Proposal for the Collaborative Research Centre 1320

EASE - EVERYDAY ACTIVITY SCIENCE AND ENGINEERING

2017 - 2021

Universität Bremen

Coordinator

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

Lena Jacobs
Universität Bremen
Am Fallturm 1
28359 Bremen, Germany
phone +49 (0)421 218 64000
email ai-office@cs.uni-bremen.de

CONTENTS

List of Figures

List of Tables

List of Abbreviations

BLORT	Block World Robotic Vision Toolbox
BMBF	Federal Ministry of Education and Research / Bundesministerium für Bildung und Forschung
BMWi	Federal Ministry for Economic Affairs and Energy / Bundesministerium für Wirtschaft und Energie
CITEC	Cluster of Excellence Cognitive Interaction Technology, Bielefeld University
CoTeSys	Cognition for Technical Systems
CRAM	Cognitive Robot Abstract Machine
CRC	Collaborative Research Center
DFG	German Research Foundation / Deutsche Forschungsgemeinschaft
DFKI	German Research Center for Artificial Intelligence / Deutsches Forschungszentrum für Künstliche Intelligenz
EASE	Everyday Activity Science and Engineering
FP7	7th Framework Program for Research and Technological Development
FSM	Finite State Machine
NAT	Naturalistic Action Test
OAC	Object-Action Complex
PCL	Point Cloud Library
PI	Principal Investigator
Prof	Professor
ROS	Robot Operating System
SPP	Priority Research Program / Schwerpunktprogramm
UB	University of Bremen / Universität Bremen
UBI	Bielefeld University / Universität Bielefeld
UIMA	Unstructured Information Management Architecture
WP	Work Package

List of Symbols and Icons

	Definition (Box)
	Citation (Box)
	Info (Box)

1 Research profile of EASE

1.1 Summary of the research program

Recently we have witnessed the first robotic agents **performing everyday manipulation activities** such as loading a dishwasher and 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**. Mastering everyday activities is an important step for robots to become the competent (co-)workers, assistants, and companions who are widely considered a necessity for dealing with the enormous challenges our aging society is facing.

We propose **Everyday Activity Science and Engineering (EASE)**, a fundamental research endeavor to investigate the cognitive information processing principles employed by humans to master everyday activities and to transfer the obtained insights to models for autonomous control of robotic agents. The aim of EASE is to boost the robustness, efficiency, and flexibility of various information processing subtasks necessary to master everyday activities by uncovering and exploiting the structures within these tasks.

Everyday activities are by definition mundane, mostly stereotypical, and performed regularly. The core research hypothesis of EASE is that robots can achieve mastery by exploiting the nature of everyday activities. We intend to investigate this hypothesis by focusing on two core principles: The first principle is **narrative-enabled episodic memories (NEEMs)**, which are data structures that enable robotic agents to draw knowledge from a large body of observations, experiences, or descriptions of activities. The NEEMs are used to find representations that can exploit the structure of activities by transferring tasks into problem spaces that are computationally easier to handle than the original spaces. These representations are termed **pragmatic everyday activity manifolds (PEAMs)**, analogous to the concept of *manifolds* as low-dimensional local representations in mathematics. The exploitation of PEAMs should enable agents to achieve the desired task performance while preserving computational feasibility.

The vision behind EASE is a cognition-enabled robot capable of performing **human-scale everyday manipulation tasks in the open world** based on high-level instructions and mastering them.

1.2 Detailed presentation of the research program

How can people...

... perform the appropriate actions with the appropriate objects in the appropriate ways when given an underspecified task such as "clean up"?

... perform everyday tasks even in unfamiliar environments with unfamiliar items?

... act competently and efficiently given the large amount of knowledge and reasoning required to do so?

Generative models for mastering everyday activity

These are some of the questions raised by one of today's most fundamental mysteries of nature: the human ability to produce efficient, flexible, and reliable complex, goal-directed behavior for vaguely specified tasks in open environments. Everyday Activity Science and Engineering (EASE) attempts to find answers to the questions above by designing, analyzing, and building new **cognition-enabled information processing models of agency** for mastering everyday household tasks, such as unloading the dishwasher or cooking a meal. The purpose of investigating the models is not only to understand how humans perform everyday activities but also how to enable robotic agents to master these activities (Figure 1).

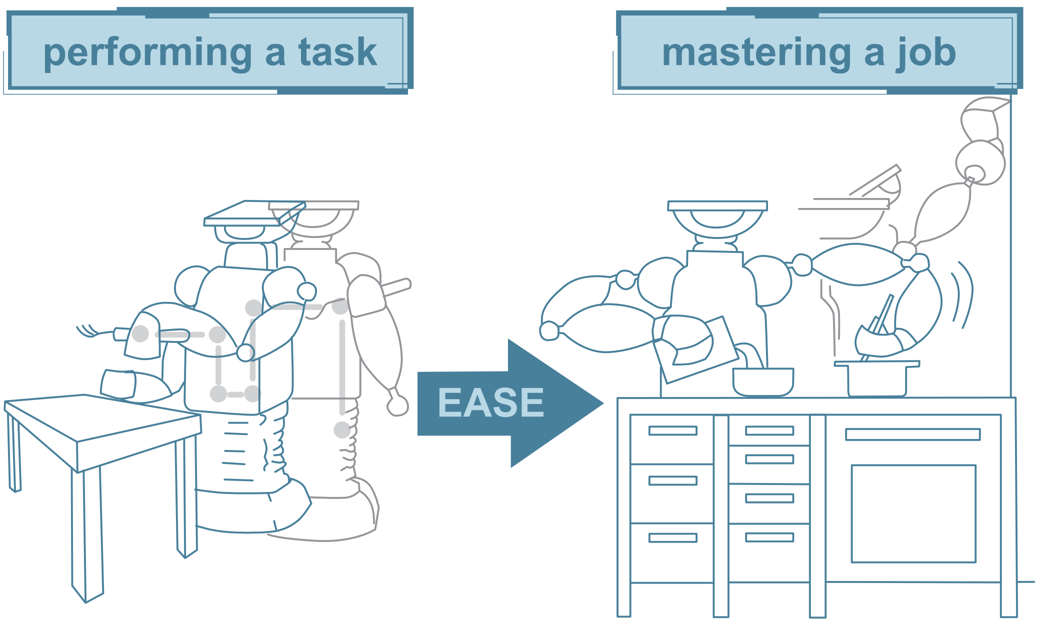


Figure 1: The EASE challenge: from performing specific tasks with specific objects in specific contexts to mastering human-scale everyday manipulation tasks.

D **Everyday Activity Science and Engineering (EASE)** is the study of the design, realization, and analysis of information processing models that enable robotic 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.

Mastery of everyday activity is important

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 live independently is determined by health professionals through assessments of their (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 tests (NATs) (Schwartz & Buxbaum, 1997), judge the ability of people "to select actions and objects at the right time and in the right order, and to engage in self-monitoring and error correction" (Lawton & Brody, 1969; Schwartz *et al.*, 2002). NATs specifically address everyday tasks that require the use of objects, the sequencing of multiple steps, and the achievement of nested goals (Giovannetti *et al.*, 2002).

D

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*[1] performance rather than expert or optimal performance is required."
— Definition by Anderson (1995)

Executing vague instructions requires a large body of knowledge

Cognition-enabled robot control (Beetz *et al.*, 2012) is a promising framework for enabling artificial agents to perform everyday activities at a level comparable to humans. To master everyday tasks, a robot cannot rely on plans that specify execution to the minute details because too much depends on the specific context and situation. Instead, a robot should be able to complete high-level tasks, such as "clean up", or execute vague instructions such as cooking recipes. Cognition-enabled robots try to do this by inferring the missing parameters in their plans, subplans and low-level control routines from what they perceive and what they know. To do this, a huge body of knowledge is required. This knowledge and associated reasoning methods include what is called commonsense and naive physics reasoning.

Achieving generality, flexibility and robustness

In this context, vague task specifications are not only a challenge but also part of the solution. We believe that the ability to execute tasks from vague instructions appropriately and competently is inherently related to and an essential resource for achieving the generality, flexibility, and robustness that is so characteristic for human everyday activity. These characteristics all depend on the ability to adapt a general activity or goal to a specific situation that may be subject to change, which is analogous to the ability of translating vague (general) instructions to specific executions.

D

Cognition-enabled agents, as defined in EASE, use information processing infrastructure for decision-making and action parametrization to enable them to satisfice their tasks in terms of performance measures such as generality, flexibility, robustness, and efficiency. They use knowledge, for example learned from (lifelong) experience, and reasoning methods, for example using *forward models* to predict the effects of intended actions, as a resource to decide on their exact course of action.

Tremendous impact

Having a better understanding of information processing models underlying the mastery of everyday activity will have a tremendous impact on our lives. These models and the application thereof to robot control are preconditions for building autonomous robotic agents such as robot

[1] *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.

(co-)workers, robot assistants, and robot companions (Dario *et al.*, 2011b). It is with good reason that many future technology observatories and roadmaps identify the development of these platforms as key to dealing with many pressing societal and industrial challenges. These challenges include population aging (Christensen, 2011; Schaal, 2007), natural disasters (Murphy *et al.*, 2008; Balakirsky *et al.*, 2007; Bradshaw *et al.*, 2003), and undesirable/dangerous work (CARETeam, 2009; Christensen, 2011; Kemp *et al.*, 2007).

Other robotic task domains

The challenges of tasks addressed as part of EASE are not particular only to everyday (household) activity but also characteristic of many other task domains. Therefore the findings of this research endeavor will be transferable to other robotic task domains such as conducting chemical experiments and the assembly of products with very low production numbers.

EASE focuses on advancing cognitive capabilities of robots

The ability to master everyday activity is hampered by both the physical and cognitive limitations of today's robots. Both capabilities interact considerably and improvements in physical capabilities may simplify the needed cognitive abilities and vice versa. However, even with their current physical limitations, advancing cognitive capabilities of robots is expected to boost their performance considerably. Robots have shown the ability to perform complex and difficult instances of everyday manipulation tasks such as cleaning a table with dexterity when remotely controlled through human operators. The inability of today's robotic agents to flexibly and robustly master these activities in various contexts and environments suggests that their cognitive capabilities are more limiting than the physical ones. Therefore, in EASE we will focus on cognitive capabilities.

In the following subsections we will describe the vision behind EASE (Section 1.2.1) and the challenges it aims to tackle (Section 1.2.2) in more detail. The goals of EASE require an integrated *complete systems* approach. The approach to building such a system is described in Section 1.2.3. The evaluation measures and expected impacts of EASE are outlined in Section 1.2.4. The two key concepts underlying the EASE research program are described in detail in Section 1.2.5 and Section 1.2.6, before outlining the research plan itself in Section 1.2.7. Finally, the infrastructure and software EASE will use for collaboration and realizing robotic agents are described in Section 1.2.8 and Section 1.2.10.

1.2.1 The EASE vision

The vision behind EASE is understanding generative information processing models underlying the mastery of everyday manipulation tasks in complete, integrated systems. These models should allow a cognition-enabled, robotic agent to autonomously perform human-scale everyday manipulation tasks competently for extended periods of time in an open world. It should be able to infer the appropriate course of action from high-level, underspecified task descriptions using the current task context and commonsense knowledge as resources. It should also be able to extend its repertoire of skills through lifelong learning: continually extending its action-relevant knowledge, learning new activities, and adapting existing skills to new objects, contexts, and environments by exploiting the nature of everyday activity.

Fundamental research endeavor

In the course of hundreds of millions of years, evolution has tailored the human brain for survival in an open environment and for flexible, robust, and efficient goal-directed action. Often, the action control performed by the brain can achieve nearly optimal performance without interrupting fluent agency. It is a mystery how near-optimal object manipulation and agency is possible, considering the perception-based understanding of situations, the multitude of decisions, the knowledge, and the foresight it requires. The research goals of uncovering and understanding the information processing and control principles that facilitate near-optimal agency, as planned in **Everyday Activity Science and Engineering** (EASE), is a research endeavor as challenging and fundamental as other research visions of science such as discovering the origin of the universe (Structure & of the Universe Roadmap Team, 2003) or decoding the human genome (Jasny & Kennedy, 2001).

