | Acknowledgements

Course Description

This course does not assume you have already studied robotics and it covers both the essentials of classical robotics and the core topics in cognitive robotics.

The focus of cognitive robotics is on flexible context-sensitive goal-directed action. A cognitive robot anticipates the need to act and the outcome of the action. The action itself is guided by prospection. A cognitive robot can also adapt to changing circumstances, adjusting existing action policies and creating new action policies when required.

After a general overview of the field, the course begins with the key elements of mobile robots, robot manipulators, and robot vision, using ROS (Robot Operating System) and OpenCV. It then progresses to the main topics in artificial cognitive systems, including the different paradigms of cognitive science and cognitive architectures. These components form the foundation for the remainder of the course, involving a detailed study of the CRAM (Cognitive Robot Abstract Machine) cognitive architecture, building on ROS, and exploiting functional programming to reason about and execute under-determined tasks in everyday activities.

The course covers both theory and practice, using robot simulators as well as low-cost robots and cameras for practical examples and exercises.

Support for the preparation of this course was provided by grant from the IEEE Robotics and Automation Society under the program Creation of Educational Material in Robotics and Automation (CEMRA) 2020.

Figure 1.60: Cognitive robotics course by David Vernon using CRAM as the reference cognitive architecture and applying the CRAM interactive tutorial.

The course teaches a holistic, system-oriented perspective of cognition-enabled robotics that focusses on AI-based methods for robot perception and control, including body motions and object manipulation, as well as for knowledge representation and reasoning, probabilistic learning and reasoning, action planning, decision-theoretic action, and machine learning. It also addresses principles of cognitive capabilities, including cognitive architectures, learning and development, prospection, memory, internal simulation, and meta cognition.

The system components and lecture modules are presented within a uniform framework and illustrated through state-of-the-art robot agents that accomplish everyday manipulations tasks.

In addition to the complete lecture courses, EASE contributes the Chapter **"Knowledge representation and reasoning for cognition-enabled robots"** (authored by Michael Beetz and Daniel Nyga) to the book "Robotics goes MOOC" (Bruno Siciliano, ed.). This book is part of the Springer MOOC & BOOK project, providing both a MOOC – offered through Federica Web Learning – and a Springer reference book based on the online course. This approach combines the quality of a scientific essay with the communicative power of an online educational product. The MOOC provides a state-of-the-art overview of various aspects of the rapidly developing field of robotics. The book is strictly linked to the MOOC and includes numerous examples and exercises in addition to those offered in the MOOC. Moreover, it features multimedia content such as videos and augmented reality which can be accessed

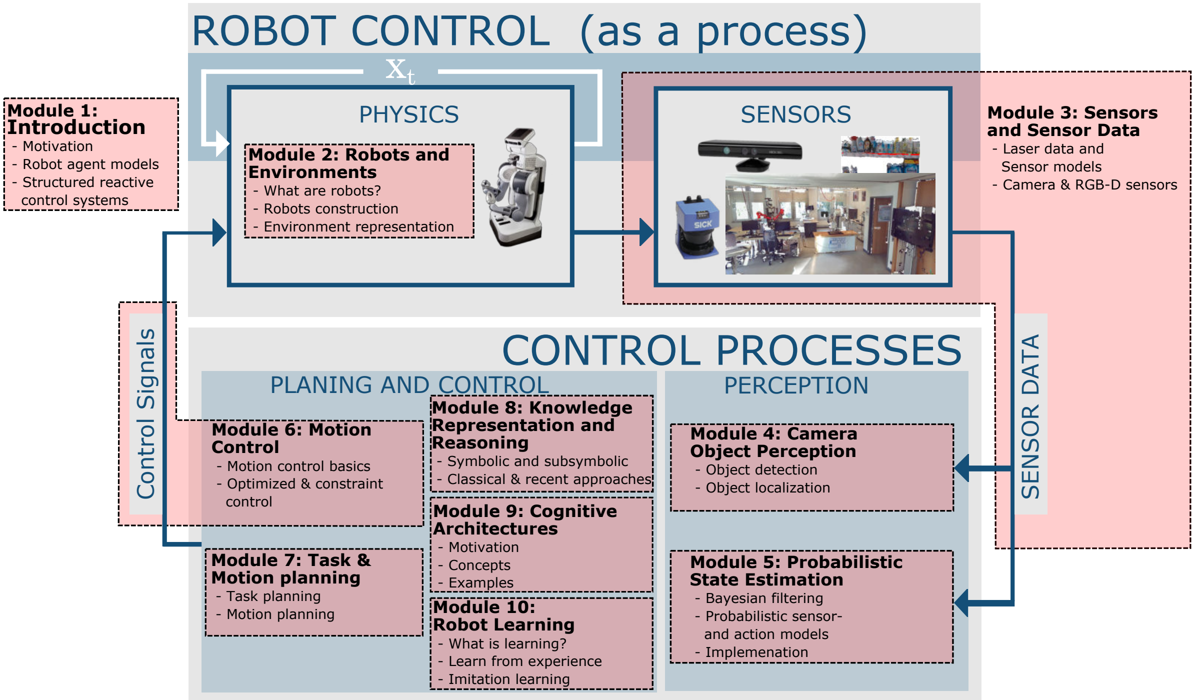


Figure 1.61: KI Campus MOOC (Massively Open Online Course) "Cognition-enabled robot agents".

via PC, tablet or any other mobile device. Students who buy the book can easily access this content through the Springer Multimedia App.

Finally, two chapters in the Cognitive Robotics Handbook, edited by Angelo Cangelosi and Minuro Asada (MIT Press, in press) are featuring EASE material (Chapter 10: Cognitive Architectures and Chapter 21: Reasoning and Knowledge Representation). They provide entry-level lecture material and references to more in-depth publications for the respective topics.

1.3.1.3.5.2 EASE Topic Modules EASE topic modules are lecture units giving in-depth presentations of EASE research topics including a video and slides. They are intended for the education and training of doctoral students who want to conduct research in cognitive robotics and senior researchers who intend to enter the field. The lecture units are given by internationally renowned experts in the field and were recorded at the EASE doctoral schools.

Cognitive architectures for robotic agents	
An overview of cognitive architectures.	David Vernon
The role of memory in cognition.	David Vernon
Cognitive Architecture design and the Common Model of Cognition.	David Vernon
Cognitive architectures for robot agents.	Michael Beetz
Open Research and the Soar Cognitive Architecture	John Laird
A short Socratic Dialogue on Action and Intelligence	Eadom Dessalene, Yiannis Aloimonos
Knowledge representation and reasoning	
Knowledge representation and reasoning for robot agents. part 1 & part 2.	Michael Beetz
Probabilistic Knowledge Acquisition and Representation for Natural-language Applications.	Daniel Nyga
Digital Twin Knowledge Bases for Robot Agents.	Michael Beetz
Mining and Explicating Instructions for Everyday Activities	Johannes Pfau

Embodied Semantics for the Language of Action and Change	Mihai Pomarlan
Perception	
Visions for Robotics. part 1 & part 2.	Markus Vincze
Robot perception for real-life applications.	Michael Suppa
Robot Perception — An Introduction.	Tim Patten
Questions Answering about 3d Scenes.	Maxim Tabachnyk and M. Katzmann
Semantic Collision Detection and Proximity Query	Toni Tan
High Precision Hand Tracking using a Marker-based Approach	Janis Roßkamp
Dataset of Binocular and RGB-D images annotated with 6d Poses	Jesse Richter Klug
Robot learning	
Deep Learning for Autonomous Robot Manipulation (part 1, part 2).	Animesh Garg
Task Instantiation from long-term Memories of Mobile Robots.	Kei Okada
Object manipulation	
Robot Manipulation in Open Environments — New Perspectives.	Frank Guerin.
Humanoid Robots in Everyday Activities.	Kei Okada
OMPL for Motion Planning	Lydia Kavraki
Human Robot Interaction	
Introduction to human-robot joint action.	Rachid Alami
Challenges for Decision Making in Human Robot Interaction.	Rachid Alami
That Ain't Right – AI Mistakes and Black Lives	Chad Jenkins
Applications of cognitive robots	
Google's Cloud Robotics.	Jürgen Sturm & Christoph Schütte
Mobile pick-and-place robots in the real world — Lessons learned from academia to startup.	Moritz Tenorth
Founding a Start-up in autonomous Robotics.	Georg Bartels and Alexis Maldonado
Tutorial on CRAM (Cognitive Robot Abstract Machine)	Gayane Kazhoyan
Miscellaneous	
How Do We Build The Next Internet?	Radu Rusu
Research Administration in Open Science	Jan Andersen
Virtual Prototype Based Verification	Vladimir Herdt

Additional lecture units are planned to be added through future EASE Fall Schools.

Figure 1.62: Interactive CRAM tutorial downloadable in a virtual box.

1.3.1.3.5.3 Interactive tutorials Another important category of teaching material developed in, and provided by EASE, are the interactive tutorials of the key software components.

A comprehensive one-hour tutorial[77] of the CRAM plan language and executive and an accompanying video[78] are accessible online. The CRAM tutorial material is organized in the CRAM tutorial website[79]. The tutorial website contains the comprehensive introductory tutorial that can be downloaded in a virtual box, as well as more than 25 other tutorials at beginner, intermediate, and advanced level. The tutorial infrastructure has been developed and is used for AI programming courses at bachelor and master level at the University of Bremen and for the practical courses of the EASE Fall Schools.

We have further developed a set of tutorials dedicated to the use of EASE research data. A uniform interface to this data is provided through the query-answering component of the OPENEASE knowledge service which has been implemented using the KNOWROB knowledge base. The webpage of OPENEASE provides several interactive tutorials covering different aspects of research data representation, retrieval and analysis[80]. Here, the programming exercises can be directly typed into the page and the answers are then graphically visualized. The tutorials have been developed in the scope of two scientific workshops and the EASE Fall schools.

[77] `cram-system.org/tutorials/demo/fetch_and_place`
[78] `youtube.com/watch?v=N-wPeBZ2Kjs`
[79] `cram-system.org/tutorials`
[80] `data.open-ease.org/tutorials`

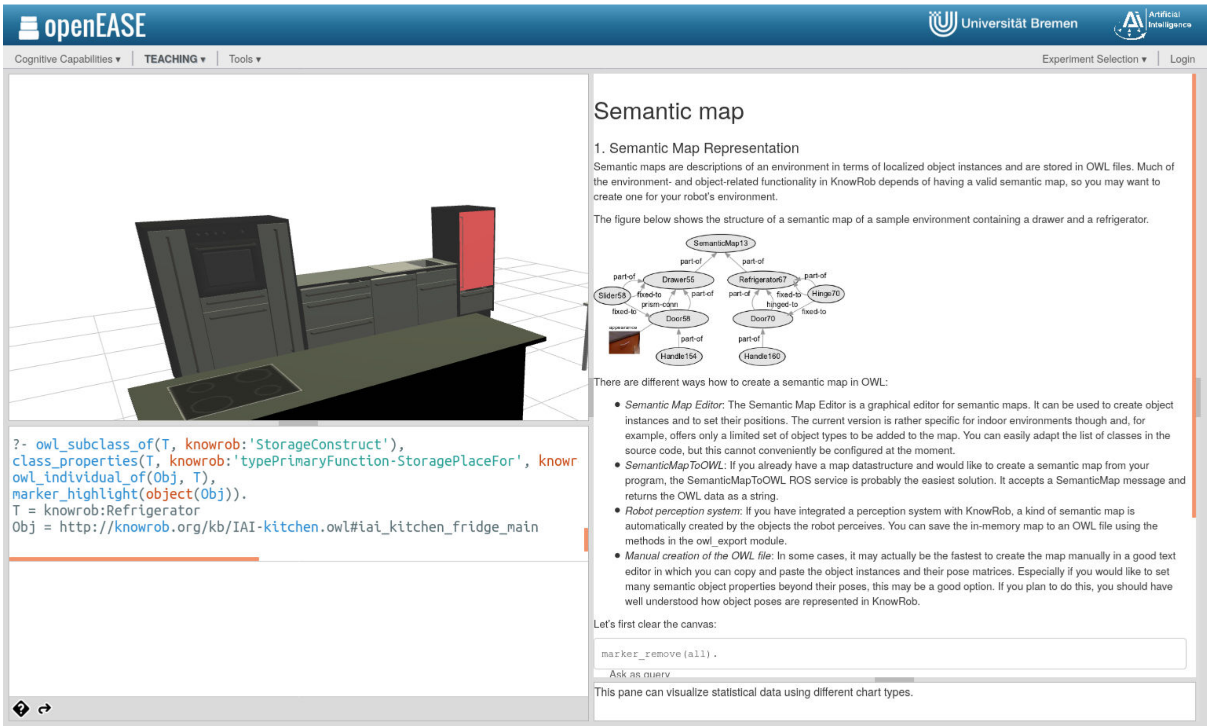


Figure 1.63: Interactive KNOWROB tutorial accessible through the OPENEASE web site.

The ROBOSHERLOCK perception executive also features extensive tutorials that can be accessed publicly[81]. To provide minimal friction when getting started with ROBOSHERLOCK, we implemented an interactive, web-based system that allows students and researchers to create containers with all the system dependencies for a full ROBOSHERLOCK installation. Our solution also includes the necessary installation of the ROS dependencies and allows to visualize the 2D and 3D outputs that perception experts can yield during their execution. This allows beginners to get immediate visual feedback how the sensor data is processed and which results have been generated.

The tutorials feature a large variety of typical tasks when adapting ROBOSHERLOCK to new domains, e.g., adding new perception experts, extending the perception knowledge base, or using the query-based perception interface. We also provide an overview of querying logged perception data to analyze the annotations generated by the perception experts retrospectively on real world robot data. Our interactive tutorials have been used to teach ROBOSHERLOCK to students as well as researchers in the EASE Fall Schools.

[81] robosherlock.org/tutorials.html

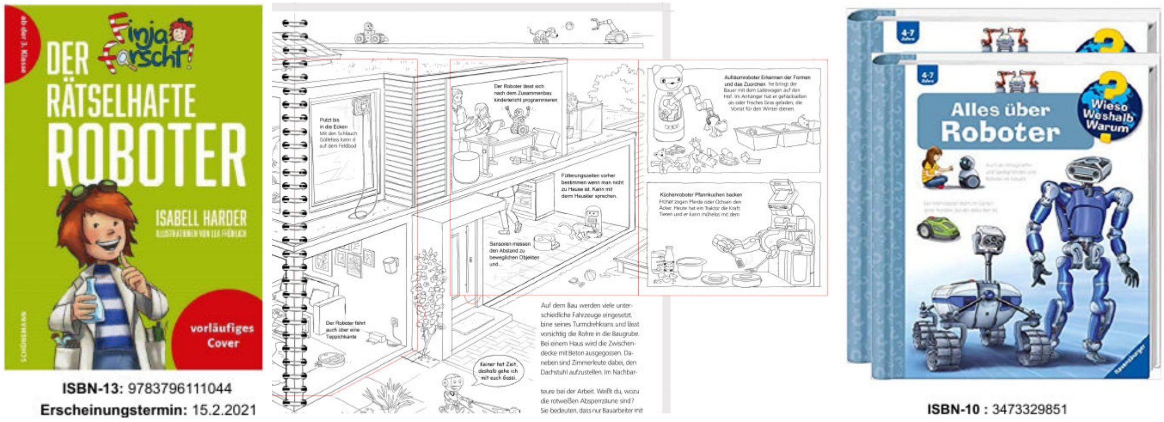


Figure 1.64: Educational children's book "Finja forscht", coloring picture, and Ravensburger children's book.

1.3.1.3.5.4 Educational young readers' book "Finja forscht" is an educational book series authored by the University of Bremen's transfer coordinator Isabell Harder. It addresses several research topics across multiple disciplines from the perspective of a 7-year-old child who is visiting researchers at the University. In the episode "Der Rätselhafte Roboter" (*"The Mysterious Robot"*), the EASE Robotics Lab (Section 1.3.3.1) is introduced and several topics related to cognition-enabled robotics in EASE are adressed.