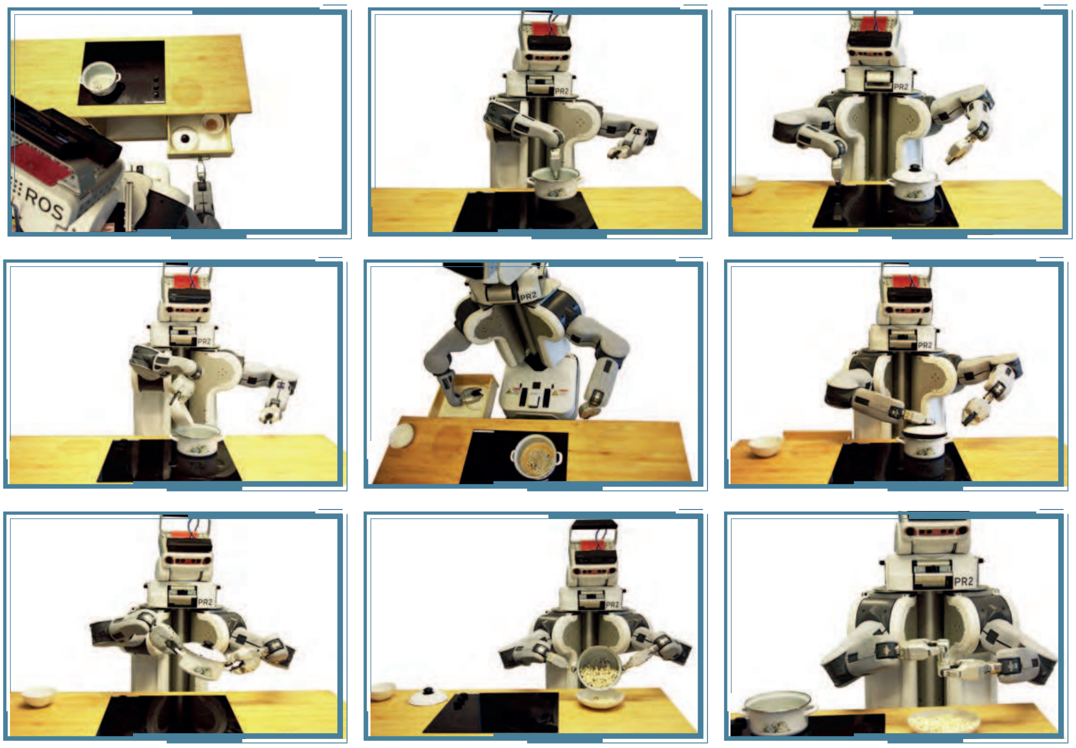


Figure 2: Autonomous mobile robot making popcorn. The snapshots show the variety of pick up, place, and other manipulation actions that are needed to make popcorn[a]

[a] `https://youtu.be/cTCJSNjTdo0`

EASE will attempt to solve some of this mystery in the context of manipulation tasks for everyday household activities. The goal is to understand how humans are able to perform these activities competently, flexibly and robustly over extended periods of time and construct information processing models that enable robotic agents to autonomously do the same.

Challenging goals

To those outside the field this task may seem deceivingly simple, because we humans perform complex daily activities with apparent ease (Bailey, 1997; Morris *et al.*, 2005; Anderson, 1995). It is something we are so familiar with and capable of that it is hard to imagine how difficult it is, since we are only aware of a fraction of all the processes that enable us to master such activities. Indeed, based on our own experiences it seems that playing chess (Newborn, 2000), expert problem-solving (Buchanan & Feigenbaum, 1978; Feigenbaum, 1993; Lenat & Feigenbaum, 1991), or playing Go (Silver *et al.*, 2016) are the hallmarks of human intelligence, because they are difficult *for us*. The field of AI has long since its conception learned that the "early" cognitive capabilities enabling humans to act successfully in a dynamic, uncertain, and open world are much harder to replicate, however (Brooks, 1991a,b; Agre, 1988, 1995; Agre & Horswill, 1997). The intuitive underappreciation for the complex nature of everyday activities also explains why science-fiction novels and movies have featured highly advanced, intelligent robots and AI systems for decades (Stork, 1997; Asimov, 1991; Slavicsek, 2000), while in reality robots have yet to materialize in human environments in a meaningful way.

These misconceptions regarding what is key to understanding and reconstructing human intelligence reflect how little grasp our intuitions have about the processes underlying our behavior. EASE is a long-term research enterprise for uncovering generative information processing models that can support the mastery of everyday activity. This includes performing actions with objects in an order and way in the given context that will achieve the desired effects as effectively and efficiently as possible given the resources. The capability to decide what is appropriate in a given task- and situational context requires solutions to a combination of very challenging problems in Robotics and Artificial Intelligence.

Holistic approach – relying on observation rather than intuition

EASE will use systematic observation of how humans perform such activities in natural and experimental settings, rather than our own introspective intuitions, to form the basis of its models. It will also use the systematic experience data from robots themselves executing household tasks to improve and master the activities. We believe that it is important to develop such models as part of a complete, integrated, physical system in order to truly advance the field. After all, simplifying the problem by focusing on solving subprocesses for components we think are important (or achievable) while ignoring the system as a whole would hold no guarantee that the set of components is complete or compatible with one another. The applicability of submodels can only be tested to a very limited extent without integration into a complete system.

Everyday household activity is an excellent domain

For investigating the information processing models and building such an integrated system, EASE will focus on everyday household tasks. They are considered an excellent domain of investigation for the following reasons.

Large, open challenge

First, mastering everyday manipulation tasks in a household setting is a large, open challenge. Everyday household tasks require the manipulation of objects in goal-directed and context-dependent ways, which makes the necessity and potential of embodied cognitive capabilities particularly evident (Rosenbaum *et al.*, 2012; Metta *et al.*, 2005)[2]. Figure 2 shows a few of the various actions needed for simple a cooking task such as making popcorn.

Complex decision making and integration required

As described by Müller *et al.* (2007), even seemingly simple everyday activities such as setting a table require considerable non-trivial judgments and decision making. To perform a high-level task such as "set the table for breakfast", the robot has to infer what to do: who to set the table for, where the participants will probably sit, what they will eat, and which utensils they need. Based on this information, the robot has to decide which objects to get, where to get them and how to arrange them on the table (Jain *et al.*, 2009). For each item the robot has to move, it has to decide where to stand, which arm to use, which grasp to apply to pick it up and where to grasp, to name only a few factors. Each of these decisions depends on the context: which objects are involved, the state of the object, the goal of the task, whether there are people present, their habits and preferences, and so on. This requires an efficient integration of current percepts, potential actions and background/commonsense knowledge. Wrong decisions can lead to plates being broken, dirty glasses being put on the table, cutlery being misplaced, etc.

Effects are direct and observable

Second, because the behavior displayed in the context of everyday manipulation and its effects is directly observable to the system, this domain is particularly well-suited for the investigation of cognition-enabled control systems that learn from experience data.

Recoverable failures

Third, failures in this domain are usually non-critical and recoverable. Even humans will sometimes forget an item for the breakfast table, spill when pouring a drink, or overboil the rice. Such events can form a valuable source for artificial systems to learn more not only about the specific task but also to learn commonsense knowledge.

Directly relevant for application areas

Finally, competence in the ability to perform everyday manipulation tasks in human working and living environments is directly relevant to many target areas for artificial agents and would boost their impact substantially (Kemp *et al.*, 2007).

[2]The application of the investigated methods to one-of-a-kind industrial assembly tasks will be researched in projects outside of EASE together with industrial partners including Bosch and Siemens.

Key research targets

To achieve the goal of constructing information processing models that enable autonomous, cognition-enabled agents to master everyday household tasks, EASE has identified the following key research targets:

- Acquiring and interpreting large, multi-modal data sets of humans performing everyday activities in experimental, artificial as well as natural settings.
- Acquiring and interpreting large, multi-modal data sets of robots performing human-scale everyday activities including complex manipulation tasks.
- Finding representational structures for activity data from different sources, both human and robotic, that optimally support the discovery of the structures and processes in everyday activity and the knowledge needed for its mastery.
- Discovering the processes and principles underlying human mastery of everyday activities.
- Deriving generally applicable knowledge, including commonsense and naive physics knowledge, from collections of experience data.
- Finding representations and reasoning mechanisms that enable artificial agents to complete tasks with competency similar to humans, by understanding the knowledge structures and processes underlying human mastery.
- Developing the software infrastructure to support the acquisition of commonsense knowledge from experience and the transformation of inference mechanisms to facilitate fast reasoning and decision making.
- Testing and improving the applicability of this knowledge and infrastructure in physical, humanoid robots by letting them perform many everyday household tasks.

In order to meet these targets, it is essential to bring together researchers from Cognitive (Neuro)science, Linguistics, Artificial Intelligence, and Robotics in a long-term collaborative research setting. Section 1.2.7 describes how the research areas and subprojects in EASE are connected. The organization and stimulation of the cooperation between the researchers from these various fields is described in Section 1.2.8.

1.2.2 The challenge

The key research problem addressed by EASE is the design, realization, and analysis of real-time, generative information processing models that provide robotic agents with effective and efficient means for flexibly and robustly executing vaguely formulated everyday tasks. These information processing models must be capable of acquiring the knowledge needed for task interpretation through means such as experience, observation, and reading. They must also be equipped with means that enable the agent to make the necessary inferences without delaying execution.

Problem in detail Here we shall consider the research problem of EASE in more detail. This will be done according to a set of key questions: (1) What is the problem? (2) Why is it important? (3) What is the common approach? (4) How do we intend to approach it? (5) Why is it possible now? (6) What will the impact be? and (7) How will progress be measured? The first 5 questions will be answered in this section, whereas the questions regarding impact and progress measure are answered in Section 1.2.4.

1. What is the problem? As touched upon in the previous section, everyday activities are commonly underspecified and the agent has to make a host of context-appropriate decisions to perform the actions appropriately. To do so, it needs a lot of background, commonsense, and naive physics knowledge, and this knowledge has to be relatable to its current perceptions and actions. Moreover, the perceptual, reasoning, decision-making and control processes should resolve at a speed that enables the robot to execute the actions smoothly and react timely.

Figure 3: Variations of performing a pouring action. The images raise the question of how a single and incomplete instruction, such as "pour some stuff into a container" can be translated to a plan with appropriate, sufficient information to execute such diverse manipulation actions.

Consider for example the case of preparing pancakes (Beetz *et al.*, 2011; Morgenstern, 2001). The robot first has to convert the recipe into a plan of actions. It then has to update the plan constantly to succeed in an unknown environment, because it will - for example - have to look for the items it needs in storage places.

It has to deal with different perceptual conditions and recognize pans even if they happen to be yellow instead of black. Moreover, perception and actuation are imperfect and might call for plan adjustments as well: The robot may not know exactly how much milk is left in the package; therefore it may pour too little or too much into the measuring cup. Using perception it should confirm whether this is the case and adjust the plan accordingly, for example by pouring again. But depending on the position of the measuring cup, its color, and the amount of daylight, the robot might not be able to read the marks on the cup. It should know that changing its own pose or the position of the cup could solve the perception problem. However, the mix will still not end up being the same every time. Therefore, the robot should know that it might have to pour less vigorously if the mix turned out thinner than usual or this particular frying pan is smaller than others.

Mastering activities is knowledge-intensive

Many of the decisions that have to be made to successfully accomplish everyday manipulation tasks are concerned with **how** to execute actions. For example, consider the subtask of "pouring something from a container". The robot should know that the container has to be held upright until it is above the target location, and that the appropriate tilting angle depends on many factors such as the fullness of the container, the type of target container (how likely is it to splash?), and the viscosity of the liquid. But even so, "pouring" actions can be very different depending on the circumstances. Figure 3 visualizes some of the different variations of actions and sophistication of skill that is needed to perform the task successfully. If a robotic agent has to perform a "pouring action", it has to consider questions such as: "Does the action require one or two hands?", "Does it require a certain tool?", and "Are there additional constraints such as holding the lid while pouring?".

The motions for flipping a pancake are even more specific. Small mistakes will lead to the pancake being broken, falling back on the same side, folding, etc. Such manipulation actions require in-depth knowledge of the causal relationship between the motion parameters and the physical effects of actions in order to execute them successfully.

Mastering activities requires fast reasoning

Finally, all these processes need to happen within a certain timeframe: If it takes the robot 15 minutes to establish how to flip the pancake, the pancake will already have burned. Ideally the execution of the action is so smooth that the "processing time" is not observable to humans and the waiting times would only occur where appropriate to the task, such as waiting for the pancake to be cooked through.

For the first phase of EASE, we will assume that the robot is mostly performing its tasks by itself. Interaction with other agents is an important aspect that we plan to integrate after after having established a solid foundation for performing competent everyday activity.

Two key research questions can be defined from this problem specification:

1. What are suitable data structures and information processing mechanisms for the acquisition of the knowledge that is required for the mastery of everyday activity?
2. How can the reasoning tasks needed for the competent execution of vague instructions be performed without noticeable computational delays?

Why is it important? As we have seen in the previous example, many eventualities come up during the execution of a single task in an unfamiliar environment. We believe that an essential requirement of competent manipulation activity is knowledge and the ability to reason about it effectively. Without knowledge, agents cannot infer the appropriate movements at the appropriate time, which means that they have to be specified by the developers.

Key approach to realizing autnomous robots

Hand-coding control programs will not enable us to scale towards human-scale manipulation activities in the real world, since it is not feasible to exhaustively cover every possible goal and every possible situation. Bestowing physical agents with the ability to act autonomously is key to realizing their full potential. Therefore, the question of how these agents can obtain the large amount of commonsense and naive-physics knowledge they need, and how they can process this knowledge and their percepts without delaying task execution, is essential.

Given their characteristics and structure, we have identified everyday household activities as a target domain of particular interest for realizing this type of autonomy and mastery as previously detailed in Section 1.2.1. It is a domain that is knowledge-intensive and instructions are typically incomplete; for a robot cleaning the table, its instruction will typically not include the exact pose of every object, the trajectory that each body part should follow, etc.

What is the conventional approach? Today the necessary knowledge is typically coded in the knowledge bases of agents. This approach reflects the *physical symbol system hypothesis* of Newell & Simon (1976), which states that the explicit representation of knowledge as physical symbols and their manipulation through symbol manipulation mechanisms are necessary and sufficient means for intelligent agency.

First we will look at conventional approaches to reasoning and robot control. Knowledge-enabled programs for agent control are typically designed as variations of the one stated in Russell & Norvig (2014)'s textbook "Artificial Intelligence — A Modern Approach." Figure 1 shows pseudocode for a program that works in this fashion. The program is iteratively called by the agent in order to compute the action to be executed in the current cycle. The first step is to take the data structure that is returned by the agent's perception system and translate it into symbolic statements that formally represent the beliefs of the agent about the world. These statements are then asserted into the agent's knowledge base, which

Algorithm 1 The classical agent program of a knowledge-based robotic agent as taught in Russell & Norvig (2014)'s textbook.

function KB-AGENT(*percept*)
  returns an action
  persistent *KB*, a knowledge base
    t, a counter, initially 0
 TELL(*KB*, make-percept-sentence(*percept*, t))
 action ← ASK(*KB*, make-action-query(t))
 TELL(*KB*, make-action-sentence(*action*, t))
 $t \leftarrow t + 1$
 return *action*
end function

contains all the beliefs of the agent that are needed in order to decide on the appropriate course of action. The action to be executed, which is also represented as an abstract symbolic expression, is computed by querying the inference system that infers the action by reasoning through the knowledge base.

There are several restrictive aspects to this approach, in particular in relation to the control of autonomous robotic agents performing complex manipulation tasks.

Classical representations are problematic: suboptimal, not scalable to the real world, and the abstractions lack mapping to robot motion

First, the robot has to abstract its percepts in such a way that they are appropriate for all possible uses in the inference processes. It has turned out that different inference tasks are often easier to solve if the reasoning agents use different representations tailored for the respective inference tasks, however.

Another problem is that the knowledge base is required to be consistent and complete with respect to the agent's beliefs. Consistent representations are very difficult and impossible to achieve if they are to represent continuous, noisy, and uncertain data and information.

Third, an unsolved problem is how an abstract representation of an action such as "add milk to the dough" can be translated into the appropriate motion and grasp specifications. This is in particular the case if a single symbolic action description has to account for behaviors that are as different as the ones depicted in Figure 3, which also vary in terms of their (side) effects.

So far, Robotics and Artificial Intelligence have not found reliable solutions for enabling robots to master everyday activities.

Even AI-planning, the field that focuses on the general ability to accomplish goals by performing the right actions from a wide range of possibilities, is only of limited use. Researchers in this field are mainly concerned with which actions to execute in which sequences, and represent actions at a high-level of abstraction. These abstract models are useful for and tailored to abstract planning, but do not afford reasoning about the details of the motion's execution. If we consider in how many ways flipping a pancake can go wrong, it becomes clear that the success or failure of actions often critically depends on the specifics of the execution.

Low-level solutions, such as those provided by neural networks, are by themselves also not able to support mastery of everyday activity. They are typically tailored to specific problems. Moreover, they have no graspable representation of conceptual information and such models do not afford reasoning about the task at a higher level. This makes it difficult to change the course of action when necessary, or adapt the models to work in other robots, environments and tasks.

Other approaches have been proposed to avoid some of the problems of logic-based commonsense reasoning and agency. Probabilistic reasoning (Thrun *et al.*, 2005) has successfully been applied primarily to considerably smaller reasoning domains or domains in which problem spaces can be effectively deconstructed and factorized.

Classical approach lacks scalable tools for acquiring naive physics and commonsense knowledge

Second, how the necessary knowledge for reasoning and control should be acquired also remains an open question. The importance of commonsense and naive physics knowledge in problem-solving has been recognized early by McCarthy (1968); McCarthy & Hayes (1969) "Programs with Common Sense") and later promoted by Hayes (1985a,b) ("Naive Physics Manifestos"). Many researchers in the area of knowledge representation and reasoning have investigated characteristic, commonsense reasoning (toy) problems by hand coding axiomatizations and developing inference methods to solve them. Scaling these approaches to so-called mid-size axiomatizations, such as the "egg-cracking" problem (Morgenstern, 2001), has proven to be very difficult and arguably so far no systems exist that comprehensively and convincingly cover the common reasoning capabilities needed for mastering everyday manipulation tasks.

The inability to scale might indicate that these are not the right approaches and representations, and the enterprise has been questioned repeatedly from in- and outside of the research community (see Brooks (1991a,b) and McDermott (1987)). Reasons for slow progress include the difficulties encountered in areas like the grounding of hand-coded symbols into the perception and action apparatus of physical robots, the search for abstractions suitable for a broad range of reasoning tasks, and the design of inference mechanisms that are correct and complete with respect to the commonsense reasoning tasks in question.

A recent, promising approach aims at acquiring commonsense knowledge through web-scale learning methods (Singh *et al.*, 2002). However, important parts of common knowledge needed for mastering everyday activity are acquired through experience and not stated explicitly in the web. Such knowledge is usually not put in text; for example how to hold a bottle when pouring, how to reach for objects, or how much force to apply when lifting and holding objects.

How do we approach the problem? EASE intends to use comprehensive experience data from various sources to acquire knowledge and optimize reasoning. The idea is to start with agents that have a set of generic plans that are carefully designed. These plans function as a starting point for acquiring the necessary commonsense and naive physics knowledge by gathering valuable experiences.

Acquiring knowledge and optimizing reasoning using experience

The human brain learns by collecting very detailed memories of activity episodes and consolidating and abstracting the highly situated knowledge in individual episodic memories into more generic, commonsense and other knowledge pieces that are applicable to a broader range of situations. We want to research an artificial memory system for artificial agents that is inspired by the role and functioning of the human episodic memory system.

The artificial episodic memories are annotated with narratives that give detailed explanations of the memorized activities so that they contain complete (low-level and high-level) information. For example, the action of picking up a filled container creates an episodic memory. The memory contains all control signals and percepts generated and received during the action. Additionally, it holds narrative information such as that the container was held at a certain angle to avoid spilling, and that the grasp was applied to the body of the container because this will make the pouring action in the next step easier to perform. These memories allow the robot to replay the execution meaningfully at a later time.

Collecting episodic memories containing complete, annotated information

An agent can learn commonsense concepts by generalizing and abstracting from the collection of episodic memories in which instances of the concept occur. A commonsense concept is for example “grasps that are suitable for picking up a heavy object with handles”. These concepts are similar to affordances, but are more closely related to actions and effects of actions than objects. For example, the robot can learn how heavy pots can be picked up by retrieving all episodic memories of picking up heavy pots and learn a classifier of situations that predicts whether or not picking up the object in a given situation will be successful. In this setup the agents can extend their commonsense and naive physics knowledge through any data mining and learning task that can be conducted on the collected episodic memories. We look at the acquisition of commonsense and naive physics knowledge as a computational problem of learning from vast amounts of subsymbolic data recorded from experience.

Episodic memories as basis for naive physics and commonsense knowledge

Moreover, the episodic memories enable the agent to form realistic expectations about the kinds of reasoning tasks it will encounter during a task and the context in which they are to be solved. These expectations enable the agent to tailor its reasoning to specific everyday activities and exploit the structure and regularities of the task. This allows the agent to reason faster and return better solutions because the process is based on experience rather than abstract performance models.

Tailoring reasoning to realistic expectations

Thus, we intend to develop mechanisms that allow the agents to analyze the structure and regularities of the complex reasoning tasks imposed by everyday activities. They will specialize reasoning methods to become more resource-efficient by exploiting these structures and regularities. Throughout this process the agents will evolve from the preliminary, designed stage to being able to complete more activities more effectively and efficiently.

Continuous evolution

Additionally, the robot plans are structured such as to make optimal use of these this knowledge and these mechanisms. Instead of searching for an abstract symbol structure that represents the appropriate action as in the classical approach, we propose that robot plans should contain high-level descriptions that query specific information from other subsystems. For example, it should be possible for the plan to query which motion parameters to use for a particular action. Detailed information such as the appropriate trajectory, grasp type, grasp force, etc. are computed and given to the plan when requested. An example of queries and answers such a control system would ask and receive is illustrated in Figure 4. Because the symbolic expressions that are inferred are very low-level and not at the abstract level of actions, the grounding of the symbolic expressions in the perception and action cycle is much easier. For example, the trajectory to reach an object, grasp type, and grasp points can be projected directly onto the RGB-D image perceived by the robot at execution time.

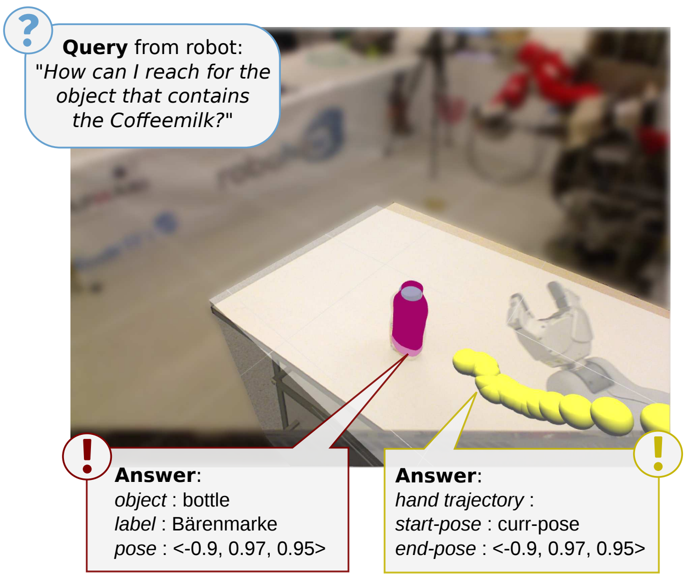


Figure 4: Example of a robot control system that asks for specific, grounded information for knowledge-enabled execution.

Finally, it should be noted that the nature of the resulting knowledge base from our proposed approach is very different from the classical one. Rather than requiring a single, complete, and consistent knowledge base we allow the knowledge base to be redundant and even inconsistent. The knowledge base contains all known information related to the queries. When tasked with answering a given query, the system proposes possible answer hypotheses and checks consistency only for the proposed answer. This requirement is much weaker than requiring whole knowledge bases to be consistent. Such an approach has been shown to scale towards open domain question answering applications, as demonstrated by the Watson system (Ferrucci *et al.*, 2010) .

Why is it possible now? Quantum leaps in information capturing and processing relevant to the control of everyday activity can be seen in several technology domains.

Spectacular progress in relevant fields

We have recently seen exciting and spectacular progress in scaling artificial systems to real-world tasks. Perhaps most prominent are the Watson system, the Siri agent, and the Google car. Watson (Ferrucci *et al.*, 2010) is a computer system that won the quiz show Jeopardy! against human champions. It demonstrated that systems can outperform humans in answering open, expert questions in previously unknown question categories by harvesting huge knowledge bases from the World Wide Web. The Siri agent (Apple) demonstrates, at least for narrow task domains, that we can equip computer systems with the capability of assisting and interacting with people when they utter vague spoken language instructions using the task, locational, and time context provided by the smartphone for disambiguation and interpretation. By October 2016 Google cars had driven, autonomously and with only one minor accident, over 2,000,000 miles through California, on highways, through inner cities, and on small mountain roads, dealing with many driving situations that are challenging even for experienced drivers. The Google car therefore is an impressive demonstration that long-term, autonomous, goal-directed activity is already within reach of the robotic systems. In all three cases, the breakthroughs were not achieved through a single novel method but by using a hybrid system that combines methods at a system level and by exploiting learning from large data and information sets. We are convinced that this kind of approach could also lead to powerful results for autonomous manipulation agents.

Breakthroughs by combining methods at system level and exploiting large datasets

In Robotics we have witnessed large improvements in hardware and information processing methods as well. They have enabled robots to perform single, experimental human-scale manipulation activities autonomously, such as making pancakes (Beetz *et al.*, 2011), serving as a butler (Srinivasa *et al.*, 2010), baking cookies (Bollini *et al.*, 2012), preparing salad (Gravot *et al.*, 2006; Okada *et al.*, 2006), folding laundry (Lakshmanan *et al.*, 2012), loading a dishwasher (Asfour *et al.*, 2006a), opening and closing furniture (Prats *et al.*, 2008), table setting (Stuckler *et al.*, 2012), or wiping (Leidner *et al.*, 2015). These efforts have substantially raised the skill level of robots in manipulation tasks.