1.3.1.4 Support from the University and Federal State of Bremen

The relationship between the EASE CRC and the University of Bremen has been excellent all the time, as expected. The University of Bremen supports EASE through numerous measures, including the recent hiring of professors in cognitive psychology Bettina von Helversen (General Psychology, since May 2019) and Markus Janczyk (Research Methods and Evaluation, since April 2019). These research areas are highly relevant for EASE and will broaden the field of available competences significantly.

The university intends to strengthen the EASE PI team by establishing an additional unscheduled 5 year W2 professorship position fitting thematically into the EASE CRC topics to be ideally filled by a promising young female researcher. At the time of proposal writing a concept for financing this position has become apparent.

The university will support EASE with a third researcher position, not only for the duration of the next funding period, but for the complete intended project runtime until 2029. This position is targeted to a female postdoctoral researcher.

As the EASE CRC is at the core of the activities in the research focus area "Minds, Media, Machines (MMM)" it will implicitly profit from the synergies of the several new support measures for the area, in particular from the (1) staffing of the focus area with proposal and research managers, (2) the founding of a university-wide data science center, and (3) the establishment of a roof for the graduate training in the area MMM.

In addition, the federal state of Bremen has provided start-up funding for establishing an AI transfer center with the EASE central lab being one of the living labs in the center, which will considerably strengthen our efforts towards establishing a virtual research infrastructure which will promote international cooperation as well as EASEINNOVATION.

The cooperation between DLR and Universität Bremen has been intensified through a joint Young Investigator Group headed by Daniel Leidner[82].

The university supports EASE with sufficient lab space, office rooms and infrastructure for smooth running of the project. Due to complicated building regulations in Bremen, the planned reconstruction

[82] `futuro.dlr.de/ease`

of the EASE robotics laboratory could not be achieved as quickly as originally planned. Extraordinary and additional consultations with the departments for construction and operation as well as with the approving authorities were necessary since the start of the project. Only this made it possible to make the necessary structural arrangements for the new laboratory and to enable us to organize alternative preliminary testing cases within the existing test facility for the meantime. After further extensive planning and preparation work in the last months, we now expect that the new laboratory can go into operation at the beginning of the next funding period. Permits and construction work did indeed lead to delays in the full use of the lab space. However, all departments of the university have always done their best to ensure that the lab can be used fully equipped as quickly as possible.

1.3.2 Staff situation

EASE has assembled a very strong team of principal investigators with four researchers among the **2020 AI 2000 Most Influential Scholars**[83] in Artificial Intelligence, with Albu-Schäffer (21) and Beetz (4) in the subfield of robotics, Lutz (30) in knowledge engineering, and Drechsler (96) in chip technology. The impact in particular in robotics are even higher with three additional researchers being supervised by EASE researchers (Rusu (2), Blodow (22), and Tenorth (87). This is based on the citation numbers in the most prestigious conferences and journals over the last ten years. All of them contribute to core components of the EASE research agenda.

Another strength of the team are the PIs that lead the investigation of complete AI-based robot control systems, including Beetz (cognition-enabled autonomous robot manipulation), Cheng (cognitive humanoid robot), Albu-Schäffer (high-performance robot control and mechatronics), Ritter (cognitive manual manipulation), and Vernon (developmental cognitive robotics in the context of the iCub project). Knowing and being aware of the pitfalls in realizing complete, autonomous robots that accomplish human-scale manipulation tasks is not possible without such expertise.

Several researchers have already led large integrated cooperative projects such as the CoTeSys excellence cluster (Beetz), the CITEC excellence cluster (Ritter), the Institute of Robotics and Mechatronics (Albu-Schäffer), the ATR research lab (Cheng), and the iCub European project (Vernon). The level of integration in the EASE project around key components reflects that huge experience in leading cooperative projects.

In addition to that unique expertise in integrated AI-based robot control systems, EASE has outstanding expertise in knowledge representation and reasoning for robots, with a particular focus on ontology representation, reasoning, and engineering. Here, Lutz brings in the expertise of basic research on ontological reasoning, Bateman advances the leading-edge of the modular design of hybrid ontologies, Malaka together with the post-doctoral researcher Porzel, who acquired and coordinates the FET-open project “Meaning and Understanding in Human-centric AI” (MUHAI), pushes embodied cognition and in particular in combination with developing formal representations of image schemata. Beetz is an expert in running knowledge representation and reasoning as essential cognitive capability in robot control systems. Beßler, who is an expert in the standardization of ontologies for autonomous robots and for leveraging the ontological knowledge for autonomous robot control, had the software lead on the KnowRob knowledge representation and reasoning system in the first phase. He is expected to defend his dissertation thesis before the end of the first funding phase and EASE intends to promote and support his career towards becoming a principal investigator in the second part of the next funding phase. The impact and synergy created through the research teams led by these PIs is enormous. Most essential for EASE is the work towards a common ontology across all EASE research areas and common data homogenization and formalization of NEEMs that is adopted by all EASE subprojects.

The third area of expertise is the modelling and understanding of human everyday activity. Here, we have established and continue the core PI team around Schultz, Schill, and Herrmann. Schultz is

[83] aminer.cn/ai2000/robotics

an expert for cognitive technical systems for human-machine interaction focussing on the application of leading-edge machine learning methods to biosignals, think aloud protocals, and full-body motion tracking. She also leads the University of Bremen initiative towards the establishment of a large-scale initiative towards the lifelong assistance of humans in everyday activity through cognitive technical systems. Schill is a renowned expert in cognitive neuro informatics and vice-coordinator of EASE, vice president of the DFG, and member of various scientific advisory boards for research competitions in the context of the AI and digitalization as well as open data and research initiatives. Manfred Herrmann is the chair of the Department of Neuropsychology and Behavioral Neurobiology and contributes to is in an essential manner with his research on identifying specific reasoning tasks for everyday activity in the human brain based on MRI data. In order to intensify the research into model building, which is the focus of the second phase, EASE has complemented the existing team Didelez, an expert in probabilistic causal modelling, and von Helversen an expert in human decision making (see below). In addition, Vernon will be included to further strengthen the cooperation between the research areas.

The core competence fields are complemented by researchers that bring in more specific, additional expertise. Zachmann brings in the know-how in leading-edge computer graphics research, most notably in very fast collision detection and distance computing and fast and realistic modelling of environments and physical processes. Drechsler and Herdt bring in the expertise of formalizing and verifying that the behavior generated by robots guarantees stated requirements. We believe that this direction of research is essential because the EASE robot agents have a very high degree of autonomy and we want to ensure that safety as well as ethical requirements can be guaranteed.

The DFKI Robotics and Innovation center is not directly integrated into the basic research thread of EASE. Rather the cooperation takes place through EASEINNOVATION. Kirchner and Beetz jointly coordinate the University of Bremen activities in the context of the University of Bremen AI strategy. In this context, common projects include the digital innovation platform and ecosystem project KNOWLEDGE4RETAIL, the Marie Curie international training network REMARO, and a precompetition project CERA4HRI (Cognition-enabled Robot Agents for human robot interaction) for establishing a competence center for human robot interaction.

New principal investigators EASE could acquire Prof. Dr. Vanessa Didelez as an additional principal investigator. Vanessa Didelez, who joined the University of Bremen in July 2016, is Professor of Statistics and Causal Inference and Deputy Head of the Department Biometry and Data Management of the Leibniz Institute for Prevention Research and Epidemiology - BIPS. Vanessa Didelez will substantially strengthen the probabilistic and causal modelling expertise in EASE. This will enable EASE to intensify the research towards proposing a probabilistic semantics of the EASE generative model (see page 20 of Section 1.2.3.1). She will be principal investigator in Subproject H01 (Sensory-motor and causal human activity models for cognitive architectures).

The EASE competence in experimental cognitive psychology is strengthened by incorporating Prof. Dr. Bettina von Helversen. Bettina von Helversen is an expert in cognitive decision making (denomination of professorship: General Psychology) in the Faculty Human and Health Sciences of the University of Bremen. Bettina von Helversen will investigate the decision making of humans in the EASE household challenge task within the conceptual framework of the EASE generative model. She will particularly focus on learning how to perform the decision making tasks in everyday activity tasks.

Prof. Dr. David Vernon is a key addition to the very core of the EASE research agenda in phase 2. David Vernon is a world-leading expert in cognitive robotics and artificial cognitive systems. He is author of the MIT Press textbook “Artificial Cognitive Systems – A Primer” and serves as one of the co-chairs of the IEEE Robotics and Automation Society (RAS) Technical Committee for Cognitive Robotics which received the 2017 RAS Most Active Technical Committee Award. David Vernon will have essential roles in investigating the CRAM 2.0 cognitive architecture, the cognitive architecture cross area research team, and strengthening the bridge between human and robot models of everyday activity.

Dr. Daniel Leidner is a homegrown young research scientist who substantially advanced his academic career within the EASE research center. Daniel Leidner is a research scientist in the Department of Autonomy and Teleoperation of the DLR Institute "Robotics and Mechatronics" and investigates with his team the planning and execution of movements on multi-arm robot systems with many degrees of freedom. He did his doctoral degree at the University of Bremen, is winner of the Georges Giralt PhD Award 2018 (best European PhD thesis in robotics), and winner of the Helmholtz Doctoral Prize 2018 in Aeronautics, Space, and Transport. He is principal investigator of a young researcher group "Failure and Uncertainty Tolerant Universal Robot Operation", which is co-located at the University of Bremen and strongly cooperates with the EASE research center. He will be principal investigator of the new EASE Subproject R06-N.

Dr. Vladimir Herdt will replace as a young principal investigator Prof. Daniel Große, who was a young principal investigator in the first phase but joined the Johannes-Kepler University Linz as a Full Professor and head of the group for Complex Systems.

New associated PIs Michael Suppa is honorary professor for Informatics/Robotics/Cognition since March 2020 and associate member of the IAI since 2015. He is an international expert in robot perception and industrial applications of advanced robot systems. Since 2013, he coordinates the topic group perception with SPARC and was appointed deputy institute director of the Institute of Roboctics and Mechatronics in August 2015. Michael Suppa published over 60 papers and was nominated for and received several best paper awards. In March 2015 he co-founded Roboception GmbH, a DLR spin-off company devoted to advance the state of the art in 3D sensors and vision. Besides being CEO of Roboception he is also responsible for business models as well as strategy and technology development. He is strengthening EASE in the area of cognitive robot perception and EASEINNOVATION in establishing industrial cooperations around deploying AI-based robot technologies.

Postdoctoral researchers with PI potential In addition, Dr. Felix Putze, who has substantially contributed to research area H and NEEM standardization and collection, has applied for an Emmy-Noether research group, in which he intends to cooperate with EASE. Further, EASE intends to promote the young research careers of Gayane Kazhoyan, who is the lead doctoral researcher for the CRAM plan executive, and Daniel Beßler, who is the lead doctoral researcher for the KNOWROB knowledge representation and reasoning systems. We intend to replace in 2023 John Bateman as a principal investigator in Subproject P04 with Daniel Beßler. John Bateman will resign from his PI-role in Subproject P04 in 2023 due to lifetime working time reasons and continue as an EASE research professor. Daniel Beßler who has already intensively cooperated with John Bateman on the EASE ontology and NEEM standardization will take over the PI role (6 joint publications in the first phase of EASE). We plan to start a strategic EASE research project focussing on realizing and integrating the concepts of the cognitive architecture CRAM 2.0 in the CRAM architecture, for which Gayane Kazhoyan would be the ideal PI. Finally, we are hiring Dr. Maria Hedblom as a postdoctoral researcher to strengthen the ontology engineering efforts of EASE in particular concerning the formalization of image schemas for everyday manipulation actions.

1.3.2.1 Selected promotions and awards

The international recognition of the EASE team of principal investigators is also reflected in the promotions and awards that they have received in the first funding period:

- Kerstin Schill
  - has become the Director of the Hanse-Wissenschaftskolleg (Institute for Advanced Study) in Delmenhorst,
  - was elected as a Vice President of German Science Foundation (DFG),
  - has been selected as a member in various committees and advisory boards including (*) the selection committee of the Carl Zeiss Foundation, (*) the DFG expert committee "Wissenschaft

im Digitalen Zeitalter", (*) the supervisory board at KIT, (*) the Academic Advisory Board at Köln University, (*) the Expert Committee AI Strategy Bavaria "KI-Wettbewerb", and (*) the Selection Committee Hessisches Ministerium für Wissenschaft und Kunst.

- Tanja Schultz
  - has become elected member of the European Academy of Sciences and Arts (2017),
  - has been selected as an IEEE Fellow in 2020 for contributions to multilingual speech recognition and biosignal processing.
  - has received a Google Faculty Research Award "EMG2Speech" recognizing and supporting world-class faculty pursuing cutting-edge research in areas of mutual interest.
  - has been selected as a Member of Board of Directors for the Leibniz Science Campus on Digital Public Health Leibniz-WissenschaftsCampi
  - has become Member of the Advisory Board of the UKRI (UK Research and Innovation) Center for Doctoral Training (CDT) in Speech and Language Technologies (SLT) at the University of Sheffield
- Karinne Ramirez-Amaro has become Assistant professor in the Electrical Engineering Department of Chalmers University of Technology (Sweden), Division of Systems and Control.
- Daniel Große has become Univ.-Prof. and Head of the Institute for Complex Systems at the Johannes Keppler University Linz.
- Mehul Bhatt has become Professor in the School of Science and Technology, Örebro University and directs the CoDesign Lab.
- Daniel Leidner
  - has received the George-Giralt Prize ror the best robotics dissertation in Europe, 2018.
  - won the Helmholtz Doctoral Prize 2018 in Aeronautics, Space, and Transport Helmholtz Association.
  - has been awarded a DLR VO-R Young Investigator Group Grant 2019. Together with University of Bremen and in cooperation with EASE.
  - was named Selected Innovator Under 35. MIT Technology Review Germany 2019.
- Rolf Drechsler
  - has been appointed as Adjunct Professor at Indian Statistical Institute, Kolkata, India,
  - received the ASP-DAC (China Asia and South Pacific Design Automation Conference) Prolific Author Award for publications between 1995 and 2020.
  - received the Berninghausen award for excellent teaching at the University of Bremen (2018),
  - Koselleck Awardee for the proposal "PolyVer: polynomial verification of electronic circuits",
  - awarded "AI 2000 Most Influential Scholars Honorable Mention" in 2020 for Chip Technology (# 96 in terms over publications in the last ten years).
- Rainer Malaka:
  - was elected as the Chair of IFIP Technical Committee 14 Entertainment Computing
- Alin Albu-Schäffer
  - won an ERC Advanced Grant M-Runners in 2019 for for mobility and running in legged robots

- Gordon Cheng:
  - has received the William Mong Distinguished Award — for AI in the real world: from neuroscience to robotic innovations
  - was elected as IEEE Fellow 2017 for contributions in humanoid robotic systems and neuro-robotics
- Michael Beetz
  - has been awarded “AI 2000 Most Influential Scholar” in 2020 in Robotics (# 4 in terms over publications in the last ten years) and three former doctoral students received “AI 2000 Most Influential Scholar/s Honorable Mention”
  - received an Honorary doctorate of the University of Örebro, Sweden in 2019.
- Carsten Lutz:
  - has been awarded “AI 2000 Most Influential Scholars Honorable Mentions” in two AI subfields: knowledge engineering and IJCAI/AAAI.