Moreover, advances in sensor hardware and accompanying interpretation algorithms are providing robots with better capabilities to perceive the world. Improvements in hardware and software also provide better means for observing human activity. For example, 3D cameras such as Kinect have accelerated progress in tracking human poses (Shotton *et al.*, 2013) for controlling video games. The KinectFusion algorithm has been applied for scene reconstruction (Newcombe *et al.*, 2011; Izadi *et al.*, 2011), and object and scene perception (Anand *et al.*, 2013). The World Wide Web has made it much easier to share such advances and powerful services are readily available. For example, Google Speech[3] can be used to transform speech into text and Google Goggles[4] can identify objects from images and retrieve webpages featuring very similar objects. These services have the potential to make the robots' "understanding" of the world much more comprehensive.

On the knowledge side, available information sources have grown rapidly. We can now readily access online knowledge bases of many types (e.g. WordNet (Miller, 1995), Open Mind Common Sense (Singh *et al.*, 2002), ResearchCyc (Matuszek *et al.*, 2006b), ConceptNet, LifeNet, and StoryNet (Liu & Singh, 2004b)), CAD model bases (e.g. Google 3D warehouse), and websites providing instructions (e.g. wikiHow and eHow), to name only a few. Many of these websites provide their information in semi-structured ways, allowing computer programs to automatically extract much of the information they need. For the enormous amount of information online that is not readily understandable to machines, Web-scale learning (Halevy *et al.*, 2009) and unstructured information processing (Ferrucci & Lally, 2004a) provide new paradigms and show promise as potential work horses in the realization of the information processing infrastructure needed for mastering everyday activity.

Finally, the rapidly developing technologies for physical simulation, physics-based animation, and scene rendering can be utilized. These technologies provide us with software tools to envision manipulation actions and their effects much more realistically, form realistic expectations about incoming percepts, etc. New hardware and the usage of GPUs have increasingly expanded the scope of what is feasible in terms of computational resources. Related to this are advances in gaming technology and the development of Games with a Purpose (GwaPs). GwaPs are games specifically designed for harvesting knowledge that is difficult to obtain through other means (Ahn, 2006). People are asked to perform tasks that are difficult for machines in a game setting. The game environment motivates people to perform these tasks. This method allows us to acquire a large amount of labeled examples for machines to learn from, converting unsupervised learning tasks into supervised ones. GwaPs allow for crowdsourcing knowledge acquisition at low cost. For instance, Walther-Franks *et al.* (2015) demonstrated the potential of GwaPs for the acquisition of manipulation skills for food preparation tasks.

In summary, the mass of recent accomplishments in natural language processing, knowledge systems, real-world systems, sensor and computing hardware, and the Internet enable us to **take the investigation of the fundamental principles of mastering everyday activity to a level that was not possible before**. It is possible to collect much more background, action and commonsense knowledge than ever before. Symbolic knowledge obtained from the various online sources can be combined with realistic physical simulations and rendering models of percepts and actions to support better reasoning and decision-making processes. Advancing game technology allows users from the general public to "teach" artificial systems intuitively. Meanwhile, improvements in autonomous manipulation platforms' hardware, perception, and control enable

[3]`https://cloud.google.com/speech/`

[4]`https://support.google.com/websearch/answer/166331`

the agents to perform a wider range of actions, perceive a wider range of effects, and provide more detailed experience from longer-term operation. We believe that by utilizing the techniques, ideas and resources from these various developments it has become feasible to build comprehensive robot knowledge bases and control systems that are able to support mastery of everyday activity by autonomous robots.

Everyday activity main scenario

Main EASE scenario We have identified an everyday activity main scenario that EASE will use to investigate the information processing models for mastering everyday activity in phase 1. The scenario is designed to reflect the structure of everyday activity, its regularities as well as variations. Thus, it offers interesting and characteristic opportunities for learning everyday knowledge and skills as well as for exploiting the learned knowledge. Importantly, it gives the Collaborative Research Center an overarching goal from which the different subprojects can draw their research targets while also working together coherently towards the vision behind EASE.

To stimulate cooperation and the development of an integrated system with high-level abilities in performing housework, regular “housework marathons” are planned. During these marathons, the robotic agent will perform a large number of everyday manipulation task cycles (“robot days”) with variations in task and setting per day. This will provide a large amount of experience data for subsequent learning and function as a benchmark for assessing the combined efforts of the subprojects. The marathons should challenge the robot with respect to its capabilities of mastering everyday activities, taking into account the dexterity and sensing capabilities of the available robot hardware. Over the lifetime of EASE, the “household marathon” will be made more and more varied, complex, and challenging, in particular with respect to research questions and methods investigated by the EASE subprojects. EASE will provide and use three identical mobile manipulation platforms to make parallel, continuous, and long-term activity research realistic.

Note that to avoid overspecialization towards a single scenario, we plan to transfer the methods and tools developed in EASE to tasks involving the assembly of individual workpieces in a factory setting. This will be done in projects outside of EASE in the context of cooperations with Bosch Research and Siemens Corporate Technology.

Activity loop of household chores

Scenario as black-box The specific everyday activity scenario to be investigated in EASE will be a robotic agent performing daily household chores. The chores will be organized in an everyday activity loop. The main scenario is to perform the daily activities three times in a “robot day”, preparing breakfast, lunch, and dinner (see Figure 5).

EASE starts with a fundamental activity loop that is focused on the kitchen: for each meal the robot will have to set the table, clean the table, load the dishwasher, and unload the dishwasher to store items where they belong. These activities were chosen because they can be used to challenge the agent to perform underspecified tasks efficiently in unfamiliar environments with unanticipated events, while at the same time being relatively simple in terms of manipulation. This allows us to focus on the information processing models underlying the activities, and gathering and combining the knowledge from experience and the various sources mentioned above.

The fundamental loop is extended with meal preparation to form the full daily activity loop. Meal preparations allow for a host of complex manipulation tasks. Increasingly complex manipulations will be included over the years to assess and demonstrate the abilities of the system.

The specifics regarding the meal tasks are selected using a probabilistic task sampler that samples who will participate in the meals, what they would want to eat, which items they will need, and additional constraints. The task description typically only includes high-level information such as “clean the table”; information gaps will have to be filled by the agent using its everyday activity knowledge.

This daily activity loop is further extended with regular tasks whose purpose is environment stabilization (Hammond *et al.*, 1995) or to make future tasks easier. These tasks are typically recurring but not daily, such as cleaning the kitchen and the living room or (online) grocery shop-

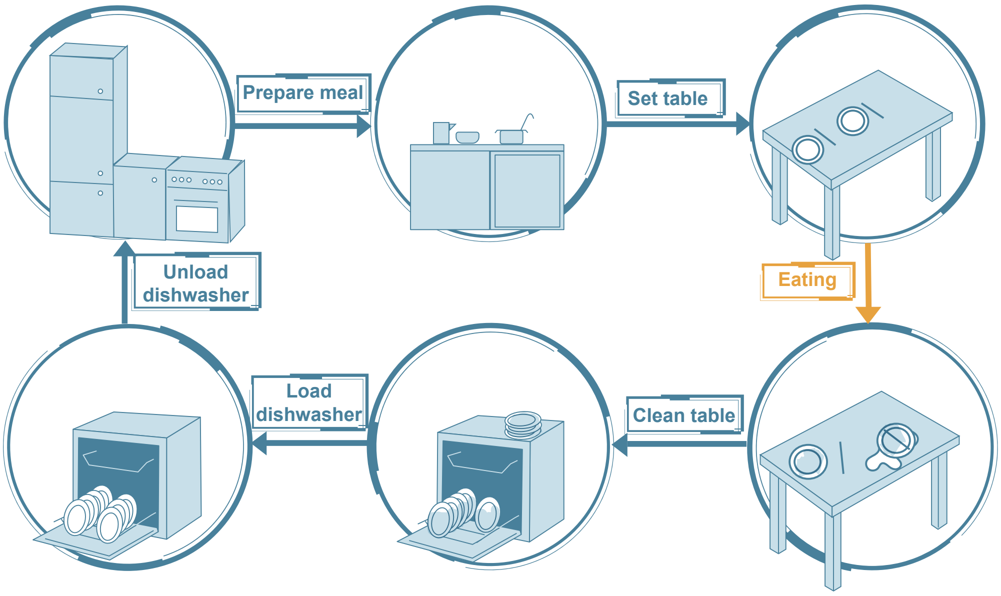


Figure 5: EASE main scenario: the everyday activity loop. The loop consists of 5 subtasks: (1) preparing the meal, (2) setting the table, (3) cleaning the table, (4) loading the dishwasher, and (5) unloading the dishwasher and putting the objects where they belong.

ping. For example, shopping restores the (filled) state of the pantry and makes the future task of meal preparation much easier (as compared to buying every single item the moment one needs it). In our scenario, shopping means that the robot has to generate a list of items currently missing from the kitchen. The ordered items are delivered in boxes that the robot has to unpack and store in the right places. This new type of task will challenge the robot in different ways, for example by introducing new objects in an existing environment. For the first phase, we will focus on the ability to autonomously perform tasks competently and learn from such experiences. Interaction with humans will not be part of this scenario until later.

At the end of phase 1, the robot should be able to perform a host of daily activities including preparing meals and cleaning up after meals, as well as regular chores such as (online) shopping and cleaning.

Milestone scenarios Household marathons with increasing complexity and difficulty will constitute the milestones for the EASE application scenario. The activities are to be demonstrated, discussed, and evaluated during the review meetings at the end of each research phase. For an overview of the phases we refer to Section 1.2.7. The complexity of these milestones will be scaled in terms of the duration of autonomy, task complexity, and other challenges provided by the environment.

Household marathon milestones

For the Year-4 Milestone we envision a household marathon in which the EASE robotic agents perform the entire activity loop for a month's worth of "robot days." Thus, the agents should prepare simple meals, set the table, clean the table, load the dishwasher, and store items from the dishwasher where they belong, clean the kitchen, and do the shopping. The measures of success according to which EASE, and the main scenario, will be evaluated are described in Section 1.2.4.

EASE research targets the intersection of several disruptive technologies and expects to achieve substantial advances herein. Therefore, rather than trying to specify the goals and details of subsequent milestone scenarios now, they will be fleshed out as part of the proposal for the second EASE phase in accordance to the goals. As touched upon in the black-box description, the scenario from the Year-4 Milestone can be extended in many ways: extending the environment from the kitchen to a complete apartment, performing tasks in previously unknown environments, adding more complex tasks, formulating tasks more vaguely, performing tasks in the presence of humans. These are the dimensions along which the challenges for the EASE research enterprise will be defined.

1.2.3 Information processing and control model

Two core concepts

The EASE concepts for everyday activity We propose to answer the two central research questions of EASE (see Section 1.2.2) through research centered on two core concepts. The first concept is that of *narrative-enabled episodic memories (NEEMs)*[5]. NEEMs are memories of activities stored in data structures that enable extraction of everyday activity knowledge. EASE will maintain and use a large knowledge base of NEEMs. The second concept is that of *pragmatic everyday activity manifolds (PEAMs)*. PEAMs are approximations of the complex computational problems EASE is aiming to solve. These approximations will be found by exploiting the structure in everyday activities and using the knowledge extracted from the NEEMs. Developing these two core concepts and tackling these challenges will result in artificial systems with better reliability, flexibility, adaptiveness and performance.

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.

NEEMs are sources of knowledge for robots (analogous to episodic memory)

Concept 1: NEEMs are a way of storing the data generated by robotic agents during everyday manipulation in such a way that enables knowledge extraction. The concept was inspired by the human episodic memory system. *Episodic memory* in humans 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, 2002a). Similarly, artificial agents should be able to acquire much of the knowledge needed for mastering everyday activity through NEEMs.

NEEMs contain comprehensive activity knowledge

While performing an activity, such as cleaning a room, 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 turns the robot into a "cognitive" agent that knows what it did, why, how, how well, etc. (Brachman, 2002). The robot can answer queries such as: "Where do the kids leave their toys after playing?", "Which is the best order to perform the cleaning sub-steps?", "When are good times for cleaning?", and "Which perception routines work best for cleaning up?".

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

[5] "narrative" is used here in the sense of "a symbolic account of connected events", analogous to the corresponding definition in the Oxford dictionary (`http://www.oxforddictionaries.com/definition/english/narrative`

past episodes to learn new information, even for aspects that were not previously considered for that particular episode. NEEMs are described in detail in Section 1.2.5.

Pragmatic everyday activity manifolds (PEAMs) are renderings of computational problems needed for mastering everyday activity, which in general form might not be computationally feasible, into problem spaces spanned by some sets of "manifolds" that allow for tractable and feasible reasoning solutions. Methods to create such "manifolds" include approximations, application of heuristics, making independence assumptions, dimensionality reduction, task specialization, and using stereotypical solutions and "generate-and-test" methods.

Concept 2: Our definition of everyday activities (page 3) characterizes them as activities "about which an agent has a great deal of knowledge, coming as a result of the activity being *common*, which is the primary contributor to its mundane nature" (Anderson, 1995). Concept 2 revolves around how the knowledge acquired through NEEMs can be used such that the perception, reasoning, and planning tasks for an action can be performed efficiently without delaying execution.