1.3.2.2 Projects contributing to EASE

	Project	Duration	PI
1	KNOWLEDGE4RETAIL	2019-2022	Beetz
2	POLYVER	2020-2025	Drechsler
3	MUHAI (FET Proactive)	2020-2024	Malaka, Porzel
4	TRANSAIR	2019-2021	Beetz, Drechsler
5	KI Campus	2020-2021	Beetz
6	REMARO	2020-2023	Beetz
7	KI-SIGS	2020-2023	Malaka
8	ILIAS	2019-2022	Beetz
9	ROPHA	2017-2020	Beetz
10	SMILE	2017-2020	Schill
11	Delphi Studie	2019	Sozialwissenschaftliches Methodenzentrum
12	SCORE (DAAD), Bozen	2018-2020	Bateman
13	ASARob: Aufmerksamkeitssensitiver Assistenzroboter	2017-2020	Schultz, Schill, Herrmann
14	Graduate School Karlsruher Decision and Design Lab (KD2Lab)	pending	Schultz, Herrmann
15	ALMED	2019-2021	Schultz
16	AI4HRI	2021-2023	Beetz
17	CERA4HRI	2021-2023	Beetz, Malaka
18	Graduate school “System Design”	2012	Drechsler
19	TRACEBOT	2012	Beetz, Vincze
20	FUTURO	2020-2023	Leidner
21	AI Transfer Center	2021	Beetz, Malaka
22	Ubica	2019	Beetz
23	Graduate School “Empowering Digital Media”	2017	Malaka
24	Data Science Center of the University of Bremen	2019	Drechsler

Table 1.4: List of the most important cooperative projects and institutionsEASE investigators were or are involved in.

1.3.3 Research infrastructure

1.3.3.1 Central EASE robot laboratory

The EASE central robotics laboratory is a kitchen environment consisting of a countertop, major appliances and storage cabinets and a dining table with chairs. The laboratory is equipped with working places for about ten researchers. EASE currently has four autonomous mobile manipulation platforms available that can be used for manipulation experiments (see Figure 1.65).

The robots are controlled through opensource ROS software libraries with the same higher level software. The robots use the same action library, that is interface layer to the low-level robot control system that is robot-specific. The kinematic structure and the sensor equipment is specified in the Semantic Robot Description Language (SRDL), which is used to program robot control software independently of the robot hardware that it runs on. Much of the software is shared between the different EASE partners.

Figure 1.65: Robots in the EASE central robotics laboratory.

Figure 1.66: The EASE BASE laboratory for recording human everyday activities (left) and the experimental kitchen setup at EASE partner DLR with the autonomous mobile manipulation platform *MobileJustin*.

1.3.3.2 EASE human activity laboratory

The Biosignals Lab at the Cognitive Systems Lab (CSL, Figure 1.66) consists of the EASE Biosignal-Aquisition-Space (EASE-BASE) for multimodal, mobile recordings and a booth for stationary, high-accuracy recordings. EASE-BASE is equipped with a nine camera OptiTack motion capture system, six stationary RGB-cameras, depth cameras, far field microphones, and two large projectors for interactive feedback to the experiment participants. The soundproof and electromagnetically shielded biosignal-recording booth contains a stationary EEG system, a high-density EMG recorder, and an fNIRS system.

Both spaces can be used for multimodel recordings with a wide range of available mobile sensors like eye tracking, wet or dry electrode based EEG, EMG, IMU, EDA, blood pressure and microphones. Additionally, EASE-BASE provides for Augmented and Virtual Reality experiments (HoloLens 1 & 2, HTC Vive, Meta 2).

1.3.3.3 EASE sandbox robot laboratory

The central robot laboratory which is introduced in Section 1.3.3.1, is also completely modeled as a virtual reality and simulation environment including the robot, as is shown on the left-hand side of Figure 1.67. The robot control programs for the simulation environment and the physical environment are identical and differ only at the low-level control, which means the EASE research software runs without change on both the real and the simulated robots.

In addition, the laboratory is also modeled with a digital twin knowledge base where its core component KnowRob introduces background knowledge and articulation models of doors and drawers, such that the robot agent can reason about all aspects of the environment, which is shown on the righ-hand side of Figure 1.67. As a result, the laboratory provides access to machine-readable state, semantic and background information about objects and robots within the virtual lab environment.

The laboratory also features, for instance, reconfigurability of the environment, replacement of objects, deployment of different robots to allow for experimentation with generalization and opentask domains. Consequently, the virtual laboratory combined with the digital twin knowledge base will provide researchers a test bed of unique level of detail, visual realism and cognitive capabilities enabling state-of-the-art contributions to AI research problems in EASE. Ultimately, such sandbox laboratory will contribute to reproducibility and comparability of experimental results and a reference setup for further research. The sandbox robot laboratory will be made available to EASE researchers through Subproject F.

Figure 1.67: The EASE central robot laboratory modeled as a virtual environment. The furniture and the appliances on the left and the hand-size objects on the right.

1.3.3.4 Web-based knowledge service infrastructure openEASE

OPENEASE is a storage, analysis, and visualisation platform for research data acquired from robots and humans performing everyday activities. It also provides software tools for visualising, reasoning with and analyzing the data. Thereby it improves understandability and reproducibility of research findings. Additional software tools are provided that allow to comprehensively record and semantically annotate everyday activities performed in several different set-ups including virtual reality, simulation and robot control. Episodes of certain tasks are stored individually, and can later be combined for analysis and visualisation purposes. The semantic annotations compile a high-level description of the activity: Which actions were performed, why and when the actions were performed, if and why an action has failed. These semantic labels are used to annotate research data such as motions carried out by the robot, forces that were recorded, and objects that were recognized. Such a comprehensive representation of everyday activities enables OPENEASE to provide useful analysis and visualisation interfaces which allow to deeply investigate the recorded data.

1.3.3.5 Computing infrastructure

The importance of available computing resources for state of the art research has increased over the last years with data-driven learning approaches. Deep learning for example requires large amounts of data to generalize well and optimize appropriately in the large parameter space of modern state of the art deep networks. This learning paradigm gave rise to the training of such deep networks with an increasing number of layers demanding extensive usage of GPU-based computing resources which.

Many approaches in the overall EASE system is dependent on available computing resources in the form of GPUs. Subproject H01 and H03 create generative models of human activity for a joint cognitive architecture, which involves the processing and inference from large sensory inputs into complex models. Subproject R02 is building advanced perception methods to be able to competently handle challenging objects from high-dimensional camera data. Subproject R05 is using deep reinforcement learning to combine our cognition-enabled plan executive with the expressiveness of machine learning methods. All of these exemplary mentioned subprojects as well as individual subtasks in the other subprojects are heavily depending on deep learning based approaches for their individual contributions.

Beside the ever growing need for GPU resources for deep learning, EASE makes extensive use of Game Engines for reasoning and learning. To simulate and render photorealistic and physically accurate environments for robot manipulation while running all cognition-enabled robot plan components, computional resources in the form of CPUs, Memory and GPUs are key to success. These resources will enable us to use Game Engines in realtime with the necessary simulation accuracy to generate episodic memories as well as supporting live prospection-based reasoning.

In order to supply the subprojects in EASE with the necessary resources, we want to provide a shared high-performance computing platform featuring the adequate CPU, memory and GPU resources for our joint research. The system will be developed as a platform with a common authentication scheme that allows the researchers to easily access the computing resources for their intended workloads. Due to the complexity of typical software dependencies and architectures in the different research areas, we are employing an container-based system to avoid interferences between the different software packages and versions. Our main goal is to provide a system for GPU-intensive workloads such as deep learning. We will also investigate methods for running Game Engine workloads on this platform, for example for long NEEM generation processes. In addition, a high-performance server architecture is needed to run the NEEMHub, an interface for upload and retrieval of NEEMs. The development of this platform within EASE is related to Subproject F (see Section F.4.2).

To support the research process even during exceptional peak loads, we will also incorporate cloud computing resources. This will allow us to flexibly scale our computing resources if necessary to cope with higher demands before milestones, conference deadlines or very large experiments.

1.4 Support structures

1.4.1 Early career support

The promotion of junior scientific personnel is a key objective of EASE and will be achieved through the following measures.

At the Bachelor and Master level, new lectures and practical courses will be introduced, in particular within the Master of Computer Science, with a focus on Artificial Intelligence, Cognition, and Robotics.

Talented students are integrated into EASE research projects as student assistant researchers from an early stage. EASE will send selected talented students to international partner institutions for conducting research for their Master theses, supervised by a principal investigator of EASE.

contract duration	no. of contracts doctoral students		no. of contracts postdocs		total
	female	male	female	male	total
1-12 months	2	6	0	3	11
12-24 months	0	3	1	0	4
24-36 months	2	4	0	0	6
36-48 months	2	17	0	5	24

Doctoral students will mainly be trained through the EASE Integrated Research Training Group (IRTG). This group provides research-oriented training in fields represented by the projects of EASE. Its purpose is the education, training, and mentoring of doctoral students participating in the projects of the CRC to support them in becoming leaders in their field. For a more detailed description of the IRTG, its activities in the first funding period and the measures planned for the next phase cf. the subproject MGK. should: 3. Young researchers in EASE are often the person in charge for an open-source software package, organize the respective website, and thereby gain visibility in the field at an early stage of their career. The core activities of the IRTG are augmented by individual measures such as the EASE mentoring programme for female researchers (cf. section 1.4.2), which offers long term research stays at internationally renowned research institutes (e.g., MIT).

1.4.1.1 Dissertation theses in the EASE CRC

Table 1.5: EASE Graduates and Doctoral Students

#	Name	Gender	Thesis Title	Status	Supervisors	Funding
1	Dr. Asil Kaan Bozcuoglu	male	Fast Robot Learning using Prospection and Experimental Knowledge A Cognitive Approach with Narrative-Enabled Episodic Memories and Symbolic Knowledge	Defense May 2019	Prof. Dr. hc. Michael Beetz, Ph.D. and Prof. Kei Okada	SFB 1320
2	Dr. Ference Balint-Benczedi	male	Task-adaptable, Pervasive Perception for Robots Performing Everyday Manipulation	Defense February 2020	Prof. Dr. hc. Michael Beetz, Ph.D.	BMWi
3	Dr. Daniel Leidner	male	Cognitive Reasoning for Compliant Robot Manipulation	Defense May/October 2017	Prof. Dr. hc. Michael Beetz, Ph.D. and Prof. Dr.-Ing. Alin Albu-Schäfer	DLR
4	Gayane Kazhoyan	female	Accomplishing Variations of Mobile Manipulations Tasks on Autonomous Robots Through Generalized Plans	Defense end 2020	Prof. Dr. hc. Michael Beetz, Ph.D.	SFB 1320
5	Jesse Richter-Klug	male	Uncertainty based pose estimation of naturally stored kitchen objects	expected in 2021	*Prof. Dr. Udo Freese*	SFB 1320
6	Constantin Uhde	male	Semantic Understanding of Human Activity	expected in 2021	Prof. Gordon Cheng	SFB 1320
7	Celeste Mason	female	Multimodal biosignal-based action recognition of semantically defined and hierarchically structured everyday activities	expected in 2021	*Prof. Dr. Tanja Schulz*	SFB 1320
8	Moritz Meier	male	Multimodal modeling of human everyday activities with an emphasis on top-down modeling based on verbal reports	expected in 2021	*Prof. Dr. Tanja Schulz*	SFB1320
9	Florian Ahrens	male	Human Neural Information Processing and Cognition Enabled Robotic Agents. Physiological Brain Recordings of Complex Everyday Activities - An fMRI and EEG Study	expected in 2021	Prof Dr. Dr. Manfred Herrmann	SFB 1320
10	Konrad Gadzicki	male	Modeling of Bio-inspired Pattern Recognition: from V1 Models to Convolutional Neural Networks	expected in 2021	Prof. Dr. Kerstin Schill	SFB 1320
11	Sebastian Koralewski	male	Improving Robotic Agents ability to perform everyday activity in a dynamic environment through learning from experience	expected in 2021	Prof. Dr. hc. Michael Beetz, Ph.D.	SFB 1320
12	Daniel Beßler	male	Representing and reasoning about robot actions with ontologies	expected in 2021	Prof. Dr. hc. Michael Beetz, Ph.D.	SFB 1320
13	Janis Roßkamp	male	-	expected in 2021	*Prof. Dr. Gabriel Zachmann*	SFB 1320
14	Thorsten Kluß	male	Multisensory and motor contributions to human flexible behavior	expected in 2021	*Prof. Dr. Kerstin Schill*	SFB 1320
15	Jan-Hendrik Worch	male	Perceiving Humans Performing Everyday Manipulation Activities	expected in 2021	Prof. Dr. hc. Michael Beetz, Ph.D.	SFB 1320
16	Jaime Maldonado	male	Modelling of human attention and stimuli integration for behaviour predictions in humans and autonomous systems	expected in 2021	*Prof. Dr. Kerstin Schill*	SFB 1320
17	Sebastian Höffner	male	Task parameter estimation for robotic agents	expected in 2021	*Prof. Dr. Rainer Malaka*	SFB 1320
18	Anneke Haga	female	-	expected in 2021	-	SFB 1320
19	Jakob Suchan	male	Declarative Reasoning about Dynamic Visuospatial Imagery	expected in 2021	*Prof. Dr. Mehul Bhatt*	SFB 1320
20	Marcel Walter	male	Design Automation for Field-coupled Nanotechnologies	expected in 2021	Prof. Dr. Rolf Drechsler	SFB 1320
21	Tim Meywerk	male	Formal Verification of Robotic Control Programs Acting in Real-World Environments	expected in 2021	Prof. Dr. Rolf Drechsler	SFB 1320
22	Toni Tan	male	Geometric Computing for Simulation-Based Robot Planning	expected in 2021	Prof. Dr. Gabriel Zachmann	SFB 1320
23	Johannes Pfau	male	Deep Player Behavior Modeling (Supervisor: Rainer Malaka - Co-Supervisor: Magy Seif El-Nasr)	plans to submit in 2020	Prof. Dr. Rainer Malaka	SFB 1320
24	Michael Neumann	male	Mental simulation and learning	expected in 2021	Prof. Michael Beetz, Ph.D.	SFB 1320

25	Petra Wenzel	female	The intelligent use of space in everyday activities	expected in 2021	PD Dr. Holger Schultheis; Prof. Michael Beetz, Ph.D.	SFB 1320
26	Mona Abdel-Keream	female	A Game-Based Learning Infrastructure for Human-Robot Interaction	expected in 2021	Prof. Michael Beetz, Ph.D.	University of Bremen / BMBF
27	Alexis Maldonado	male	-	expected in 2021	Prof. Michael Beetz, Ph.D.	BMWi
28	Georg Bartels	male	-	expected in 2021	Prof. Michael Beetz, Ph.D.	BMWi
29	Jose Rojas	male	Photorealistic Learning Infrastructure	expected in 2021	Prof. Michael Beetz, Ph.D.	University of Bremen
30	Martin Meier	male	Perceptual Grouping in Oscillator Networks	expected in 2021		SFB 1320
31	Simon Stelter	male	Knowledge Enabled Effect Aware Motion Control	expected in 2021	Prof. Michael Beetz	BMWi / EU
32	Alexander Wich	male	Learning Generative Models for Autonomous fetch-and-place tasks from observing human demonstrations (Robots Learning from Active Observation)	expected in 2021	Prof. Michael Beetz	BMwi
33	Feroz Ahmed Siddiky	male	Deep learning from Episodic Memory of Robotic Agents	expected in 2021	Prof. Michael Beetz	Grundhaushalt
34	Patrick Mania	male	Knowledge-enabled Robot Belief States in Game Engines	expected in 2021	Prof. Michael Beetz	DFG / Landesstelle
35	Andrei Haidu	male	Automated Models of Everyday Manipulation Activities in Virtual Reality Environment	expected in 2021	Prof. Micheal Beetz	SFB 1320
36	Jeroen Schäfer	male	Action-Aware Bimanual Manipulation	expected in 2024	Prof. Michael Beetz	SFB 1320
37	Dr. Sebastian Brunner	male	A Framework for Analyzable, Resource-Aware and Self-Optimizing Robot Longterm Autonomy	Septermber 2020	Prof. Michael Beetz; Prof. Dr. Christian Schlegel	DLR
38	Dr. Vladimir Herdt	male	Efficient Modeling, Verification and Analysis Techniques to Enhance the Virtual Prototype based Design Flow for Embedded Systems	Februar 2020	Prof. Dr. Rolf Drechsler	DFKI

1.4.1.2 Early career success stories in EASE

Several young researchers in EASE have achieved important new career stages during the first funding period. EASE PI *Mehul Bhatt* has become a professor for computer science at the School of Science and Technology of the Örebro University in Sweden. EASE researcher *Karinne Ramirez-Amaro* is an assistant professor in the research group Mechatronics at the department of Electrical Engineering of Chalmers University of Technology, Sweden, since September 2019. EASE PI *Daniel Große* has become professor and head of the 'Institut für Complex Systems' at the Johannes Kepler Universität Linz, Austria.

Daniel Leidner, who will be PI of the new EASE subproject R06-N, has successfully finished his PhD in the first funding period. His "summa cum laude" doctoral thesis was honored with many awards, including the Georges Giralt PhD Award 2018 (best European PhD thesis in robotics) and the Helmholtz Doctoral Prize 2018 in Aeronautics, Space, and Transport. Daniel Leidner is listed on the Business Punk Watchlist 2020 for Tech & Engineering and is a Selected Innovator Under 35 by MIT Technology Review Germany 2019. He is leader of the Rollin' Justin team and the Semantic Planning group at German Aerospace Center (DLR).

Zhou (Yuen) Fang became co-founder and CTO of the startup company ZoeysRooms. Former EASE researchers *Ferenc Balint-Benczedi, Georg Bartels*, and *Alexis Maldonado* founded the robotics startup *Ubica* supported by EXIST funding under PI Michael Beetz, in EASE incubator.

1.4.2 Gender equality and family-friendly policies

Gender equality is one of the guiding principles at the University of Bremen and there is a long-standing commitment to it. Its gender-equality concept received top ranks in the "Professorinnenprogramm" of the German Federal Ministry for Research and Education (BMBF) in 2008 and in the 2013 reevaluation. The latest gender-equality concept "geschlechtergerecht 2028" was rated with excellency in 2018. Additionally, the university is certified by the "Audit familiengerechte Hochschule" (2016) as well as the "Audit Diversity-Hochschule" (2017). With the staff unit "Equal opportunities and Antidiscrimination" the University of Bremen has established a working group, which counsels the university leadership and coordinated programs such as CRCs.

At the staff unit "Equal opportunities and Antidiscrimination" the section for "Gender Consulting" provides researchers with information and advice on their steps to reach more gender equality during the whole process of application for DFG-grants to the development of tailor-made measures to increase the number of and empower women scientist. The "Gender Consulting" section also coordinates the "Network Gender Equality in Research Collaborations". The CRC participates actively in this network and supports the "Gender Consulting" by financing the section in collaboration with other coordinated programmes at the University of Bremen. To enhance the proportion and the success of women in sciences, the University of Bremen has implemented various measures, which helped to successfully recruit and empower a number of excellent women scientists and increased the proportion of women professors in the natural sciences:

- The mentoring programme *plan m – Mentoring in Science* provides women PhD-students as well as women postdocs with support through a one to one mentoring with a successful professor of their research field as well as with workshops on career planning.
- The *navigare* programme offers workshops for women scientists in research collaborations over a period of one year. This programme is available in German or in English.
- With *perspektive promotion* the University of Bremen offers workshops, writing-groups and individual counselling for women PhD-students to support them to successfully finish their thesis.
- With the *informatica feminale* and the *women engineers summer school* the *Centre for Women in STEM* offers high quality teaching, lectures and opportunities for networking for undergraduates and graduate-students once a year.

To increase the number of women scientists, the University of Bremen implemented the project *go diverse*, which concentrates on the recruiting processes for excellent staff.

The **gender disbalance in the EASE team** could be improved substantially in many respects during the first funding period. After starting with two female PIs in the first project phase, we now have two additional female PIs (**Bettina von Helversen** and **Vanessa Didelez**) in the team, which doubles the number of female PIs in EASE Ṙecently, EASE was successful to hire a female postdoc (**Maria M. Hedblom**) for the rest of the first funding period who will bring in additional competence into the field of ontological models of complex activities. Though not foreseen, this topic has become one of the most prominent cross-research-area topics during the project's first funding period. Another example of an successful attempt to correct the gender disbalance is that number of female researchers (WiMi) at Cognitive Systems Lab increased from 0% in 2017 to 28% in 2020 (4/14 female doctorate students). However, at the career stages of PhD students and postdocs, we still need to improve our recruiting strategy a lot in order to get at targeted proportions of women in EASE. Due to the inherent gender disbalance in the field of robotics and a strong international competition with many other players both in academia and industry, we could not completely obtain the intended proportion of 20% and 30%, respectively. We will strengthen our activities to gain more attention for our field among female students and young researchers worldwide by creating innovative learning and teaching materials (such as massive open online courses) on Every Activity Science and Engineering and making them easily accessable and available for free.

	1st Funding Period				2nd Funding Period
	Targeted proportion of women (%) According to proposal for 1^{st} funding period	**Current number of men/-women Status Quo** Reference date: proposal for 2^{nd} funding period		**Current proportion of women (%) Status Quo** Reference date: proposal for 2^{nd} funding period	**Targeted proportion of women(%)**
		m	f		
Doctoral researchers	20	19	3	14	20
Postdoctoral researchers	30	6	0	0	30

We are planning to intensify a recently started collaboration with *Women in AI*[84], WAI) in the next funding period. Women in AI is a worldwide operating network with Germany being one of its largest groups (more than 400 members). Currently a WAI chapter in Bremen is to established. EASE will offer internships for female students at (or close to) master level in the EASE main lab. The interns will have the opportunity to get in touch with the female researchers of our Collaborative Research Center in Bremen and their respective international networks. Another field of collaboration between WAI and EASE will be online teaching and online learning materials. We strongly believe that the cooperation with WAI is a great opportunity to improve the recruiting of female researchers in EASE.

Gender equality activities in EASE adress many different target groups, including the doctoral students and young researchers in EASE, schools and children in Bremen and at the EASE partner locations, and the general public. In cooperation with the staff unit “Equal opportunities and Antidiscrimination” a customized programme of special courses for female researchers was developed for EASE, which included a presentation workshop for women given by the renowned coach Dr. Saskia Schottelius.

Due to the COVID-19 situation the Arbeitsstelle Chancengleichheit started to offer the online training series ”Confident in the digital workplace” for female researchers and the virtual talk series ”Gender and diversity to go” in summer 2020. Several gender-equality-related events were co-organized, hosted, and/or sponsored by EASE and EASE researchers, e.g.

- **WeLead - Women Leaders in Artificial Intelligence Engineering and Robotics** 2018 in Munich (co-organized by EASE researcher Karinne Ramirez-Amaro). WeLead goals were to showcase women key players in artificial intelligence, robotics and engineering to break the ”glass ceiling”

[84] `womeninai.co`

	1^{st} Funding period				2^{nd} Funding period		
	Proportion of women (%) According to proposal for 1st funding period	**Current number of men/-women Status Quo** Reference date: proposal for 2^{nd} funding period		**Current proportion of women (%) Status Quo** Reference date: proposal for 2^{nd} funding period	**Number of men/ women**		**Proportion of women (%)**
		m	f		m	f	
Postdoctoral researchers	0	3	0	0	4	0	0
research group leader/junior research group leaders/junior professors	0	0	0	0	1	0	0
professors C3/W2	0	2	0	0	2	1	33
professors C4/W3	18	9	2	18	9	3	25
Total	18	14	2	18	16	4	58

that prevent women to access senior positions. WeLead also aimed at increasing the presence of female researchers in robotics and AI and defining a road map of goals and events involving several gender organizations. EASE constributed as a co-organizer through Karinne Ramirez-Amaro, and as a sponsor of the event. All in all, four EASE researchers participated in this event.

- EASE constributed to a **cartoon postcard design project with an exhibition** organized by the Arbeitsstelle Chancengleichheit of the Universität Bremen ("aufgezeichnet - gender dynamics in MINT"[85])
- talk by **Dr. Sandra Buchmüller**, TU Braunschweig, on "Gender, Technik und Mobilität" sponsored by EASE
- in 2021 EASE will contribute to a **poetry slam** event on gender topics ("equality slam")

The target group of girls and young women was adressed through events like the "Girls' Day", lab visits for classes, and many workshops coorganized with the BMBF-funded project SMILE, in which two EASE PIs (Kerstin Schill and Michael Beetz) were directly involved with Kerstin Schill being the coordinator of the project. More than 40 events were organized with a total number of *1,900 female participants* from schools in the Bremen region.

The **EASE mentoring program** for female EASE researchers was implemented and EASE researcher **Gayane Kazhoyan** had a three-month research visit at the MIT, where she did research in the group of her mentor Leslie Kaelbling and had an excellent networking opportunity at one of the most prominent AI research centers worldwide. EASE researcher **Zhou (Yuen) Fang** decided to leave academia and became co-founder and CTO of the startup company ZoeysRooms. EASE researcher **Karinne Ramirez Amaro** has been appointed assistant professor, Electrical Engineering, Research Group Mechatronics Chalmers University in Gothenburg, Sweden.

[85] `uni-bremen.de/chancengleichheit/ausstellungen/aufgezeichnet-gender-dynamics-in-mint`

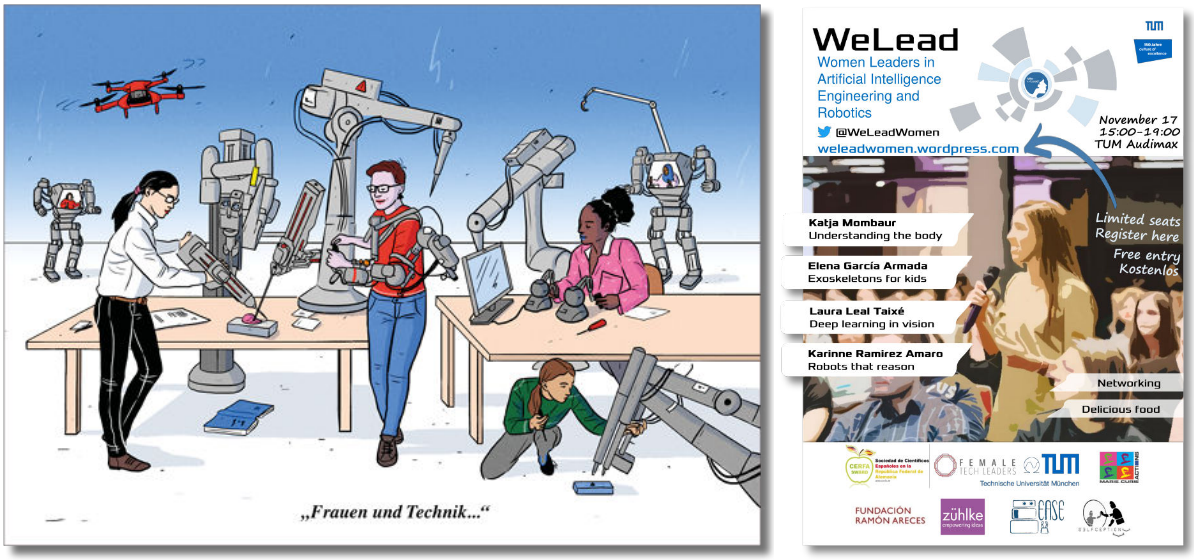


Figure 1.68: *Left:* Postcard from the gender equality cartoon project designed in collaboration with EASE. *Right:* The WeLead conference poster.

Figure 1.69: Collaborative event of the SMILE project and EASE.

In the future EASE will continue to provide its female researchers with customized measures for supporting their career within our project and beyond, be it in academia or elsewhere.

Since 2007 the University is evaluated by the "audit familiengerechte hochschule" and offers a variety of services for researchers and their families. Kindergardens are available at the university and at the Technology Park on campus. The university provides funding for off-campus childcare. Especially for international scientist the KLEX-childcare offer is important. At KLEX scientists can ask for childcare on an irregular basis such as during conferences or in peripheral hours. Furthermore, the university works together with the child care service "Die Notfallmamas", which provides professional childcare in case of emergencies, such as in case of illness or unexpected meetings. The "Welcome Centre" provides information and assistance regarding the search for childcare or schools for researchers from abroad during their stay at the University of Bremen. Dual-career couples are supported through regional networks of research institutions, through contacts with non-academic employers, and by dual-career arrangements at the University of Bremen.

1.4.3 Management of research data and knowledge

Experimental data generated using various agents and means lies at the center of EASE. Since beginning, EASE Consortium is not just committed to **produce high-quality experimental data** to put the research forward but also to **share this data publicly** to democratize robotics research. Producing the data from conducted experiments aligned with EASE Ontology is a standard practice for EASE researchers and in the research pipeline of every EASE subproject.

On the other hand, the latter would require special attention in the course of project. Open-data and data management are hot topics in research field that focus on the organisation of data, from its creation to the research cycle through to the dissemination and publishing of valuable results. INF-Subproject is dedicated to implement and setup the necessary infrastructure within EASE to adress these topics.

EASE Consortium's efforts towards implementation such a research infrastructure have been started even before the first funding period with the implementation of openEASE. As a web-based knowledge service providing robot and human experimental data called Narrative-enabled Episodic Memories, openEASE has bridged the gap between robotics research labs worldwide. Using openEASE, the researchers can access each other's data, which are standardized using EASE Ontology, and use it for their own agenda. An example of openEASE-based collaboration is illustrated in Figure 1.70. openEASE made a big impact on robotics community. Thus, openEASE-related papers (Beetz et al., 2015b; Bozcuoglu et al., 2018) have been nominated for three different best paper awards at IEEE ICRA, one of the top conferences in robotics, in 2015 and 2018 respectively.

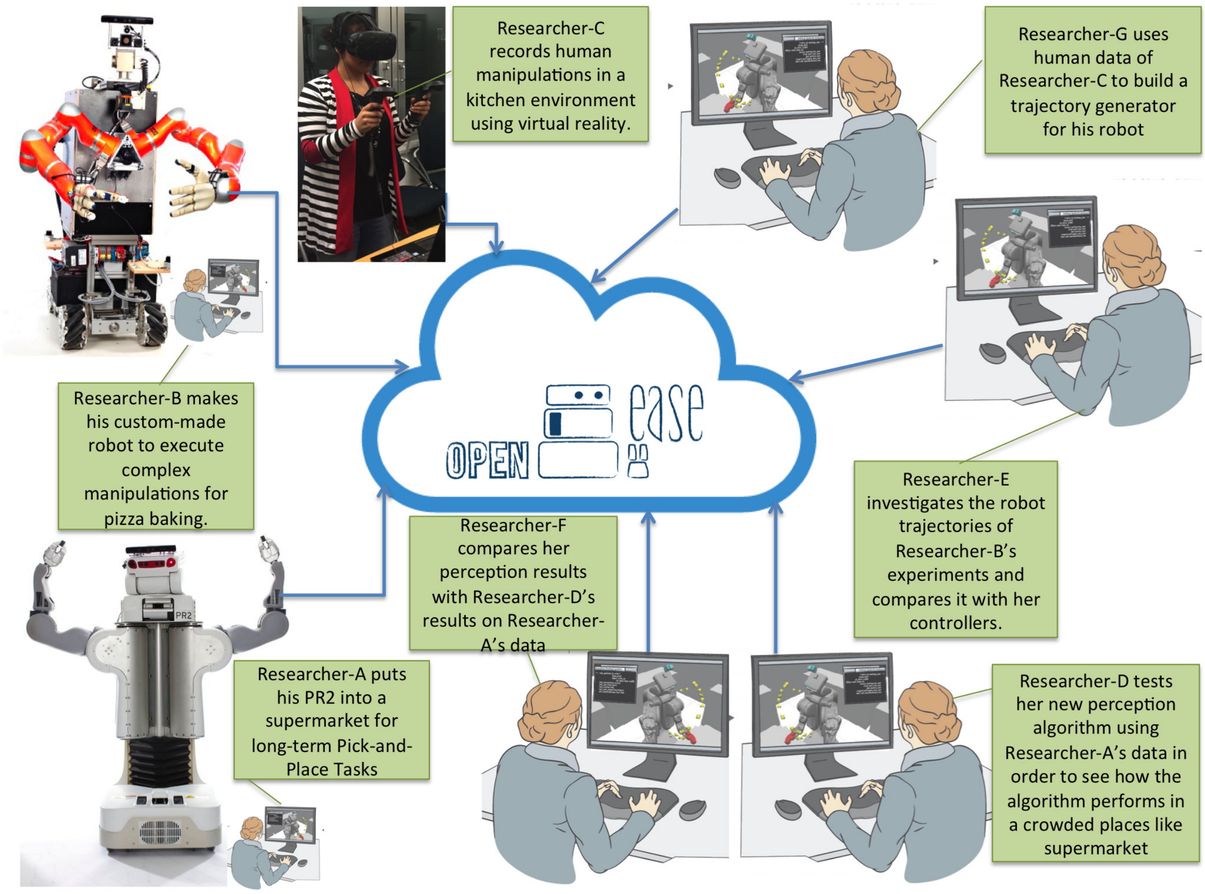


Figure 1.70: Possible Collaborations using openEASE.