PEAMs boost the speed of robotic agents' reasoning processes

Since everyday activities are predictable and structured in nature, we expect that underlying low-dimensional spaces can be identified in which the computational problems for everyday activities become easier to solve than in their original formulation, which often involves large search spaces. We refer to these spaces as PEAMs, in analogy to the mathematical concept of manifolds referring to low-dimensional local representations but used in a much broader notion beyond the mathematical term. We aim at uncovering structures (and methods operating on them) that often solve the original problems approximately, probably[6], or in a satisfactory[7] manner, but are still expressive enough to produce the required task performance.

PEAMs make the computational problems of everyday activities tractable

Examples of complex computational problems that we expect to benefit from PEAMs are scene recognition, logic-based query answering, action planning, spatial and temporal reasoning, and diagnosis. Especially in spatial and temporal reasoning, the intractability appears to be partly caused by inadequate formalization and representation, which often blows up the physical problem into a large set of mostly impossible (in the real world) substates. Consider, for example, reaching trajectories for picking up and placing objects. Instead of trying to plan optimal trajectories for each pick and place instance, the robot considers the set of stereotypical trajectories, a kind of "manifold" embedded in the domain of all possible reaching motions. The reduced entropy of stereotypical motions has several advantages: they are easier to learn, are more legible by others, and facilitate failure detection, prediction, and diagnosis.

PEAMs are described in detail in Section 1.2.6.

NEEMs and PEAMs will be incorporated into **generative models[8] of everyday activity**, which address the research question of how the necessary knowledge can be acquired and how efficient knowledge-enabled problem solving can be realized.

The two core concepts will be investigated analytically by studying the way humans master their everyday activities (Research Area H), the information processing principles behind efficient execution of everyday activities (Research Area P), and the generative information processing

[6] approximately and probably in the sense of *probably approximately correct (PAC) learning*

[7] satisfactory in the sense of Simon's *satisficing* decision making strategy

[8] generative model as opposed to an analytical one, i.e., a model that can not only analyze but also generate behavior

models that will be embodied into robotic agents (Research Area R). The activities are analyzed on the basis of collected NEEMs and knowledge derived from NEEMs in order to identify PEAMs that can be exploited in order to improve the performance of the aforementioned tasks.

Details of the research plan are described and discussed in Section 1.2.7.

The proposed model The results from the research on the two core concepts will be incorporated into a complete information processing and control model that enables robotic agents to perform and later master human-scale manipulation activities. As the research in EASE progresses, the realization of the concepts will be refined and revised. The implementation of the model will be adapted accordingly.

The model has three main tasks: first, plan-based control for executing vaguely described everyday tasks; second, collection of experiences from everyday activities (NEEMs) and deriving generalized knowledge from NEEMs; and third, the creation and optimization of task-specific plans through plan specialization based on detected and exploited PEAMs.

The first task of plan-based control will be based on past research in cognition-enabled control (Beetz *et al.*, 2012). Implementations of this control model will be available from the start of EASE and will be built upon to construct a more effective and efficient control model integrated with the other two components.

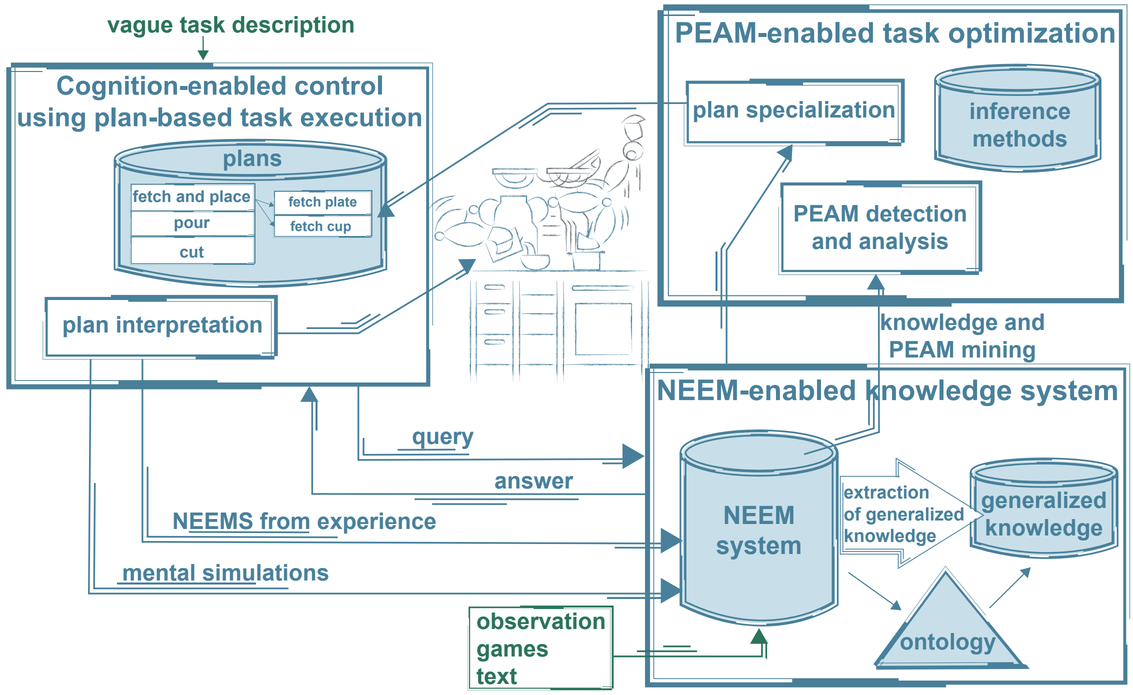


Figure 6: EASE's initial proposal for an information processing model to enable robotic agents to master everyday activity.

Our initial proposal for an information processing model is visualized in Figure 6. The components work together as follows:

1. **Cognition-enabled control using plan-based task execution.** The control model receives a vague instruction such as "put a plate on the table" in the context of table setting. The model looks up a suitable plan from a library of predefined, generic plans. For the example instruction, it might find a generic plan for fetching an object from a place and placing it somewhere. This plan is made specific using the (incomplete) instruction received and the generalized knowledge the model has access to (using the NEEM-enabled knowledge system). Information such as where to get the plate, how to hold the plate, where exactly to put down the plate,

etc. are determined by reasoning modules using the knowledge system, in interaction with perception and execution models.

2. **NEEM-enabled knowledge system**. Knowledge is of key importance for the control model to instantiate plans appropriately, since the instruction will most likely not include all the information a robotic agent requires to execute a task. It has to "complete" the instructions on both a high level (where are the plates?) as well as on a low level (what are the appropriate control parameters for grasping that plate?). Every time a plan is executed, there is an opportunity for the model to learn from this experience as well. It can also obtain knowledge from other sources, such as (natural language) text about table setting, observing humans performing the task, and mentally simulate variations on parts of the task such as different ways to pick up the plate. Thus, the second component is concerned with storing *experiences* as NEEMs and using them to obtain (generalized) knowledge for the control model to use.

3. **PEAM-enabled task optimization**. Putting a plate on the table is an everyday activity that is repeated many times. The model can use the wealth of NEEMs to find ways to optimize its plan and reasoning process (using PEAMs). This means the robot has to tailor to the specifics of the reasoning task and exploit the structure of the context in which the reasoning task is to be conducted. The plan is specialized by (1) learning the distribution of queries and processes required for putting a plate on the table in this table setting context. In other words, learning the regularities and knowledge/reasoning requirements typically needed for this task. (2) learning knowledge from past experiences such as where the clean plates are likely to be in this particular environment, how to optimize recognition of plates in a cupboard, and which grasps are most effective; and (3) analyzing the information from 1 and 2 to generate special purpose reasoners that are optimized for these requirements to provide answers to queries and paramaterizations more effectively and efficiently.

These components are described in detail in the remainder of this section.

Cognition-enabled control using plan-based task execution Mastering everyday activities involves performing the same action appropriately in different tasks and environments. This is only feasible using flexible and adaptive task descriptions. Cognition-enabled control uses generic plans to address this need for flexibility and robustness. Generic plans describe an action at a high enough level to be applicable irrespective of the particular conditions. The plans contain abstract descriptions of what the subactions, motions, and parameters should adhere to without stating the specific solutions. For example, in a plan in which an object has to be grasped, the plan does not specify exactly which grasp to use with which control parameters, e.g. the specific pose of the gripper and the force to apply. Rather it contains a high-level description of the requirements of the grasp – for example, that pots should be grasped by the handles and that glasses should be touched with less than maximum force. We represent these abstract vague descriptions as entities with their corresponding list of constraints in the form of key-value pairs. For example, the grasping action can be represented as follows:

```
(an action (type grasping) (object pot) (from handle) (hand both-hands))
```

The plan interpreter then resolves the description to the grasp most suitable in this particular task by querying the robot's knowledge base and combining the knowledge with its current percepts and the respective task context. Thus, generic plans are parametrized *on the spot* according to the situation at hand. This approach to planning and control has been successful in generating adaptive behavior (McDermott, 1992a; Beetz, 2001) and its features and feasibility have been tested in previous experiments (Beetz *et al.*, 2012).

Starting with a library of engineered, generic plans

We will equip robotic agents with a library of carefully designed generic plans to start with. These plans are associated with action verbs common in everyday activity such as *fetch*, *put*, *open*, *close*, *pour*, *stir*, *wipe*, etc. Because of the small set of general action verbs that can describe the majority of everyday household activity (Nyga & Beetz, 2012), manually designing generic plans is effortful but doable, in contrast to highly specific plans. We expect that the main EASE scenario can be covered by around 30 plans in total. The plan structure is such that a plan can call subplans. So-called high-level plans will mostly call other plans whereas low-level plans translate the symbolic expressions into motion and perception.

An example of a generic plan for the high-level activity of tablesetting is sketched in Algorithm 2.

Algorithm 2 A generic plan for setting a table.

```
def-plan SET-TABLE(meal-constraints)
    needed-items ← (all objects (needed-for meal) meal-constraints)
    arrangement ← infer set of spatial constraints for needed-items
    for all obj ∈ needed-items do
        FETCH(obj)
        PLACE-AT(obj, (a location (satisfies arrangement)))
    end for
    drinks ← (all drinks (needed-for meal) meal-constraints)
    for all drink ∈ drinks do
        container ← (an object (type container) (contains drink))
        drinkware ← (an object (type drinkware) (suitable-for drink))
        FETCH(container)
        POUR(drink, container, (a location (in drinkware)))
    end for
end def-plan
```

The plan states that the robot should get the items that are needed for a given meal, put them on the table in the appropriate arrangement, and pour drinks. When the plan is read and *needed-items* are unspecified, the system computes a set. The probabilities that items are needed for the meal are computed based on all information available to the robot: specifications by the user, online text sources, past experiences, etc. If the probability exceeds a specified threshold, the object is added to the local plan variable *needed-items*. Subsequently, the set of qualitative geometric constraints constituting the arrangement of the objects on the table is computed. For example, it will specify that the knife goes right of and in proximity of the plate. Once *needed-items* and *arrangement* are computed, the plan states to get each object and place it according to the arrangement.

The corresponding calls FETCH and PLACE-AT require the *obj* parameter. If *obj* is not yet bound to a perceived object in the world, the control system will call the perception system to locate the object in the current state of the world.

The same process holds for *drinks* as for *needed-items*: the plan should infer a set of *drinks* that go with the meal, infer which containers could contain them and which drink should be served into which drinkware, fetch the container and execute the pouring action.

Here we will not describe in detail how the symbolic queries are translated and answered by the perception system. Instead, we refer to Bálint-Benczédi *et al.* (2016). Similarly, the translation of the plans to motion constraints and objective functions is described by Bartels *et al.* (2013). A short explanation of the systems EASE intends to use is given in Section 1.2.10. Given the demonstrations (in household applications, among others) already developed by the members of the consortium, the existing systems are assumed to provide the necessary core perception and manipulation abilities to acquire experience data for EASE. These abilities are then extended through our research in PEAMs (as in Subproject R02).

The generic plan for pouring is shown in Algorithm 3. The plan consists of a signature that states the name of the plan and its formal parameters. The parameters (e.g., *an object*, *a location*, *a substance*) are part of an object ontology from the knowledge base, containing information about these concepts. For example, an object of type container as specified in the pouring plan cannot be a container that is too big to fit into the hand, because then it wouldn't have the affordance of *picking up*. The body of the plan specifies pouring as follows: GRASP the container that contains the substance to be poured, HOLD it above the destination, and TILT the container until the desired amount of the substance is at the destination. When GRASP is called, the validity of *source* as established previously by the perception system can be checked to make sure the object is still at the expected position and state. Then the control system will request the computation of motion parameters to move towards the object, based on the information it has received and which higher-level plan it is being called by.