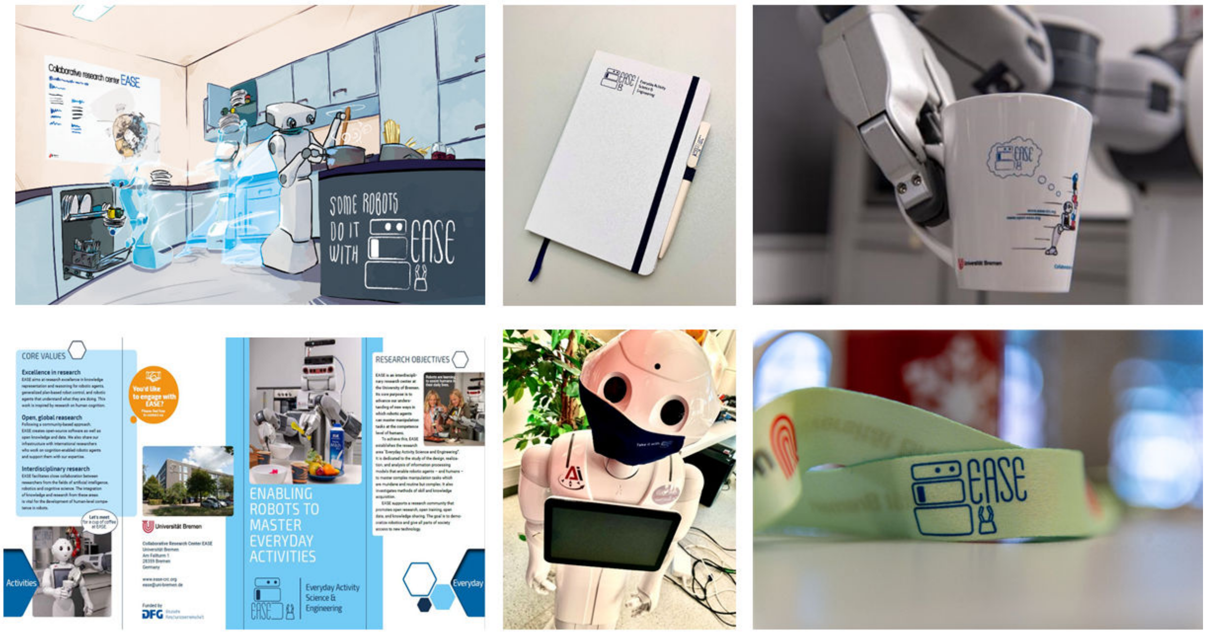


Figure 1.71: Giveaways for promoting the visibility of EASE.

During the first funding phase of EASE, INF-Project has continued the development of openEASE according to the needs of EASE. The most significant change/adoptation in this year is to setup and implement a full-fledged big data infrastructure, called NEEMHub, based on Apache HADOOP[86] and Data Version Control (DVC)[87]. At initial versions, openEASE was using a FTP server to store NEEMs. On the other hand, intense continuous developments in Knowrob Framework and EASE Ontology had required a versioning mechanism for NEEMs as well. To this end, INF-Subproject has setup a high-end HADOOP cluster. DVC, a data versioning tool by machine learning engineers, is adopted on top of HADOOP to provide data versioning using hdfs, HADOOP's remote storage. Replacing FTP server with this storage enabled openEASE to load the correct version of NEEMs with the appropriate versions of KnowRob and EASE Ontology. Moreover, it enables researchers to publish new versions of their research data without sacrificing present usage.

1.4.4 Knowledge transfer and public relations

During the first funding period EASE launched many activities for knowledge transfer and public relations. The EASE website [88] and printed information material (flyers etc.) were acompanied by appearances of EASE in the print media, TV, radio, exhibitions, physical and virtual events, movies, and hackathons.

The list below shows a selection of EASE PR activities

- Most recently EASE was featured in the "Kinderradionacht 2020", broadcasted on the 27th of November 2020.
- Many EASE researchers contributed to the online conference "Democratize AI with open research" organized by the Institute for Artificial Intelligence (IAI) at Bremen University running from September 30 until October 14, 2020. The virtual conference included: presentations by international speakers, tutorials on open projects, interview, panel discussion, networking opportunities and col-

[86] `hadoop.apache.org`
[87] `dvc.org`
[88] `ease-crc.org`

laboration and covering the advances, challenges, as well as potentials of open science in Artificial Intelligence and Robotics.

- To further increase awareness for EASE, we reached out to the student and public community. We engaged students to participating in the Bremen Big Data Challenge[89] (BBDC), a yearly event at the University of Bremen that invites students to compete for attractive prices. In 2020, the students were challenged to solve everyday activity recognition tasks by interpreting biosignal data recorded from EMG and motion capture sensors (2020). The overarching goal of the BBDC is to excite students about problem solving in Big Data, and to deepen their understanding of machine learning. While our initial intention was to raise students' interest in EASE, we indeed benefited from the challenge since student uncovered unusual and unexpected artifacts in the recordings.
- we leveraged the crowd-sourcing platform Zooniverse to perform additional spatial annotations of objects. Selected NEEMs from EASE-TSD are uploaded and can be annotated without any additional instructions or software. The resulting annotations are then post-processed, (in collaboration with subproject P03-E) supplemented with spacial relations and joined on NEEM-Hub to the main dataset.
- In March 2020 SAT1 Nord was shooting in the EASE main lab for a show broadcasted on the 4th of March.
- "Roboter als Pfleger oder Helfer: Bremer laut Umfrage noch skeptisch" (14.01.2020, Radio Bremen TV).
- exhibition "Schaufenster Wissenschaft – Highlights der Bremer Forschung" at the Haus der Wissenschaft (House of Science) from November 14, 2019, until January 27, 2020. EASE participated in the exhibition with an exhibit especially designed for this exhition which showcased our research as an interactive journey.
- The *Wissenschaft im Dialog* initiative has launched the competition *Beats & Bits* about music composed by Artificial Intelligence in 2019. EASE member Daniel Nyga contributed a blog article about different perspectives on Artifical Intelligence and creativity (Kunst oder Künstlich? Perspektiven auf die Kreativität Künstlicher Intelligenz).
- The "ScienceGoesPublic" initiative is a very popular format in which scientists present their research topics to the broad public in various bars and pubs in Bremen and Bremerhaven in a cozy and casual environment. On Thursday November 21, 2019, IAI and EASE member Daniel Nyga gave a presentation about Artificial Intelligence and robotics. Title: "... denn sie wissen, was sie tun – wie Roboter lernen" (because they know what they are doing – how robots learn).
- In the BBDC 2019 the students were challenged to solve everyday activity recognition tasks by interpreting biosignal data recorded from inertial sensors. A total of 115 students in 61 teams participated in the 2019 challenge.
- In February 2019 the Institute for Artificial Intelligence and EASE hosted the 3rd meeting of the artificial intelligence cluster BREMEN.AI. The cluster provides the opportunity for companies, startups and researchers in the field of artificial intelligence to network and to discuss current topics from practice and research.
- EASE contributed to the exhibition "Einfach Wissenswert: Robotik und KI" in the Bremen 'Haus der Wissenschaft' (February 15 – June 15, 2019). EASE member Johannes Pfau from the AG Digital Media Lab (Prof. Dr. Rainer Malaka) developed a video game (EASEY) for this exhibition. EASE member Dr. Daniel Nyga gave a presentation to the general public. Title: "... denn sie wissen, was sie tun – wie Roboter lernen" (because they know what they are doing – how robots learn)
- The Institute for Artificial Intelligence and EASE appeared in the upcoming documentary film "Hi, AI" by Isa Willinger[90]. Shooting took place in 2018. The movie won many international awards, including the 2019 Max Ophüls Prize for Best Documentary Film.
- On November 16, 2018 a TV team from RTL Nord was visiting the EASE main lab to check out the

[89] `https://bbdc.csl.uni-bremen.de/`
[90] `hiai-film.de`

robots and Virtual Reality environment.

- The episode “Können Computer denken?” (Can computers think?) was shown in the ZDF TV show ”Pur+” on the 14th of September 2018, ZDF.
- In January 2018 Radio Bremen TV show ”Buten un Binnen featured” reported about EASE in a story entitled “Die Roboterschule” (the robot school).
- The EASE main lab was visited by many school classes, guests and partners of the Universität Bremen, visiting researchers from many countries, and visitors from companies.

1.5 Other sources of third-party funding for principal investigators

Principal investigator	Project	Project title	Funding period	Funding agency
Albu-Schäffer Leidner	R04, R06	M-RUNNERS	2019 - 2024	EU
		SMiLE2gether	2019 - 2024	Bavaria
		SMiLE	2017 - 2023	Bavaria
		GINA	2018 - 2021	BMBF
		DIH-HERO	2019 - 2022	EU
		CeTI	2019 - 2025	DFG
		AnDy	2017 - 2020	EU
Beetz	R01, R03, R04, R05	AI4HSR	2021 - 2023	DFG
		CERA4HRI	2021 (applied)	BMBF
		FAME	36 months (applied)	EU
		ILIAS	2019 - 2022	BMBF
		IMPROVER	2019 - 2022	BMBF
		Knowledge 4 Retail (K4R)	2020 - 2022	BMWi
		PIPE	2018 - 2022	DFG
		REFILLS	2017 - 2022	EU
		REMARO Marie-Curie Network	2020 - 2024	EU
		RoPha	2017 - 2020	BMBF
		TraceBot	2021 - 2025	EU
		TransAIR	2019 - 2021	BMBF
		Ubica	2019 - 2021	BMWi
		SMILE	2017 - 2020	BMBF
		SFB 1232 Farbige Zustände	2016 - 2020	DFG
Cheng	R01	OpenWalker	2019 - 2020	EU
		Artificial Intelligence and the Mobility of the Future – Between Trust and Control	2020 - 2021	VW-Stiftung
		Selfception	2017 - 2019	EU
		Factory-in-a-Day	2013 - 2017	EU
Didelez	H01	Causal discovery for cohort data	2018 - 2021	DFG
Drechsler Herdt	P04	CONFIRM	2017 - 2019	BMBF
		INTUITIV	2018 - 2021	BMBF
		PolyVer	2020 - 2025	DFG
		SATiSFy	2018 - 2021	BMBF
		Scale4Edge	2020 - 2023	BMBF
		SecProPort	2018 - 2021	BMVi
		VerA	2020 - 2022	DFG
		VerSys	2019 - 2022	BMBF
		Fast and Slow	2019 - 2022	BMBF
		AdaMeKoR	2020 - 2023	BMBF
		AUTOASSERT	2020 - 2023	BMBF
		VeryHuman	2020 - 2024	BMBF
		OptiSecure	2021 - 2024	DFG
		PLiM	2019 - 2022	DFG
		SMILE	2017 - 2020	BMBF
		SYMVIR	2018 - 2021	Zentrale Forschungsförderung Uni Bremen
		Formalisierung und Eigenschaften von Plänen	2017 - 2021	DFG
		Formale Methoden zur Energie-sicheren Testerzeugung für digitale Schaltunge	2016 - 2019	DFG

Principal investigator	Project	Project title	Funding period	Funding agency
		Entwicklung einer qualitätsorientierten, kosteffektiven und robusten Testumgebung für Nanometer-Schaltkreise	2013 - 2017	Exzellenz
		Hochflexible Materialsynthese und Mikrostruktureinstellung	2020 - 2024	DFG
		KI-SIGS	2020 - 2023	BMWi
		QuadCore SoC	2016 - 2021	Industrie
		ATB: General Technology Assessment	2018 - 2019	Industrie
		ai-Philos	2018 - 2019	Industrie
		CONVERS	2017 - 2020	BMBF
		SecRec	2017 - 2020	BMBF
		SELFIE	2016 - 2019	BMBF
		MANIAC	2016 - 2019	DFG
		SaferApps	2015 - 2017	BMWi
		Neuartige rekonfigurierbare Transistoren für den Knowhow-Schutz von Elektronikkomponenten	2020 - 2023 (applied)	BMBF
		VE-HEP	2020 - 2023 (applied)	BMBF
		SFB 1232 Farbige Zustände	2016 - 2020	DFG
Frese	R02	ZaVI - Zustandsschätzung allein durch Vorwissen und Inertialsensorik	2018 - 2022	DFG - FR2620/3-1
		DoF-Adaptiv	2021 - 2024	BMBF - V5KI039
Herrmann	H04	KD2School – Designing Adaptive Systems for Economic Decisions	2021 - 2025 (final review due 2021)	DFG
Lutz	P05	Conservative Extensions Beyond Description Logics (CEO)	2015 - 2018	DFG
		Custom-Made Ontology-Based Data Access (CODA)	2015 - 2020	ERC
		Query Evaluation in Open and Closed Worlds: Testing, Enumeration and Counting (QTEC)	2020 - 2023	DFG
Malaka	P01, P05	Adaptify	2015 - 2018	BMBF
		Empowering Digital Media	2017 - 2022	Klaus-Tschira Stiftung
		first.stage	2016 - 2019	EU
		MAL	2016 - 2019	BMBF
		VIVATOP	2018 - 2021	BMBF
		UsableSecAtHome	2020 - 2023	BMBF
		Muhai	2020 - 2024	EU
		KI-SIGS	2020 - 2023	BMWi
		SmartOT	2019 - 2021	BMBF
		Aimdata	2016 - 2019	DFG
		Pervasive Health	2021 - 2024	DFG
		Exist – Gründerstipendium Waldfleisch	2020 - 2021	BMBF
		SecProPort	2018 - 2021	BMVI
Ritter	R05	SARAFUN	2015 - 2018	EU
		NeuTouch	2019 - 2023	EU
		TACT-HAND	2016 - 2019	DFG
		DEEP-HAND	2020 - 2022	DFG
Schill	H01, H03	EnEx-CAUSE	2015 - 2018	BMWi
		SMARTFARM	2016 - 2019	BMWi
		UAgriCo	2016 - 2020	BMBF
		AO-CAR	2016 - 2018	BMWi
		SMILE	2017 - 2020	BMBF
		BOB - Bee Observer	2018 - 2020	BMBF
		BORUS - Bee Observer Russia	2019 - 2021	BMBF
		[at]CITY-AF	2019 - 2021	Continental
		PRORETRA 5-urbAn driving	2019 - 2022	Contintental
		KANARIA-K2I	2019 - 2020	BMWi
		OPA3 L	2019 - 2023	BMWi
		Bee Var	2020 - 2021	BAB
		TRIPLE-nanuAUV1	2020 - 2022	BMWi
		VIPE	2017 - 2022	BMBF/NSF
Schultz	H03, H04	Privacy-preserving Natural Language Processing	2021 - 2024	Sparkasse Bremen
		SmartHelm	2020 - 2023	BMVI
		Google Faculty Research Award	2020 - 2021	Google Inc.