Algorithm 3 A generic plan for a pouring action.

```
def-plan POUR( theme,       : (a substance)                        )
               source,      : (an object
                                  (type container)
                                  (contains theme)
                                  (affordance (an action
                                                 (type picking-up))))
               destination  : (a location)
    GRASP(source)
    HOLD-AT(source, (a location (above destination)))
    repeat TILT(source)
    until (the amount (of theme)) − (the amount (of substance) (at destination)) ≤ θ
end def-plan
```

Generic plans are interpreted and parameterized on the spot

The plan is structured in such a way that when the robot executes it, vague action descriptions are automatically transformed into queries for the underspecified/missing information. Interpretation and parametrization of the plan are done on the fly using a lot of commonsense and naive physics reasoning. The plan library must, therefore, be complemented with a system that can answer the queries posed to complete the plan. For example, consider that milk should be poured into a glass. If the milk is in a tetrapack, the robot must reason whether it could grasp the pack with one hand and how much force it should apply so the pack doesn't slip while also not squeezing so hard that the milk will spill out. If the milk was heated, the same action description would imply grasping the pot in which the milk is contained. In that case the robot has to infer that two hands will be needed, that it should grasp the handles on the side, that the pouring requires tilting the pot around the axis between the two handles, and so on.

In reality, plans for mastering everyday activity are much more complex than the examples given here. Plans should not only specify the intended course of action, but also detect and respond appropriately to asynchronous relevant events. The system has to detect execution failures, diagnose them, and try to recover appropriately. It must also be capable of interrupting plan execution safely to perform tasks of higher urgency (e.g., the robot notices that the milk boils over) and continue the interrupted activity afterwards.

The plans are not only designed to be flexible and robust but also to facilitate reasoning about the intended course of action and the effects that it causes (Brachman, 2002). To this end, plans are designed to be modular and transparent (Beetz, 2002b). **Modularity** means that subplans serve exactly one purpose. **Transparency** means that the purpose is explicitly stated in the subplan. This means that the agent can assume that a plan is designed to achieve all the goals of its subplans and if a goal was not explicitly stated, the plan was not trying to achieve that. Under these assertions many inference tasks needed for understanding a plan can be realized through simple pattern-directed subplan retrieval.

Specializing plans

Finally, we will enable the robots to use experiences to generate specialized plans for specific activities by modifying generic plans. For example, the robot could refine the generic pouring plan to a generic plan for pouring pancake batter, pouring beer into a glass, or pouring milk from a pot. In the first stage, programmers will design these plans based on suggestions that are generated by the reasoning mechanisms of the robot. By the end of the first four-year phase, plans that can be automatically learned are investigated. The specialized plans are expected to optimize execution and eliminate the need for complicated reasoning by exploiting the experience of the robot in performing the respective task. This is an important step towards realizing mastery of everyday activity.

NEEM-enabled knowledge system To determine how generic actions should be instantiated and parametrized given a specific situation, the control model needs massive amounts of knowledge and a comprehensive system for using it. To know how one should grasp an object to lift it, the control model requires the ability to predict the effects of the positions and forces of the fingers in interaction with the physical structure of the object. It also has to be able to predict how effects unfold and interact given certain actions in order to achieve or avoid certain effects reliably. Much of this knowledge is generally applicable across tasks and domains and consists in large part of commonsense and naive physics knowledge.

The set of rules required to describe this knowledge is vast and infeasible to generate manually. A more feasible approach is to equip the system with the ability to autonomously learn such rules. EASE's approach is to learn this knowledge from large collections of NEEMs. NEEMs contain comprehensive recordings of an activity. They contain detailed, low-level recordings of what has happened during the execution, including sensory data (images, body poses, etc.) and the interpretation thereof by perception routines. The design of the plans and the semantics of the plan language enables the plan interpreter to link and annotate these recordings with a narrative that explains the activity in terms of beliefs, desires, intentions, causes and effects.

From NEEMs we extract highly situated information and abstract it into broadly applicable knowledge. The specific and general knowledge are used to complete underspecified tasks and adapt to the events and situations where necessary.

Creating comprehensive knowledge bases

To realize this approach, we intend to build infrastructure for collecting, storing, and managing NEEMs of the executed activities and integrate this with the plans. This infrastructure can be seen as an artificial episodic memory system and will be essential to learning the knowledge needed for the mastery of everyday tasks. EASE will develop methods for effectively compressing and managing large collections of NEEMs, and analyze and acquire knowledge from them. This includes, for example, knowledge about the environment and the objects in it, the capabilities of the robot, and how to parametrize actions to achieve certain effects. The expected results are comprehensive declarative knowledge bases about the environment, tasks, commonsense knowledge, and naive physics knowledge that facilitate the relevant everyday activities.

Abstracting and generalizing episodic memories

Consider a robot that has collected a large dataset of SET TABLE and CLEAN TABLE tasks. The knowledge contained in the individual episodes is highly situated in the respective execution context. By abstracting and generalizing over these sets of episodic memories the robotic agent can learn commonsense and naive physics knowledge. For example, it can learn where objects are likely to be found in the contexts of specific tasks, what grasps are successful in which conditions, and how the visual appearances of objects change depending on where they are, etc. The behavioral results of actions, i.e. success/failure and other effects, are also stored. Therefore, the data will be suited for supervised as well as unsupervised learning methods.

The symbolic and subsymbolic information contained in semantically similar experiences can be combined to make predictions. For example, in the context of SET TABLE one likely has to get the plates from the cupboard, whereas during CLEAN TABLE one has to get the plates that are on the table.

Not only general knowledge can be queried, but also specific queries in relation to current percepts can be made. For example, if the plan queries "which container should I use to put water in?", the control model retrieves the knowledge that in this task, the water has to be boiled and therefore should be put in a pot. This pot should have a volume larger than 1 liter to hold the required amount of liquid and prevent overboiling. This knowledge is used for the perceptual interpretation of the current scene to identify a container that matches these characteristics.

Accelerating through cultural learning

To accelerate learning, the robots can gather more information by testing (variations on) their behavior in simulation. The robot can obtain NEEMs from other sources as well. It can boost the acquisition of everyday knowledge through cultural learning.[9] Our agent employs two forms of cultural learning: (1) learning through the observation of other agents performing the respective activities, and (2) reading about everyday activity.

NEEMs from multiple sources

Thus, we intend to create NEEMs not only from the experiences of robots performing activities, but also from observing humans and parsing their behavior. They can also be formed using Games with a Purpose (GwaPs), where the system generates simulation games in which people perform tasks the system is trying to understand. Additionally, the system can collect NEEMs from reading (instruction) texts. Feldman & Narayanan (2004) argue that people use mental simulation to understand the meaning and implications of texts. We can use a similar setup wherein the system generates behavior in simulation based on the texts and stores what happens in order to understand the text. The fact that the information from these different sources will be stored using the same representational structure makes it much easier to combine and learn knowledge from these sources.

PEAM-enabled task optimization To master everyday activity, we do not only need knowledge, but also methods for using this knowledge in relation to current percepts. The robotic agents can start by using general methods for solving the perception tasks and answering the queries contained in the generic plans. However, this often requires the robotic agents to search through very large search spaces and apply very complex reasoning methods. Mastery of everyday activities can be substantially advanced by optimizing the inference and perception mechanisms for the specific tasks that are to be expected, thereby limiting the search space.

Exploiting problem structure to optimize reasoning

Consider for example the perception and reasoning tasks required to fetch an object. In general terms, the perception system has to detect and localize partially described objects in potentially very cluttered scenes and reconstruct a shape model of the object that is sufficient for grasping it in the right way. This perception task in this general form is very hard and computationally very expensive to solve. However, one might want to consider that a fetch plan called in a table setting task is likely to have a very different distribution of perception and reasoning tasks compared to when it is called in the context of unpacking a shopping bag. This knowledge can be used to limit search space and use specialized methods for solving that particular task. In the context of table setting, the robot should look for plates and cups in the cupboard. It can expect the plates to be stacked in the cupboard. It knows that it should take the plate on top of the pile, and it can form strong expectations about how a pile of plates looks and what a good "top plate detector" for the cupboard would be. In the context of unpacking shopping bags, the robot can expect many of the objects to be packed goods from the supermarket. For these objects it can apply appearance-based recognition methods, using Google Goggles to read the labels on boxes and cans and look for bar codes. It can use these clues to retrieve semantic information about the products from the Internet. The robot might even have strong expectations about what is in the bag because it has seen the shopping list or knows which required objects are missing. Thus by considering the characteristics of a particular context, the perception problem can be solved more effectively and efficiently.

[9] Cultural learning refers to the way individuals of a group pass on information to each other, enabling the individuals to acquire knowledge and skills far beyond what they could acquire independently in a lifetime. See for example the "cultural intelligence hypothesis" Herrmann *et al.* (2007)

NEEMs are well-suited for detecting these characteristics and regularities in tasks. Using memories of applying the generic methods to specific everyday tasks in specific environments, the robotic agents can analyze which inferential capabilities are not needed in the respective tasks and what additional structures and regularities of these tasks can be exploited in order to speed up the inference and search processes.

For example, the agent can use NEEMs to:

- Infer the set of all reasoning problems it has to solve by parsing through all plans in its plan library and extracting each query.
- Learn the distribution of queries in plans, plan contexts, etc. It can learn which queries to expect for tasks and what common answers are. For example, the agent can learn that for fetch-and-place plans it always has to search for the appropriate place to put down the object at the destination. It could also learn that for the table setting task, the destination is very often the kitchen table.
- Learn to predict under which conditions which queries will be asked in the course of a particular task. Knowing when queries will be asked allows the system to prepare for computing the query in advance, for example, by precomputing the parts of the answer that only depend on knowledge that is already known. This reduces the time a robot is waiting for an answer and doing nothing meanwhile.

Knowledge such as from the examples given above can be exploited to optimize reasoning and control. We will enable agents to more efficiently complete their inference tasks by analyzing the perception, reasoning, and control tasks of generic plans in their respective activity contexts and tailor the knowledge bases and inference mechanisms accordingly. As suggested by Horswill (1994), architectures work better by specializing in what they have to do, especially given that many other situations of the general case never actually come up.

The exploitation of the structural characteristics in everyday activities is a key research goal of EASE. We aim to construct models that enable artificial agents to represent, use, and reason about their knowledge in connection to perception and executions without delaying execution, despite the fact that in their general formulation these inference tasks are often unsolvable, undecidable, or computationally intractable.

Our hypothesis is that we can find mappings from the problem space of the inference task into a collection of other problem spaces that allow for more efficient solutions by exploiting inherent structures of everyday activity. We call these mappings PEAMs. PEAMs will be investigated at the low- and sub-symbolic level (e.g., reaching trajectories) as well as the abstract, symbolic level (e.g., semantic environment). They will be used to reformulate perception, reasoning, and control problems in order to improve performance.

Identifying PEAMs from NEEMs

PEAMs are to be detected, analyzed and learned from NEEMs by researchers as well as by the robotic agents themselves. The NEEM datasets are used to analyze the subtasks associated with a task and translate general reasoning methods into task-specific combinations of specialized inference mechanisms. If the specialized methods for exploiting PEAMs fail, the robot should detect it and automatically apply the available more general reasoning methods and use the experience to make the action skill more robust.

Having many special-purpose reasoners reflects the view of Minsky (1986): “What magical trick makes us intelligent? The trick is that there is no trick. The power of intelligence stems from our vast diversity, not from any single, perfect principle.”