Principal investigator	Project	Project title	Funding period	Funding agency
		DEEP-HAND	2020 - 2022	DFG
		ALMED	2019 - 2022	DFG
		Graduiertenkolleg	2018 -2021	Klaus-Tschira Stiftung
		ASARob	2018 - 2020	BMBF
		Response (USA - Germany)	2017 - 2020	NSF-BMBF
		Arthrokinemat	2016 - 2019	BMWi
		I-CARE	2016 - 2019	BMBF
von Helversen	H04	Understanding the Role of Memory in Judgments and Decisions: The Influence of Exemplar	2016 - 2020	Swiss national fund
		Modeling Human Judgment: Integrating Memory and Rule-based Processes	2013 - 2018	Swiss national fund
Zachmann	R03	Dynamischer Hüftimplantatssimulator (Dynamic HIPS)	2020 - 2022	BMBF - 16SV8355
		Kognitionsbasierte, autonome Navigation am Beispiel des Ressourcenabbaus im All (Kanaria)	2013 - 2017	DLR - 50 NA 1318
		Kognitive Autonome Navigation am Beispiel des Ressourcenabbaus im All, Teilvorhaben Raumfahrzeuge und Rover (Kanaria2-K2I-RR)	2019 - 2022	DLR - 50NA1916
		Steigerung von Ergonomie und Effizienz im OP durch Smarte Beleuchtung und Smarte Steuerung (SmartOT)	2019 - 2021	BMBF - 13GW0264D
		Optimal assistierte, hoch automatisierte, autonome und kooperative Fahrzeugnavigation und Lokalisation (OPA3L)	2019 - 2023	DLR - 50NA1909
		Vielseitiger, immersiver, virtueller, und augmentierter Tangible OP (Vivatop)	2018 - 2021	BMBF - 16SV8077
		Visual Autonomous Robotics (VAR)	2018 - 2019	BMBF - 01DS18006
		Hüftimplantat-Pfannenfrässimulator (HIPS)	2016 - 2018	BMWi - 16KN036252
		Modulares virtuelles Test-bed für die VaMex-Vorhaben (VaMex-VTB)	2017 - 2019	DLR - 50NA1712
		Intra-Operative Information	2013 - 2017	Universität + DFG

References

Rachid Alami, Raja Chatila, Sara Fleury, Matthieu Herrb Felix Ingrand, Maher Khatib, Benoit Morisset, Philippe Moutarlier, and Thierry Simeon. Around the Lab in 40 days. . . . In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2000)*, 2010.

John Eric Anderson. *Constraint-Directed Improvisation for Everyday Activities*. PhD thesis, Department of Computer Science, University of Manitoba. Winnipeg-Manitoba, Canada., 1995.

Tamim Asfour, Fabian Paus, Mirko Wächter, Lukas Kaul, Samuel Rader, Pascal Weiner, Simon Ottenhaus, Raphael Grimm, You Zhou, and Markus Grotz. ARMAR-6: A high-performance humanoid for human-robot collaboration in real-world scenarios. *IEEE Robotics Autom. Mag.*, 26(4):108–121, 2019.

Naomi Azrieli. BRAIN Canada. braininitiative.nih.gov, 2011.

Ferenc Balint-Benczedi. *Task adaptabled, Pervasive Perpception for Robots Performing Everyday Manipulation*. PhD thesis, University of Bremen, 2020.

Ferenc Bálint-Benczédi and Michael Beetz. Variations on a Theme:“It's a Poor Sort of Memory that Only Works Backwards”. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 8390–8396. IEEE, 2018.

Ferenc Balint-Benczedi, Patrick Mania, and Michael Beetz. Scaling Perception Towards Autonomous Object Manipulation — In Knowledge Lies the Power. In *International Conference on Robotics and Automation (ICRA)*, Stockholm, Sweden, 2016.

Ferenc Balint-Benczedi, Zoltan-Csaba Marton, Maximilian Durner, and Michael Beetz. Storing and Retrieving Perceptual Episodic Memories for Long-term Manipulation Tasks. In *Proceedings of the 2017 IEEE International Conference on Advanced Robotics (ICAR)*, pages 25–31, Hong-Kong, China, 2017. IEEE. **Finalist for Best Paper Award**.

Georg Bartels, Daniel Beßler, and Michael Beetz. Episodic Memories for Safety-Aware Robots. *KI-Künstliche Intelligenz*, 33 (2):123–130, 2019.

Chiara Bartolozzi, Daniele Pucci, Agnieszka Wykowska, and Giorgio Metta. iCub: The not-yet-finished story of building a robot child. *Sci. Robotics*, 2(13), 2017.

John Bateman, Michael Beetz, Daniel Beßler, Asil Kaan Bozcuoğlu, and Mihai Pomarlan. Heterogeneous Ontologies and Hybrid Reasoning for Service Robotics: The EASE Framework. In Anibal Ollero, Alberto Sanfeliu, Luis Montano, Nuno Lau, and Carlos Cardeira, editors, *ROBOT 2017: Third Iberian Robotics Conference*, pages 417–428, Cham, 2018a. Springer International Publishing.

John Bateman, Mihai Pomarlan, and Gayane Kazhoyan. Embodied contextualization: Towards a multistratal ontological treatment. *Journal of Applied Ontology*, 14(1):1–35, 2019.

Tamas Bates, Karinne Ramirez-Amaro, Tetsunari Inamura, and Gordon Cheng. On-Line Simultaneous Learning and Recognition of Everyday Activities from Virtual Reality Performances. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Vancouver, Canada, 2017.

Peter W. Battaglia, Jessica B. Hamrick, and Joshua B. Tenenbaum. Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, 110(45):18327–18332, 2013.

Adrian S Bauer, Peter Schmaus, Freek Stulp, and Daniel Leidner. Probabilistic Effect Prediction through Semantic Augmentation and Physical Simulation. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9278–9284. IEEE, 2020.

Michael Beetz. *Concurrent Reactive Plans: Anticipating and Forestalling Execution Failures*, volume LNAI 1772 of *Lecture Notes in Artificial Intelligence*. Springer Publishers, Berlin, 2000.

Michael Beetz, Lorenz Mösenlechner, and Moritz Tenorth. CRAM – A Cognitive Robot Abstract Machine for Everyday Manipulation in Human Environments. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1012–1017, Taipei, Taiwan, October 18-22 2010b.

Michael Beetz, Dominik Jain, Lorenz Mösenlechner, Moritz Tenorth, Lars Kunze, Nico Blodow, and Dejan Pangercic. Cognition-Enabled Autonomous Robot Control for the Realization of Home Chore Task Intelligence. *Proceedings of the IEEE*, 100(8):2454–2471, 2012.

Michael Beetz, Ferenc Balint-Benczedi, Nico Blodow, Daniel Nyga, Thiemo Wiedemeyer, and Zoltan-Csaba Marton. RoboSherlock: Unstructured Information Processing for Robot Perception. In *IEEE International Conference on Robotics and Automation (ICRA)*, Seattle, Washington, USA, 2015a. **Best Service Robotics Paper Award**.

Michael Beetz, Moritz Tenorth, and Jan Winkler. Open-EASE – A Knowledge Processing Service for Robots and Robotics/AI Researchers. In *IEEE International Conference on Robotics and Automation (ICRA)*, Seattle, Washington, USA, 2015b. **Finalist for the Best Cognitive Robotics Paper Award.**

Michael Beetz, Daniel Beßler, Andrei Haidu, Mihai Pomarlan, Asil Kaan Bozcuoğlu, and Georg Bartels. Knowrob 2.0—a 2nd generation knowledge processing framework for cognition-enabled robotic agents. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 512–519. IEEE, 2018. **Is on the list of most important publications**

over past 3 years in the area of cognitive robotics selected by the IEEE RAS Technical Committee on Cognitive Robotics.

Thorsten Belker and Michael Beetz. Learning to Execute Robot Navigation Plans. In *Proceedings of the 25th German Conference on Artificial Intelligence (KI 01)*, Wien, Austria, 2001.

Benjamin Bergen and Nancy Chang. Embodied Construction Grammar in simulation-based language understanding. *Construction grammars: Cognitive grounding and theoretical extensions*, 07 2003.

Christopher Berner, Greg Brockman, Brooke Chan, Vicki Cheung, Przemyslaw Debiak, Christy Dennison, David Farhi, Quirin Fischer, Shariq Hashme, Chris Hesse, Rafal Józefowicz, Scott Gray, Catherine Olsson, Jakub W. Pachocki, Michael Petrov, Henrique Pond'e de Oliveira Pinto, Jonathan Raiman, Tim Salimans, Jeremy Schlatter, Jonas Schneider, Szymon Sidor, Ilya Sutskever, Jie Tang, Filip Wolski, and Susan Zhang. Dota 2 with Large Scale Deep Reinforcement Learning. *ArXiv*, abs/1912.06680, 2019.

Daniel Beßler, Mihai Pomarlan, Aliakbar Akbari, Mohammed Diab, Jan Rosell, John Bateman, Michael Beetz, et al. Assembly planning in cluttered environments through heterogeneous reasoning. In *Joint German/Austrian Conference on Artificial Intelligence (Künstliche Intelligenz)*, pages 201–214. Springer, 2018b.

Daniel Beßler, Mihai Pomarlan, and Michael Beetz. OWL-enabled Assembly Planning for Robotic Agents. In *Proceedings of the 2018 International Conference on Autonomous Agents*, AAMAS '18, Stockholm, Sweden, 2018c. **Finalist for the Best Robotics Paper Award.**

Daniel Beßler, Robert Porzel, Pomarlan Mihai, and Michael Beetz. Foundational Models for Manipulation Activity Parsing. In Timothy Jung, M. Claudia tom Dieck, and Philipp A. Rauschnabel, editors, *Augmented Reality and Virtual Reality – Changing Realities in a Dynamic World*. Springer, 2019b.

Daniel Beßler, Robert Porzel, Pomarlan Mihai, Michael Beetz, Rainer Malaka, and John Bateman. A Formal Model of Affordances for Flexible Robotic Task Execution. In *Proc. of the 24th European Conference on Artificial Intelligence (ECAI)*, 2020b.

Mehul Bhatt. Reasoning about Space, Actions and Change: A Paradigm for Applications of Spatial Reasoning. In *Qualitative Spatial Representation and Reasoning: Trends and Future Directions*. IGI Global, USA, 2012.

Mehul Bhatt and Seng W. Loke. Modelling Dynamic Spatial Systems in the Situation Calculus. *Spatial Cognition & Computation*, 8(1-2):86–130, 2008.

Mehul Bhatt and Jakob Suchan. Cognitive Vision and Perception. In Giuseppe De Giacomo, Alejandro Catalá, Bistra Dilkina, Michela Milano, Senén Barro, Alberto Bugarín, and Jérôme Lang, editors, *ECAI 2020 - 24th European Conference on Artificial Intelligence, 29 August-8 September 2020, Santiago de Compostela, Spain, August 29 - September 8, 2020 - Including 10th Conference on Prestigious Applications of Artificial Intelligence (PAIS 2020)*, volume 325 of *Frontiers in Artificial Intelligence and Applications*, pages 2881–2882. IOS Press, 2020b.

Mehul Bhatt, Jae Hee Lee, and Carl P. L. Schultz. CLP(QS): A declarative spatial reasoning framework. In Max J. Egenhofer, Nicholas A. Giudice, Reinhard Moratz, and Michael F. Worboys, editors, *Spatial Information Theory - 10th International Conference, COSIT 2011, Belfast, ME, USA, September 12-16, 2011. Proceedings*, volume 6899 of *Lecture Notes in Computer Science*, pages 210–230. Springer, 2011.

Mehul Bhatt, Carl P. L. Schultz, and Christian Freksa. The 'Space' in Spatial Assistance Systems : Conception, Formalisation and Computation. In Thora Tenbrink, Jan Wiener, and Christophe Claramunt, editors, *Representing Space in Cognition: Interrelations of behaviour, language, and formal models*. Oxford University Press, 2013a.

Mehul Bhatt, Jakob Suchan, and Carl P. L. Schultz. Cognitive interpretation of everyday activities - toward perceptual narrative based visuo-spatial scene interpretation. In Mark A. Finlayson, Bernhard Fisseni, Benedikt Löwe, and Jan Christoph Meister, editors, *2013 Workshop on Computational Models of Narrative, CMN 2013, August 4-6, 2013, Hamburg, Germany*, volume 32 of *OASICS*, pages 24–29. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2013b.

Erik A. Billing, Henrik Svensson, Robert Lowe, and Tom Ziemke. Finding Your Way from the Bed to the Kitchen: Reenacting and Recombining Sensorimotor Episodes Learned from Human Demonstration. *Frontiers in Robotics and AI*, 3:9, 2016.

Peter Blouw, Eugene Solodkin, Paul Thagard, and Chris Eliasmith. Concepts as Semantic Pointers: A Framework and Computational Model. *Cogn. Sci.*, 40(5):1128–1162, 2016.

R. Peter Bonasso, R. James Firby, Erann Gat, David Kortenkamp, David P. Miller, and Marc G. Slack. Experiences with an architecture for intelligent, reactive agents. *J. Exp. Theor. Artif. Intell.*, 9(2-3):237–256, 1997.

Christoph Borst, Thomas Wimböck, Florian Schmidt, Matthias Fuchs, Bernhard Brunner, Franziska Zacharias, Paolo Robuffo Giordano, Rainer Konietschke, Wolfgang Sepp, Stefan Fuchs, Christian Rink, Alin Albu-Schäffer, and Gerd Hirzinger. Rollin' Justin - Mobile platform with variable base. In *2009 IEEE International Conference on Robotics and Automation, ICRA 2009, Kobe, Japan, May 12-17, 2009*, pages 1597–1598. IEEE, 2009.

Anneke Bötcher, Carsten Lutz, and Frank Wolter. Complete Approximations of Horn DL Ontologies. In Magdalena Ortiz and Thomas Schneider, editors, *Proceedings of the 31st International Workshop on Description Logics co-located with 16th International Conference on Principles of Knowledge Representation and Reasoning (KR 2018), Tempe, Arizona, US, October 27th - to - 29th, 2018*, volume 2211 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2018.

Anneke Bötcher, Carsten Lutz, and Frank Wolter. Ontology Approximation in Horn Description Logics. In Sarit Kraus, editor, *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*, pages 1574–1580. AAAI Press, ijcai.org, 2019.

Asil Kaan Bozcuoglu, Gayane Kazhoyan, Yuki Furuta, Simon Stelter, Michael Beetz, Kei Okada, and Masayuki Inaba. The Exchange of Knowledge using Cloud Robotics. *Robotics and Automation Letters*, 3(2):1072–1079, April 2018. **Finalist for Best Cognitive Robotics Paper**.

Ronald J. Brachman. Systems That Know What They're Doing. *IEEE Intell. Syst.*, 17(6):67–71, 2002.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. Language Models are Few-Shot Learners, 2020.

Sebastian G Brunner, Peter Lehner, Martin J Schuster, Sebastian Riedel, Rico Belder, Daniel Leidner, Armin Wedler, Michael Beetz, and Freek Stulp. Design, execution, and postmortem analysis of prolonged autonomous robot operations. *IEEE Robotics and Automation Letters*, 3(2):1056–1063, 2018.

Gordon Cheng, Karinne Ramirez-Amaro, Michael Beetz, and Yasuo Kuniyoshi. Purposive learning: Robot reasoning about the meanings of human activities. *Science Robotics*, 4(26), 2019.

Agnese Chiatti, Enrico Motta, and Enrico Daga. Towards a Framework for Visual Intelligence in Service Robotics: Epistemic Requirements and Gap Analysis, 2020.

Mohammed Diab, Mihai Pomarlan, Daniel Beßler, Aliakbar Akbari, Jan Rosell, John Bateman, and Michael Beetz. An Ontology for Failure Interpretation in Automated Planning and Execution. In *Fourth Iberian Robotics conference*, ROBOT '19, pages 381–390. Springer, 2019.

Mohammed Diab, Mihai Pomarlan, Daniel Beßler, Aliakbar Akbari, Jan Rosell, John Bateman, and Michael Beetz. SkillMaN–A skill-based robotic manipulation framework based on perception and reasoning. *Robotics and Autonomous Systems*, 134:103653, 2020a.

Alexander Dietrich, Kristin Bussmann, Florian Petit, Paul Kotyczka, Christian Ott, Boris Lohmann, and Alin Albu-Schäffer. Whole-body impedance control of wheeled mobile manipulators - Stability analysis and experiments on the humanoid robot Rollin' Justin. *Auton. Robots*, 40(3):505–517, 2016.

Pedro Domingos, Stanley Kok, Hoifung Poon, Matthew Richardson, and Parag Singla. Unifying Logical and Statistical AI. In *Proceedings of the Twenty-First National Conference on Artificial Intelligence*, Boston, MA, 2006. AAAI Press.

Chris Eliasmith, Terrence C. Stewart, Xuan Choo, Trevor Bekolay, Travis DeWolf, Yichuan Tang, and Daniel Rasmussen. A Large-Scale Model of the Functioning Brain. *Science*, 338(6111):1202–1205, 2012.

Mustafa Ersen, Erhan Oztop, and Sanem Sariel. Cognition-Enabled Robot Manipulation in Human Environments: Requirements, Recent Work, and Open Problems. *IEEE Robotics Automation Magazine*, 24(3):108–122, Sept 2017.

David Ferrucci, Eric Brown, Jennifer Chu-Carroll, James Fan, David Gondek, Aditya A. Kalyanpur, Adam Lally, J. William Murdock, Eric Nyberg, John Prager, Nico Schlaefer, and Chris Welty. Building Watson: An Overview of the DeepQA Project. *AI Magazine*, 31(3):59–79, Jul. 2010.

R. James Firby, Peter N. Prokopowicz, Michael J. Swain, Roger E. Kahn, and David Franklin. Programming CHIP for the IJCAI-95 Robot Competition. *AI Magazine*, 17(1):71–81, 1996.

J Randall Flanagan, Miles C Bowman, and Roland S Johansson. Control strategies in object manipulation tasks. *Current opinion in neurobiology*, 16(6):650–659, 2006.

Maria Fox and Derek Long. PDDL+ level 5: An extension to PDDL2.1 for modeling planning domain with continuous time-dependent effects, 2001.

Christian Freksa. Strong Spatial Cognition. In Sara Irina Fabrikant, Martin Raubal, Michela Bertolotto, Clare Davies, Scott Freundschuh, and Scott Bell, editors, *Spatial Information Theory*, Lecture Notes in Computer Science, pages 65–86. Springer International Publishing, 2015.

Christian Freksa, Thomas Barkowsky, Zoe Falomir, and Jasper van de Ven. Geometric problem solving with strings and pins. *Spatial Cognition & Computation*, 19(1):46–68, 2019.

Lex Fridman. MIT 6.S099 course: Artificial General Intelligence. `agi.mit.edu`, 2020.

Konrad Gadzicki, Razieh Khamsehashari, and Christoph Zetzsche. Early vs Late Fusion in Multimodal Convolutional Neural Networks. In *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*, pages 1–6, 2020.

Malik Ghallab, Adele Howe, Dave Christianson, Drew McDermott, Ashwin Ram, Manuela Veloso, Daniel Weld, and David Wilkins. PDDL - The Planning Domain Definition Language. *AIPS98 Planning Committee*, 78(4):1–27, 1998.

Malik Ghallab, Dana S. Nau, and Paolo Traverso. Automated Planning: Theory and Practice. In *The Morgan Kaufmann Series in Artificial Intelligence*. Morgan Kaufmann, 2004.

Yolanda Gil and Bart Selman. A 20-Year Community Roadmap for Artificial Intelligence Research in the US. Technical report, AAAI and CCC, 2019.

Giovannetti, Tania, Schwartz, Myrna, Buxbaum, and Laurel. The Coffee Challenge: A new method for the study of everyday action errors. *Journal of Clinical and Experimental Neuropsychology*, 29(7):690–705, oct 2007.

Tania Giovannetti, David J Libon, and Tessa Hart. Awareness of naturalistic action errors in dementia. *Journal of the International Neuropsychological Society JINS*, 8(5):633–644, 2002.

Anneke Haga, Carsten Lutz, Johannes Marti, and Frank Wolter. A Journey into Ontology Approximation: From Non-Horn to Horn. In Christian Bessiere, editor, *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020*, pages 1827–1833. ijcai.org, 2020.

Andrei Haidu and Michael Beetz. Automated models of human everyday activity based on game and virtual reality technology. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 2606–2612. IEEE, 2019.

Irene Hartigan. A comparative review of the Katz ADL and the Barthel Index in assessing the activities of daily living of older people. *International Journal of Older People Nursing*, 2(3):204–212, 2007.

Demis Hassabis, Dharshan Kumaran, Christopher Summerfield, and Matthew Botvinick. Neuroscience-Inspired Artificial Intelligence. *Neuron*, 95(2):245–258, jul 2017.

Germund Hesslow. Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences*, 6(6): 242–247, 2002.

Thomas Hinterecker, Paolo Pretto, Ksander N. de Winkel, Hans-Otto Karnath, Heinrich H. Bülthoff, and Tobias Meilinger. Body-relative horizontal–vertical anisotropy in human representations of traveled distances. *Experimental Brain Research*, 236(10):2811–2827, 2018.

C. L. Hull. *Principles of behavior: an introduction to behavior theory*. Appleton-Century, Oxford, England, 1943.

Félix Ingrand and Malik Ghallab. Deliberation for autonomous robots: A survey. *Artif. Intell.*, 247:10–44, 2017.

Kathryn J. Jeffery, Aleksandar Jovalekic, Madeleine Verriotis, and Robin Hayman. Navigating in a three-dimensional world. *Behavioral and Brain Sciences*, 36(5):523–543, 2013.

Mark Johnson. The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason. volume 5, pages 58–60. University of Illinois Press, 1987a.

Matt Jones and Bradley C. Love. Bayesian Fundamentalism or Enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behavioral and Brain Sciences*, 34(4):169–188, 2011.

Jean Christoph Jung, Carsten Lutz, Hadrien Pulcini, and Frank Wolter. Logical Separability of Incomplete Data under Ontologies. In *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning, KR 2020, Rhodes, Greece, September 12-18, 2020*. AAAI Press, 2020b.

Daniel Kahneman. *Thinking, fast and slow*. Farrar, Straus and Giroux, New York, 2011.

Takeo Kanade. Quality of Life Technology. *Proceedings of the IEEE*, 100(8):2394–2396, 2012.

Gayane Kazhoyan and Michael Beetz. Programming Robotic Agents with Action Descriptions. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 103–108, Vancouver, Canada, Sept 2017.

Gayane Kazhoyan and Michael Beetz. Accomplishing Variations of Mobile Manipulation Actions through Generalized Robot Plans. In *Proceedings of the Seventh Annual Conference on Advances in Cognitive Systems*, 2019a.

Gayane Kazhoyan and Michael Beetz. Executing Underspecified Actions in Real World Based on Online Projection. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5156–5163, 2019b.

Gayane Kazhoyan, Alina Hawkin, Sebastian Koralewski, Andrei Haidu, and Michael Beetz. Learning Motion Parameterizations of Mobile Pick and Place Actions from Observing Humans in Virtual Environments. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020a.

Gayane Kazhoyan, Arthur Niedzwiecki, and Michael Beetz. Towards Plan Transformations for Real-World Mobile Fetch and Place. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2020b.

Charles C Kemp, Aaron Edsinger, and Eduardo Torres-Jara. Challenges for robot manipulation in human environments. *IEEE Robotics and Automation Magazine*, 14(1):20, 2007.

Franklin Kenghagho Kenfack, Feroz Ahmed Siddiky, Ferenc Balint-Benczedi, and Michael Beetz. RobotVQA — A Scene-Graph- and Deep-Learning-based Visual Question Answering System for Robot Manipulation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Las Vegas, USA, 2020.

Frank Kirchner, Sirko Straube, Daniel Kühn, and Nina Hoyer, editors. *AI Technology for Underwater Robots*. Intelligent Systems, Control and Automation: Science and Engineering. Springer International Publishing, 2020.

Wouter Kool, Joseph T. McGuire, Zev B. Rosen, and Matthew M. Botvinick. Decision Making and the Avoidance of Cognitive Demand. *Journal of experimental psychology. General*, 139(4):665–682, 2010.

Sebastian Koralewski, Gayane Kazhoyan, and Michael Beetz. Self-Specialization of General Robot Plans Based on Experience. *IEEE Robotics and Automation Letters*, 4(4):3766–3773, 2019.

Stephen M. Kosslyn. Mental Images and the Brain. *Cognitive Neuropsychology*, 22:333–347, 2005.

Robert A. Kowalski. Algorithm = logic + control. *Commun. ACM*, 22:424–436, 1979.

Bernd Krieg-Brückner, Serge Autexier, Martin Rink, and Sidoine Ghomsi Nokam. Formal Modelling for Cooking Assistance. In Rocco De Nicola and Rolf Hennicker, editors, *Software, Services, and Systems - Essays Dedicated to Martin Wirsing on the Occasion of His Retirement from the Chair of Programming and Software Engineering*, volume 8950 of *Lecture Notes in Computer Science*, pages 355–376. Springer, 2015.

Benjamin Kuipers, Edward A. Feigenbaum, Peter E. Hart, and Nils J. Nilsson. Shakey: From Conception to History. *AI Magazine*, 38(1):88–103, 2017a.

Michaela Kümpel, Anna de Groot, Ilaria Tiddi, and Michael Beetz. Using Linked Data to Help Robots Understand Product-related Actions. In *JOWO 2020, The Joint Ontology Workshops*, volume 2708. CEUR-WS, 2020.

Lars Kunze and Michael Beetz. Envisioning the qualitative effects of robot manipulation actions using simulation-based projections. *Artif. Intell.*, 247:352–380, 2017.

Brenden M. Lake, Tomer D. Ullman, Joshua B. Tenenbaum, and Samuel Gershman. Building Machines That Learn and Think Like People. *The Behavioral and brain sciences*, 40:e253, 2016.

George Lakoff. The contemporary theory of metaphor. In *Metaphor and Thought*, page 202–251. Cambridge University Press, 1993.

Michael F. Land and Mary Hayhoe. In what ways do eye movements contribute to everyday activities? *Vision Research*, 41 (25):3559–3565, 2001.

M. Powell Lawton and Elaine M. Brody. Assessment of older people: self-maintaining and instrumental activities of daily living. *The Gerontologist*, 9(3):179–186, 1969.

Daniel Leidner, Georg Bartels, Wissam Bejjani, Alin Albu-Schäffer, and Michael Beetz. Cognition-enabled robotic wiping: Representation, planning, execution, and interpretation. *Robotics and Autonomous Systems*, 114:199–216, 2019.

Daniel Sebastian Leidner. *Cognitive reasoning for compliant robot manipulation*, volume 127. Springer, 2019b. **Dissertation Thesis won the Georges Giralt PhD Award**.

Hector Levesque and Gerhard Lakemeyer. Cognitive robotics. In *Chapter 24 in Handbook of Knowledge Representation*. Elsevier, 2007.

Gérard Ligozat. John Wiley & Sons, Ltd, 2013.

Jaime Leonardo Maldonado Cañon, Thorsten Kluss, and Christoph Zetzsche. Adaptivity of End Effector Motor Control Under Different Sensory Conditions: Experiments With Humans in Virtual Reality and Robotic Applications. *Frontiers in Robotics and AI*, 6:63, 2019.

Jaime Leonardo Maldonado Cañon, Thorsten Kluss, and Christoph Zetzsche. Categorization of Contact Events as Intended or unintended using Pre-Contact Kinematic Features. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 46–51, 2020b.

Patrick Mania and Michael Beetz. A Framework for Self-Training Perceptual Agents in Simulated Photorealistic Environments. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 4396–4402. IEEE, 2019.

Patrick Mania, Franklin Kenghagho Kenfack, Michael Neumann, and Michael Beetz. Imagination-enabled Robot Perception. `https://arxiv.org/abs/2011.11397`, 2020.

Gary Marcus and Ernest Davis. *Rebooting AI. Building Artificial Intelligence We Can Trust*. penguin Random House, 2019a.

Eitan Marder-Eppstein, Eric Berger, Tully Foote, Brian Gerkey, and Kurt Konolige. The office marathon: Robust navigation in an indoor office environment. In *International Conference on Robotics and Automation*, 2010.

Henry Markram. Biology - The blue brain project. In *Proceedings of the ACM/IEEE SC2006 Conference on High Performance Networking and Computing, November 11-17, 2006, Tampa, FL, USA*, page 53. ACM Press, 2006.

Henry Markram. www.humanbrainproject.eu, 2013.

Claudio Masolo, Stefano Borgo, Aldo Gangemi, Nicola Guarino, and Alessandro Oltramari. WonderWeb Deliverable D18 Ontology Library (final). Technical report, IST Project 2001-33052 WonderWeb: Ontology Infrastructure for the Semantic Web, 2003.

Celeste Mason, Konrad Gadzicki, Moritz Meier, Florian Ahrens, Thorsten Kluss, Jaime Maldonado, Felix Putze, Thorsten Fehr, Christoph Zetzsche, Manfred Herrmann, Kerstin Schill, and Tanja Schultz. From Human to Robot Everyday Activity. In *IROS 2020*, Las Vegas, USA, 2020.

D. McDermott. Robot Planning. *AI Magazine*, 13(2):55–79, 1992.

Colin McGinn. *Prehension: The Hand and the Emergence of Humanity*. MIT Press, 2015.

Martin Meier, Robert Haschke, and Helge J. Ritter. From Geometries to Contact Graphs. In *Artificial Neural Networks and Machine Learning - ICANN 2020 - 29th International Conference on Artificial Neural Networks, Bratislava, Slovakia, September 15-18, 2020, Proceedings, Part II*, volume 12397, pages 546–555. Springer, 2020. **Best Paper Award**.

Andrew N. Meltzoff. Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, 31:838—850, 1995.

Tim Meywerk, Marcel Walter, Vladimir Herdt, Daniel Große, and Rolf Drechsler. Towards Formal Verification of Plans for Cognition-Enabled Autonomous Robotic Agents. *22nd Euromicro Conference on Digital System Design (DSD)*, pages 129–136, 2019.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin A. Riedmiller, Andreas Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. Human-level control through deep reinforcement learning. *Nat.*, 518(7540):529–533, 2015.

Lorenz Mösenlechner, Nikolaus Demmel, and Michael Beetz. Becoming action-aware through reasoning about logged plan execution traces. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2231–2236, 2010.

Armin Müller, Alexandra Kirsch, and Michael Beetz. Transformational Planning for Everyday Activity. In *Proceedings of the 17th International Conference on Automated Planning and Scheduling (ICAPS'07)*, pages 248–255, Providence, USA, September 2007.