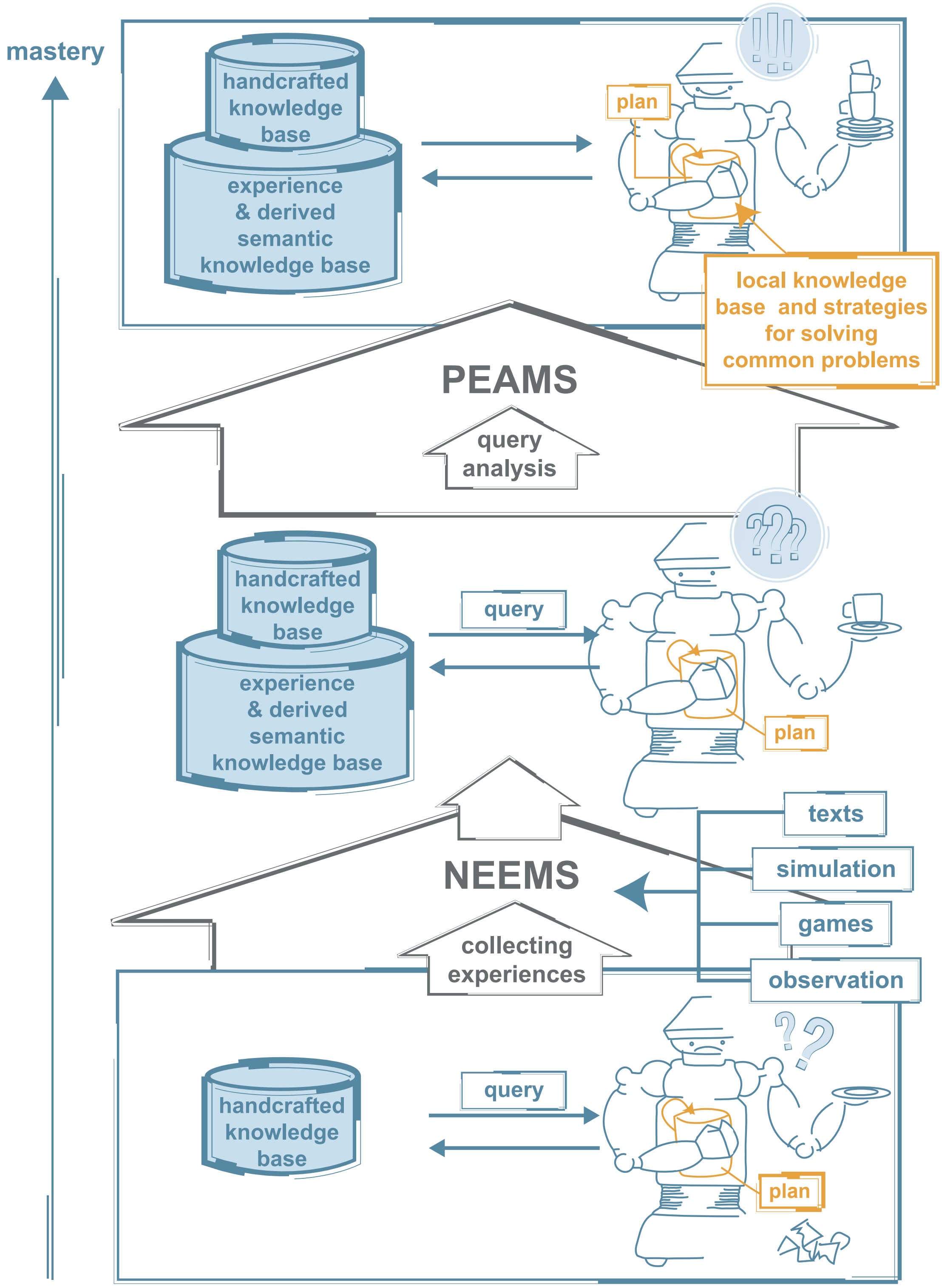


Figure 7: The figure shows the role of NEEMs and PEAMs in transforming knowledge-poor robotic agents into ones that can master everyday activity. In the first step, NEEMs are generated from collecting experiences and other sources (such as reading or simulating), which enables the robotic agent to complement the handcrafted knowledge with commonsense and naive physics knowledge obtained from NEEMs. In the second step, the PEAMs hidden in the structure of everyday activities are detected and exploited in order to equip the control system with fast query-specific reasoners.

The hypothesis underlying the EASE research project is that the investigation and usage of NEEMs and PEAMs will substantially advance the capability of the agent to master everyday activity. NEEMs are designed with the aim of acquiring huge commonsense and naive physics knowledge bases that are grounded into the perception and action systems of agents. PEAMs facilitate the optimization of reasoning methods by detecting and exploiting the structures and regularities underlying reasoning tasks in everyday activities. The concepts of NEEMs and PEAMs will form the foundations of a novel information processing framework that will be embodied into robotic agents to achieve the mastery of everyday activities.

Evolution of mastery How the research on the tasks outlined above is related to achieving mastery in everyday activity is depicted in Figure 7. We start with robotic agents that are equipped with generic action plans and knowledge bases that are carefully hand-coded by experts. The robotic agents then complement the hand-coded knowledge bases with additional commonsense and naive physics knowledge that is learned from collections of NEEMs. NEEMs are studied and analyzed to identify PEAMs that can be used to optimize performance on the (expected) tasks within their respective contexts.

The information processing models that result from our investigations of everyday activities, NEEMs, and PEAMs will be embodied into robotic agents. In EASE we consider embodiments to be the integration of the information processing principles with the perception-guided activity that changes the state of the world. The role of embodiment for robust activity has been pointed out more than a decade ago (Brooks, 1999; Pfeifer & Scheier, 2001). EASE goes beyond this line of research in that it combines the existing ideas of behavior-based control with the requirements for accomplishing various **human-scale activities** and using rich background and context knowledge (both symbolic and subsymbolic) for achieving competence in everyday activity.

1.2.4 EASE evaluation and impact

This section considers the measures that will be used to evaluate the results of EASE and the impacts EASE aims to achieve in terms of scientific contributions as well as broader impacts.

Measures of success The success of the EASE CRC will be measured along three dimensions:

1. The **outcome** of the EASE research activities. We consider the outcome to be the progress towards EASE's research goals of understanding and validating information processing models underlying everyday activity and the mastery of such activities by embodied robotic agents. These outcomes relate to the abilities and improvements of models and systems as a result of the research. The outcome includes but is not limited to the performance on the main EASE scenario (see Section 1.2.2 page 14). The criteria for evaluating the outcomes are described in more detail below.

2. The **output** that EASE generates. EASE aims to enhance our knowledge and understanding of the information processing principles underlying the mastery of everyday activity. Thus, one dimension of evaluation is formed by how much progress was made in terms of publications, their citations, and the standing of the publishing venues. Besides the publications and presentations, key outputs will be the open data and open software that EASE intends to produce and make openly available. The software will provide new methods and approaches for challenging problems in reasoning, motion planning, and vision that are unique and unprecedented.

3. The **impacts** of the EASE research activities. The impact of EASE will be evaluated as the effects that EASE has on the transformation of robot programming. We predict that the manipulation tasks autonomous robots have to perform will become increasingly complex, the environments more open, and the programming necessarily more knowledge-intensive. To be able to meet these demands, EASE is foreseen to play a key role in meeting this demand. EASE also targets wider impacts including promoting the (research) environment at UB and Bremen as a whole, furthering training and education through its graduate school and beyond, involving the general public, and advancing the promotion of young researchers.

The outcome of EASE is evaluated in more detail according to these five success criteria:

- **Outcome Criterion 1:** flexibility, robustness, efficiency, and predictability of the behavior generated by the robotic agents. The first measure of outcome success is the performance of the autonomous robotic agents realized by EASE on the main scenario (see Page 14). This includes the perception and manipulation skills that are needed for performing these tasks. Flexibility is defined as the ability to adapt at execution time; it reflects the ability of the robot to make use of opportunities and look for alternatives if the planned course of action cannot be executed. Robustness refers to the ability to detect/diagnose failures and recover from them. Efficiency is defined as accomplishing goals without large delays due to processing or processes interfering with task execution. Finally, a high predictability/low entropy in behavior is also desired for agents acting in human environments. Tangible progress will be provided in the form of open experiments that will be publicly available on the OPENEASE research platform (described in more detail at the end of the section and at the end of Section 1.2.8).

Flexibility

Robustness

Efficiency

Predictability

- **Outcome Criterion 2:** generalizability of everyday activity plans. The second success measure evaluates the quality of the plans that are used to generate the behavior measured by Criterion 1. The plans are assessed in terms of generalizability, ability to cope with underspecification, and ability to specialize themselves. Generalizability refers to being able to use the plans for different robots, tasks, and environments. Consider for example a plan for fetching and placing objects. Such a plan has to produce different behavior depending on whether it is called in the context of setting the table or loading the dishwasher. Different behavior is also necessary if the plans are employed by robotic agents that have different capabilities, depending on the perceptual capabilities and the dexterity of their manipulation apparatus. Another quality is how much vagueness in task formulation the system can deal with: the more specific task formulations are required to be, the more programming efforts are required for the plan, while a plan that tolerates more vagueness is easier to specify and tends to be more flexible. Finally, we consider plans that can specialize themselves in order to automatically improve the robustness, flexibility, and efficiency of the robotic agents to be superior.

Generalizability

Coping with underspecification

Improvement through specialization

- **Outcome Criterion 3:** knowledge content and usability. A key characteristic of everyday activity is the amount of knowledge that is needed for mastering it. Therefore, measuring how much relevant knowledge a robotic agent has available is a natural criterion for measuring its capability for mastering everyday activity. The knowledge content can be measured through the set of queries that a robotic agent is capable of answering. We will develop catalogs of queries that can be used to objectively measure the knowledge of robotic agents, which will include (but will not be limited to) queries that test the following capabilities:

Knowledge content

  1. Competent interpretation of vaguely, ambiguously, and incompletely formulated tasks.
  2. Successful answering of queries regarding what the robot has done, how, and why, and what it is capable of accomplishing.
  3. The robot's ability to answer queries about its operation environment.

4. The ability to understand scenes and form memories and models of the environment.
5. The ability to answer queries about the expected consequences of actions depending on the action parametrizations and the contexts they are executed in.

Any control program for robots mastering everyday activity must provide answers to these interpretation tasks and queries, be they hard-coded in the program or in the form of knowledge accessible to the program. The latter approach is in line with the view of Brachman (2002), who characterizes cognitive computer systems as “systems that know what they are doing”.

Generalizable knowledge

A secondary criterion is the generalizability of the agent’s knowledge: how much of the knowledge can be used by other robots for other tasks and environments? Generalizability includes the transfer of knowledge from one kitchen environment to another one but also from one application field to another. For example, from household to manufacturing and rescue applications. Finally, we will consider the expected performance gain of this knowledge.

Knowledge acquisition abilities

- **Outcome Criterion 4:** amount of knowledge that a robotic agent can autonomously acquire through performing long-term everyday activity. Another criterion will be the ability of agents to extend their body of commonsense and naive physics knowledge and improve their performance of everyday tasks through long-term or life-long activity. This will be done by comparing the question answering capabilities and task performance before and after knowledge acquisition from NEEMs.

Reasoning performance

- **Outcome Criterion 5:** performance gains in reasoning and perception tasks through PEAMs. The execution of tasks without delay due to inference tasks is a key goal of EASE and PEAMs are a key concept for achieving this. Therefore, we will assess the performance gains in the quality of solutions on perception and reasoning tasks as well as the computation time needed.

Scientific contributions The research contributions of EASE will be as follows:

- **Methods** for acquiring, interpreting, and analyzing data about everyday activity from different sources and forming NEEMs from this data. Sources include humans performing everyday activity, robots performing everyday activity, text sources, and simulation.
- **Comprehensive collection of NEEMs** on performing everyday activities. This collection contains information about how people use context-specific, implicit knowledge for mastering everyday activities, and how robotic agents have interpreted and executed these tasks.
- **Formalizations** of NEEMs, narratives, background knowledge, and plans and their properties as foundations of mastering everyday activities.
- **Methods** for abstracting NEEMs to generalizable knowledge and store these in comprehensive **knowledge bases**.
- **Framework** for NEEM-based, cognition-enabled robot control that can use the knowledge to satisfy the information need for mastering everyday activities.
- **Reformulations** of information processing problems in everyday activities into computationally more feasible ones through the exploitation of PEAMs.
- **Cognitive mechanisms**, including perception, knowledge processing, temporal projection[10], and transformational learning and planning[11] that are tailored to using these PEAMs.

[10] temporal projection in AI planning is the computational problem of predicting what will happen as a result of a robot executing its plan.

[11] transformational in the sense that general plans in a particular context and environment are transformed in order to improve their performance using various reasoning techniques and the knowledge from collected experience.

- **Robotic agents** that are capable of mastering long-term everyday activities and that learn to improve their performance with experience.

EASE broader impacts In addition to research impact, EASE aims and is expected to contribute to the achievement of specific, desired societal outcomes. This includes the impact of its research on important technological developments, creating a sustainable infrastructure for research and education, the participation of underrepresented groups, promoting learning in young pupils, and the promotion of scientific and technological understanding.

Broader Impact 1: Quality of life Due to the ongoing demographic changes in the European society there will be an increasing imbalance between the number of care givers and those who need care. Recent statistics estimate that within the 21st century, one third of Europe's population will at some point in their life be affected by brain-related diseases (Human Brain Project Team, 2012, p. 17). Many of these people will have problems in accomplishing their everyday tasks, which is strongly related to how independently they will be able to live. EASE will substantially contribute towards better Quality of Life Technology (QoLT) and Active and Assisted Living (AAL).

QoLT and AAL technologies are expected to improve lives in the large, growing subpopulation of people with reduced functional capabilities due to aging and disability. The aim is to develop systems that can *monitor* and *communicate* with the person, *understand* the person's daily needs and tasks, and *provide* reliable and happily-accepted assistance by compensating and substituting for diminished capabilities.

The societal impact of everyday activity science and engineering will be along two dimensions. Firstly, the improvement of cognitive orthotics, software-based personal reminder systems for people with cognitive impairment, such as memory loss. EASE research increases our understanding of the nature of everyday tasks and the cognitive capabilities involved. EASE will substantially advance the knowledge that cognitive orthotics can be equipped with by achieving a better understanding of how people master their everyday activities, building generative computational models for everyday activity, and mining the structure and regularities of everyday activity. This can be used to improve the current methods for assessing the capabilities of humans with physical challenges and assist those with cognitive challenges.

Secondly, EASE will substantially advance the ability of robotic agents to assist and cooperate with people in everyday activity task contexts. Many people are not able to perform certain manipulation actions themselves and are dependent on caregivers. Future personal robots might be powerful tools for these people to retain or reclaim some of their independence. Such robots could increase living standards while being more cost effective than other alternatives.

Broader Impact 2: Furthering scientific and technological understanding EASE intends to organize broad dissemination activities, including the creation of targeted media content and participation in information events for the general public.