Masaki Murooka, Ryohei Ueda, Shunichi Nozawa, Yohei Kakiuchi, Kei Okada, and Masayuki Inaba. Global planning of whole-body manipulation by humanoid robot based on transition graph of object motion and contact switching. *Adv. Robotics*, 31 (6):322–340, 2017.

Karen L. Myers, Pauline Berry, Jim Blythe, Ken Conley, Melinda T. Gervasio, Deborah L. McGuinness, David N. Morley, Avi Pfeffer, Martha E. Pollack, and Milind Tambe. An Intelligent Personal Assistant for Task and Time Management. *AI Magazine*, 28(2):47–61, 2007.

Lorenzo Natale, Ali Paikan, Marco Randazzo, and Daniele Domenichelli. The iCub Software Architecture: Evolution and Lessons Learned. *Frontiers Robotics AI*, 3:24, 2016.

Allen Newell. *Unified Theories of Cognition*. Harvard University Press, Cambridge MA, 1990.

Andrew Y. Ng and Michael Jordan. On Discriminative vs. Generative classifiers: A comparison of logistic regression and naive Bayes. In *Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic*, NIPS'01, pages 841–848. MIT Press, 2001.

Nils J. Nilsson. Shakey the Robot. Technical Report 323, AI Center, SRI International, Menlo Park, CA, USA, 1984b.

Nils J. Nilsson. Eye on the Prize. *AI Magazine*, 16(2):9–17, 1995a.

Donald A. Norman and Tim Shallice. Attention to action: willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartz, and D. Shapiro, editors, *Consciousness and Self-Regulation. Advances in Research and Theory*, pages 1–18. Plenum Press, 1986.

Natasha Noy and Deborah McGuinness. Final Report on the 2013 NSF Workshop on Research Challenges and Opportunities in Knowledge Representation. National Science Foundation Workshop Report, 2013.

Andrew M Nuxoll and John E Laird. Enhancing intelligent agents with episodic memory. *Cognitive Systems Research*, 17: 34–48, 2012.

Alberto Olivares-Alarcos, Daniel Beßler, Alaa Khamis, Paulo Goncalves, Maki Habib, Julita Bermejo, Marcos Barreto, Mohammed Diab, Jan Rosell, João Quintas, Joanna Olszewska, Hirenkumar Nakawala, Edison Pignaton, Amelie Gyrard, Stefano Borgo, Guillem Alenyà, Michael Beetz, and Howard Li. A Review and Comparison of Ontology-based Approaches to Robot Autonomy. *The Knowledge Engineering Review*, 34:e29, 2019.

David Paulius and Yu Sun. A Survey of Knowledge Representation in Service Robotics. *Robotics Auton. Syst.*, 118:13–30, 2019.

Johannes Pfau, Robert Porzel, Mihai Pomarlan, Vanja Sophie Cangalovic, Supara Grudpan, Sebastian Höffner, John Bateman, and Rainer Malaka. Give MEANinGS to Robots with Kitchen Clash: A VR Human Computation Serious Game for World Knowledge Accumulation. In Erik van der Spek, Stefan Göbel, Ellen Yi-Luen Do, Esteban Clua, and Jannicke Baalsrud Hauge, editors, *Entertainment Computing and Serious Games*, volume 11863, pages 85–96, Cham, 2019a. Springer International Publishing.

Mihai Pomarlan and John A. Bateman. Embodied Functional Relations: a Formal Account Combining Abstract Logical Theory with Grounding in Simulation. In *Proceedings of the 11th International Conference on Formal Ontology in Information Systems*, FOIS, 2020.

David Poole and Alan K. Mackworth. *Artificial Intelligence - Foundations of Computational Agents*. Cambridge University Press, 2010.

Wolfgang Prinz, Miriam Beisert, and Arvid Herwig, editors. *Action Science — Foundations of an Emerging Discipline*. MIT Press, 2013.

Karinne Ramirez-Amaro, Michael Beetz, and Gordon Cheng. Transferring skills to humanoid robots by extracting semantic representations from observations of human activities. *Artif. Intell.*, 247:95–118, 2017.

Karinne Ramirez-Amaro, Yezhou Yang, and Gordon Cheng. A survey on semantic-based methods for the understanding of human movements. *Robotics and Autonomous Systems*, 119:31–50, 2019c.

Jesse Richter-Klug and Udo Frese. Towards Meaningful Uncertainty Information for CNN Based 6D Pose Estimates. In

International Conference on Computer Vision Systems, pages 408–422. Springer, 2019.

Stuart J. Russell and Peter Norvig. *Artificial Intelligence — A Modern Approach (3. internat. ed.)*. Pearson Education, 2010.

Werner Schneider, Josefine Albert, and Helge Ritter. Enabling cognitive behavior of humans, animals, and machines: A situation model framework. *ZiF-Mitteilungen*, 2020(1):21–34, 2020.

Gregor Schöner. Development as Change of Dynamic Systems: Stability, Instability, and Emergence. In J. P. Spencer, M. S. C. Thomas, and J. L. McClelland, editors, *Toward a Unified Theory of Development: Connectionism and Dynamic Systems Theory Re-Considered*, New York, 2009. Oxford University Press.

Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, Timothy P. Lillicrap, and David Silver. Mastering Atari, Go, Chess and Shogi by Planning with a Learned Model. *CoRR*, abs/1911.08265, 2019.

Holger Schultheis and Thomas Barkowsky. Casimir: An Architecture for Mental Spatial Knowledge Processing. *Topics in cognitive science*, 3 4:778–95, 2011.

Holger Schultheis, Sven Bertel, and Thomas Barkowsky. Modeling Mental Spatial Reasoning About Cardinal Directions. *Cognitive Science*, 38(8):1521–1561, 2014.

Carl P. L. Schultz, Mehul Bhatt, Jakob Suchan, and Przemyslaw Andrzej Walega. Answer Set Programming Modulo 'Space-Time'. In Christoph Benzmüller, Francesco Ricca, Xavier Parent, and Dumitru Roman, editors, *Rules and Reasoning - Second International Joint Conference, RuleML+RR 2018, Luxembourg, September 18-21, 2018, Proceedings*, volume 11092 of *Lecture Notes in Computer Science*, pages 318–326. Springer, 2018.

Gerhard Schurz and Paul D. Thorn. The Revenge of Ecological Rationality: Strategy-Selection by Meta-Induction Within Changing Environments. *Minds and Machines*, 26(1):31–59, mar 2016.

Myrna F. Schwartz and Laurel J. Buxbaum. *Naturalistic action*, pages 269–289. Psychology Press, 1997.

Myrna F. Schwartz, Mary Segal, Tracy Veramonti, Mary Ferraro, and Laurel J. Buxbaum. The Naturalistic Action Test: A standardized assessment for everyday action impairment. *Neuropsycological Rehabilitation*, 12(4):311–339, 2002.

Yun-Ming Shih, Collin Gordon, Munehiro Fukuda, Jasper van de Ven, and Christian Freksa. Translation of String-and-Pin-based Shortest Path Construction into Data-Scalable Agent-based Computational Models. In *2018 Winter Simulation Conference (WSC)*, pages 881–892. IEEE, 2018.

David Shurtleff, Michael Steinmetz, Bradley Wise, Antonio Noronha, Jill Heemskerk, Deborah Henken, Susan Weiss, Amy Poremba, Greg Farber, and Edmund Talley. The BRAIN initiative. braincanada.ca, 2013.

Bruno Siciliano and Oussama Khatib, editors. *Springer Handbook of Robotics*. Springer Handbooks. Springer, 2016.

David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the Game of Go with Deep Neural Networks and Tree Search. *Nature*, 529(7587):484–489, 1 2016.

Herbert Simon. Rational choice and the structure of the environment. *Psychological Review*, 63:129–138, 1956.

Herbert A. Simon. A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, 69(1):99–118, 1955.

Michael Spranger, Jakob Suchan, and Mehul Bhatt. Robust natural language processing - combining reasoning, cognitive semantics, and construction grammar for spatial language. In Subbarao Kambhampati, editor, *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016*, pages 2908–2914. IJCAI/AAAI Press, 2016.

Simon Stelter, Georg Bartels, and Michael Beetz. Multidimensional Time Series Shapelets Reliably Detect and Classify Contact Events in Force Measurements of Wiping Actions. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Vancouver, Canada, 2017. **Finalist for the Best Cognitive Robotics Paper Award.**

Jakob Suchan and Mehul Bhatt. Semantic question-answering with video and eye-tracking data: AI foundations for human visual perception driven cognitive film studies. In Subbarao Kambhampati, editor, *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016*, pages 2633–2639. IJCAI/AAAI Press, 2016.

Jakob Suchan and Mehul Bhatt. Deep Semantic Abstractions of Everyday Human Activities - On Commonsense Representations of Human Interactions. In Aníbal Ollero, Alberto Sanfeliu, Luis Montano, Nuno Lau, and Carlos Cardeira, editors, *ROBOT 2017: Third Iberian Robotics Conference - Volume 1, Seville, Spain, November 22-24, 2017*, volume 693 of *Advances in Intelligent Systems and Computing*, pages 477–488. Springer, 2017b.

Jakob Suchan, Mehul Bhatt, and Carl P. L. Schultz. Deeply Semantic Inductive Spatio-Temporal Learning. In James Cussens and Alessandra Russo, editors, *Proceedings of the 26th International Conference on Inductive Logic Programming (Short papers), London, UK, 2016*, volume 1865 of *CEUR Workshop Proceedings*, pages 73–80. CEUR-WS.org, 2016.

Jakob Suchan, Mehul Bhatt, Przemyslaw Andrzej Walega, and Carl P. L. Schultz. Visual Explanation by High-Level Abduction: On Answer-Set Programming Driven Reasoning About Moving Objects. In Sheila A. McIlraith and Kilian Q. Weinberger, editors, *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018*, pages 1965–1972. AAAI Press, 2018b.

Jakob Suchan, Mehul Bhatt, and Srikrishna Varadarajan. Out of Sight But Not Out of Mind: An Answer Set Programming Based Online Abduction Framework for Visual Sensemaking in Autonomous Driving. In Sarit Kraus, editor, *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*, pages 1879–1885. ijcai.org, 2019. **(received Distinguished Paper Honorable Mention Award)**.

Ron Sun. Desiderata for cognitive architectures. *Philosophical Psychology*, 17(3):341–373, 2004.

Leonard Talmy. Force dynamics in language and cognition. *Cognitive Science*, 12(1):49–100, 1988.

Toni Tan, René Weller, and Gabriel Zachmann. SIMDop: SIMD optimized Bounding Volume Hierarchies for Collision Detection. In *IROS*, pages 7256–7263, 2019.

Moritz Tenorth and Michael Beetz. KnowRob – A Knowledge Processing Infrastructure for Cognition-enabled Robots. *Int. Journal of Robotics Research*, 32(5):566 – 590, April 2013.

Moritz Tenorth and Michael Beetz. Representations for robot knowledge in the KnowRob framework. *Artificial Intelligence*, 2015.

Moritz Tenorth and Michael Beetz. Representations for robot knowledge in the KnowRob framework. *Artif. Intell.*, 247: 151–169, 2017.

Madhura Thosar, Sebastian Zug, Alpha Mary Skaria, and Akshay Jain. A Review of Knowledge Bases for Service Robots in Household Environments. In Antonio Chella, Ignazio Infantino, and Antonio Lieto, editors, *Proceedings of the 6th International Workshop on Artificial Intelligence and Cognition, Palermo, Italy, July 2-4, 2018*, volume 2418 of *CEUR Workshop Proceedings*, pages 98–110. CEUR-WS.org, 2018.

Paul Tibbetts. Cognitive Neuroscience: The Biology of the Mind. 2009.

Endel Tulving. Episodic and semantic memory 1. *Organization of Memory. London: Academic*, 381(e402):4, 1972.

M.T. Turvey. Affordances and Prospective Control: An Outline of the Ontology. *Ecological Psychology*, 4(3):173–187, 1992.

Constantin Uhde, Nicolas Berberich, Karinne Ramirez-Amaro, and Gordon Cheng. The Robot as Scientist: Using Mental Simulation to Test Causal Hypotheses Extracted from Human Activities in Virtual Reality. In *IROS 2020*, Las Vegas, USA, 2020.

Jasper van de Ven, Munehiro Fukuda, Holger Schultheis, Christian Freksa, and Thomas Barkowsky. Analyzing Strong Spatial Cognition: A Modeling Approach. In Sarah Creem-Regehr, Johannes Schöning, and Alexander Klippel, editors, *Spatial Cognition XI*, Lecture Notes in Computer Science, pages 197–208, Cham, 2018. Springer International Publishing.

Frank van Harmelen and Annette ten Teije. A Boxology of Design Patterns for Hybrid Learning and Reasoning Systems. *CoRR*, abs/1905.12389:97–124, 2019.

David Vernon. Editorial: special issue on cognitive vision. *Image and Vision Computing*, 26(1):1–4, 2008.

David Vernon. *Artificial Cognitive Systems: A Primer*. MIT Press, 2014.

David Vernon and Markus Vincze. Catalogue of Cognitive Systems Capabilities. Technical Report Deliverable D3.2, RockEU2Robotics Coordination Action for EuropeTwo, 2017.

Markus Vincze and David Vernon. Industrial Priorities for Cognitive Robotics. Technical Report Deliverable D3.1, RockEU2Robotics Coordination Action for EuropeTwo, 2017.

Przemyslaw Andrzej Walega, Mehul Bhatt, and Carl P. L. Schultz. ASPMT(QS): non-monotonic spatial reasoning with answer set programming modulo theories. In Francesco Calimeri, Giovambattista Ianni, and Miroslaw Truszczynski, editors, *Logic Programming and Nonmonotonic Reasoning - 13th International Conference, LPNMR 2015, Lexington, KY, USA, September 27-30, 2015. Proceedings*, volume 9345 of *Lecture Notes in Computer Science*, pages 488–501. Springer, 2015.

Petra Wenzl and Holger Schultheis. Optimality and Space in Weakly Constrained Everyday Activities. In S. Denison, M. Mack, Y. Xu, and B.C. Armstrong, editors, *Proceedings of the 42nd Annual Conference of the Cognitive Science Society*, pages 1866–1872. Cognitive Science Society, 2020a.

Petra Wenzl and Holger Schultheis. Spatial Representation in Sequential Action Organization of Weakly Constrained Everyday Activities. In Jurǵis Šķilters, Nora S. Newcombe, and David Uttal, editors, *Spatial Cognition XII*, Lecture Notes in Computer Science, pages 59–75, Cham, 2020b. Springer International Publishing.

Petra Wenzl and Holger Schultheis. What Everyday Activities Reveal About Spatial Representation and Planning Depth. In Terry C. Stewart, editor, *Proceedings of the 18th International Conference on Cognitive Modeling*, pages 302–308, University Park, PA, 2020c. Applied Cognitive Science Lab, Penn State.

Thomas Wimböck, Christoph Borst, Alin Albu-Schäffer, Christian Ott, Florian Schmidt, Matthias Fuchs, Werner Friedl, Oliver Eiberger, Andreas Baumann, Alexander Beyer, and Gerd Hirzinger. DLRs zweihändiger Humanoide Justin: Systementwurf, Integration und Regelung (DLR's Two-Handed Humanoid Justin: System Design, Integration and Control). *Automatisierungstechnik*, 58(11):622–629, 2010.

Patrick Henry Winston. The Right Way. *Advances in Cognitive Systems*, 1:23–36, 2012.

Andreas Zwergal, Florian Schöberl, Guoming Xiong, Cauchy Pradhan, Aleksandar Covic, Philipp Werner, Christoph Trapp, Peter Bartenstein, Christian la Fougère, Klaus Jahn, Marianne Dieterich, and Thomas Brandt. Anisotropy of Human

Horizontal and Vertical Navigation in Real Space: Behavioral and PET Correlates. *Cerebral Cortex*, 26(11):4392–4404, 2016.