Communication to the general public, media, policy-makers, leaders, and industry will be maintained according to a strategic communication plan made in the early stages of Subproject Z. The plan will consider different kinds of media such as: websites and social media, enhanced by engaging animations and movies, technical and non-technical literature, participation in key public, technical and non-technical events, and press kits. We intend to involve Prof. Prendas (Invenio University, Costa Rica), director and founder of OLIVAfilms[12], to develop and execute a broad communication concept for EASE. Prendas is a distinguished expert in communication of scientific research targeted to different stakeholders and audiences. Key targeted measures will be: analysis of the communications environment in R&D robotics, determination of EASE's communications' goals, identification and characterization of target audiences, determination of

[12] `http://olivafilms.com/`

resources, identification of key messages, decisions on channels of communication, budget, execution and, evaluation and impact assessment.

EASE will also organize science shows and public talks in cooperation with the Universum Science Museum in Bremen (see Letter of Intent). The Universum is a popular, interactive Science Museum with many changing exhibitions that also organizes events to inform and generate interest in the general public for scientific and societal topics. Public events organized by EASE in cooperation with the Universum are expected to garner much general interest.

Broader Impact 3: Enhancing infrastructure for research and education The core of EASE's educational and training infrastructure will be formed by its Integrated Research Training Group (IRTG) (see Subproject MGK). The IRTG will organize biannual 5-day international spring/fall schools with courses taught by renowned lecturers and targeted at EASE's research areas. The schools will cover relevant topics in a combination of lectures and practical exercises. Successful examples of such schools that EASE 's principal investigators have organized in the past according to these principles include the Player Summer School on Cognitive Robotics (PSSCR) and the CoTeSys-ROS Fall School on Cognition-enabled Mobile Manipulation in 2010. Summer school courses will also count towards fulfilling course requirements of the EASE graduate school.

EASE will also enable its doctoral students and post-doctoral researchers to cooperate with and make research stays at leading research laboratories around the world. We intend to establish double degree programs at different levels of academic degrees including a doctoral degree with the University of Rome *La Sapienza* and the University of Toulouse. We also intend to exchange early stage and postdoctoral researchers with the Seoul National University (Center for Human-level Machine Learning) to be supported by the DAAD Genko program. We provide letters of intent for cooperation with EASE from leading institutes, such as the Robotics Institute at Carnegie Mellon University in Pittsburgh and the Graduate School of Information Science and Technology in Tokyo.

In addition to raising the quality of teaching and training for the students and young researchers, EASE will provide excellent hardware infrastructure. Leading-edge research laboratories with modern sensing, robot, and computing infrastructure will be available for research and teaching. The facilities are detailed in Section 1.2.10.

Broader Impact 4: Participation of underrepresented groups Anticipated doctoral students of the EASE graduate school will include foreign students with DAAD and government scholarships from countries including Bangladesh (Feroz Siddiky), Mexico (Lisset Y. Salinas Pinacho), Costa Rica (Sebastian Chinchilla Gutierrez, Juan Carlos Saborio Morales), and Argentina (Ricardo Garro). Another doctoral student, Elias Dinter, who is suffering from severe multiple disabilities due to a *amyotrophic lateral sclerosis (ALS)*, has received support from a joint project of the Integrationsamt Bremen and UB to enable him to obtain a doctor degree in Assistance Robotics for people with disabilities.

Broader Impact 5: Promoting learning in young pupils In order to promote teaching and learning, EASE will offer educational material and training courses for senior highschool students and provide advice for teachers in the highschool partner program. For example, EASE intends to provide material to support senior students with entry-level information for research essays.

EASE will also contribute to programs for schools (science classes) at the Universum[13], the Science Center and Museum of Bremen.

We will continue the go4ITcampus[14] initiative, which has been successfully offered by the Institute for Artificial Intelligence. Its goal is to inform (prospective) students about what it means to study computer science. On the website the reviewers can see that the programs in past years

[13] `http://www.universum-bremen.de/`
[14] `http://www.go4it.uni-bremen.de/`

already featured courses for programming Lego Mindstorm robots and we will adjust the direction towards Cognitive Robotics. Videos of the results are available on the website as well. Through go4ITcampus we already have a large number of partner high schools and a well-established network of connections with teachers. The EASE activities in this respect will be managed by Sabine Veit, who has extensive experience in this area. She will also connect these activities to the supraregional SMILE project, which has a focus on girls and young women, their teachers and parents.

Dissemination with OPENEASE EASE will aim at achieving broader impact in the Robotics and Artificial Intelligence research community by extending and making its outputs openly available through a remote knowledge representation and processing service for researchers called OPENEASE[15] (Beetz *et al.*, 2015a). OPENEASE enables researchers to share experiment data with each other, as well as to share knowledge with artificial systems, and for artificial systems to share data as knowledge as well. The platform is built upon results from the European project ROBOEARTH[16], using a more advanced version of its knowledge system KNOWROB[17]. It contains tools to visualize, analyze and learn from (activity) data using a common platform. An impression of the website is given in Figure 8.

OPENEASE is operational and has been used for several cooperative projects, including ROBOHOW[18], ACAT[19], and SAPHARI[20]. The current developments of OPENEASE are partly funded by the EU projects ROBOHOW and ACAT. It is unique because of (1) the comprehensiveness with which real execution data can be stored and made openly accessible to the research community; (2) the representational infrastructure through which inhomogeneous experience data from different robots and even human manipulation episodes are semantically accessible in a uniform and standardized concept vocabulary; and (3) the suite of software tools that enable researchers and robots to interpret, analyze, visualize, and learn from the experience data.

OPENEASE will be a very useful platform for EASE to demonstrate and share its outputs with the research community in an easily accessible manner. EASE intends to share its data and methods using this platform and develop the extensions it needs as part of the data management plan of the project.

OPENEASE will be used for dissemination through the following:

1. **eLearning for AI-based Robotics.** OPENEASE is already being used as a tool for teaching a course in Intelligent Robotics at the University of Bremen. It enables students to explore the hardware of robots, their sensors and effectors, and get better intuitions about the data that sensors generate ("Can you detect the handles of cups using images where the camera is positioned at least 1.5m away from the cup?", or "Which objects or object parts in the kitchen environment cannot be detected with the Kinect sensor of the robot?"). In addition, the students do exercises with real robot data, such as learning object classifiers for the objects that stand on the kitchen counter during a set of manipulation episodes. EASE will expand upon the current tools and use the platform to teach and let students experiment with the data collected by EASE.

2. **Reproducing and extending experiments.** OPENEASE enhances the use of experimental data. For example if an experimental evaluation of a scientific publication has to be extended

Unique platform for reproducing recorded experiments

[15] OPENEASE stands for "open Everyday Science Activity and Engineering", and it is a remote service that gives its users access to experience data of robots and humans performing everyday manipulation tasks

[16] `http://roboearth.org/`

[17] `http:www.knowrob.org`

[18] `https://robohow.eu/`

[19] `http://www.acat-project.eu/`

[20] `http://www.saphari.eu/`

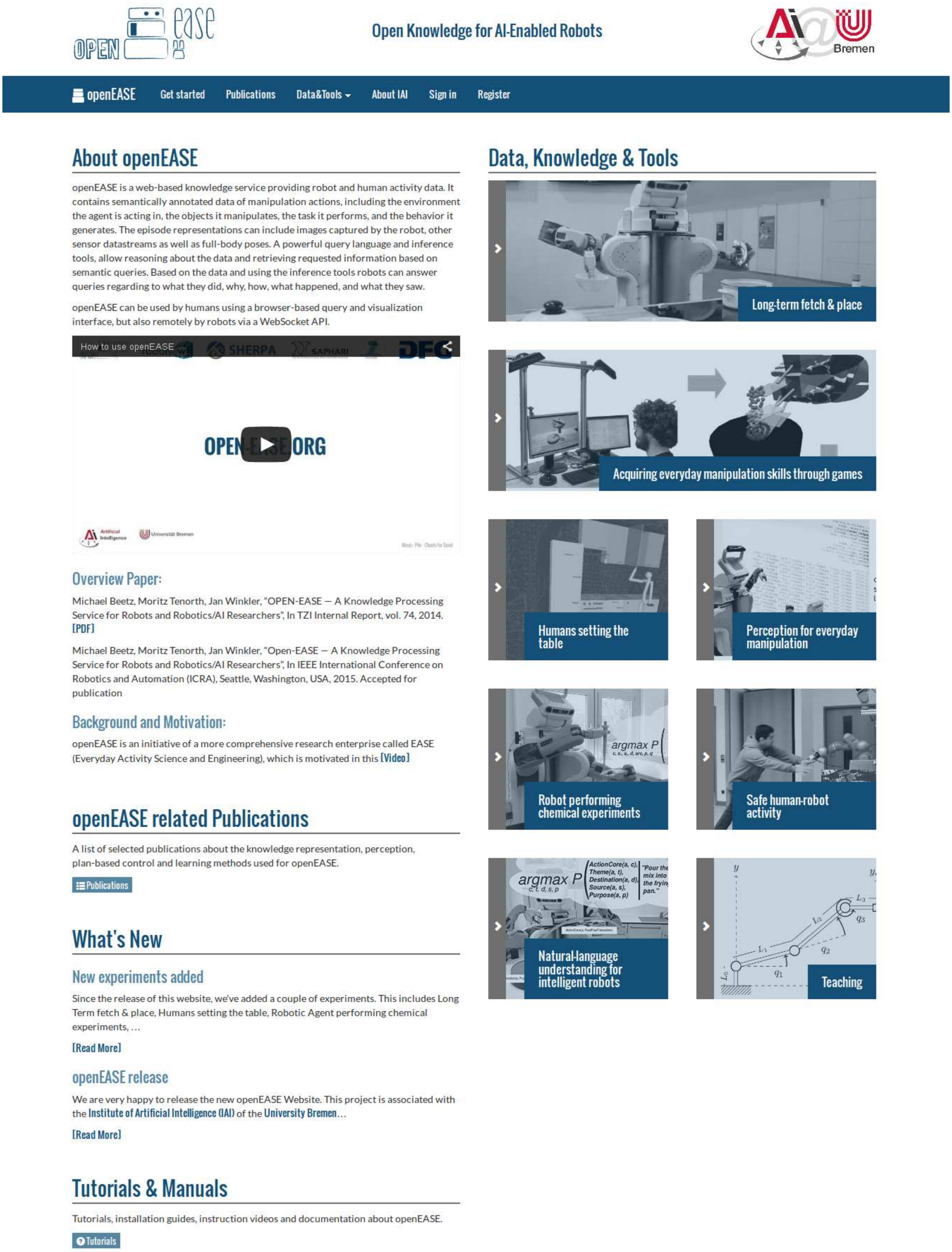


Figure 8: OPENEASE web page offering open knowledge bases of autonomous robots performing complex manipulation tasks as well as knowledge about human activities that can be used for robot learning.

after some time; rerunning experiments is tedious and time consuming, and requires a hardware setup that might not be available anymore. The comprehensive storage and semantic retrieval facilities support researchers in making additional analyses on existing experimental data. Researchers can also give reviewers and readers access to the experiment data through OPENEASE. This allows reviewers to assess the experiments in more detail and clarify questions regarding the experimental setting, such as where the robot stood when the object recognition mechanisms succeeded, or in which scenes an object could not be recognized. Making robot and human experiments reproducible through such a platform is unprecedented worldwide.

Advancing Robotics through sharing data and tools

3. **Open Robotics research.** Recently, progress in many fields has been fueled by making large volumes of data and corresponding analytics tools openly available. This is in the spirit of Nielsen's vision of "Reinventing Discovery" (Nielsen, 2012), which promotes new ways of conducting research more effectively through the cooperation facilities provided by modern internet technology. Inspiring blueprints for web services that promote open research witnessed in other domains include the Allen Human Brain Atlas[21] (Hawrylycz *et al.*, 2012) and the HapMap project (Gibbs *et al.*, 2003), which enable networked science in human and animal brain anatomy and human genome research correspondingly. Similarly, EASE intends to support open research in everyday manipulation and perception tasks through OPENEASE.

   OPENEASE currently provides data from robotic agents performing fetch and place tasks in a kitchen environment, users demonstrating pancake making in a virtual reality game, and people setting the table and cleaning up (Tenorth *et al.*, 2009). We plan to include experience data from the everyday household activities performed in the EASE CRC, thereby enormously increasing the data available on the platform and the semantic information that can be derived from them. This will make OPENEASE the most comprehensive and detailed activity knowledge bases relevant for autonomous robotics research in the world.

4. **Creating realistic benchmark problems for Machine Learning and Robot Perception.** OPENEASE will also be used to create realistic benchmark datasets for everyday activity. For example if one wants to test a newly developed robot perception method on a realistic set of perception tasks, one can take characteristic everyday activities and query OPENEASE for the set of perception tasks that a robot issues to perform such an activity. This informs the user which types of perception tasks are important for this task and which ones are not. If needed, the user can assert additional knowledge or correct knowledge in the knowledge base. Finally, the user can create realistic situations in which the perception tasks are to be performed.

5. **Tool for assessing how realistic action representations and and action reasoning methods are.** Most knowledge representation languages and methods for symbolically reasoning about actions and change are based on modeling assumptions. Using OPENEASE gives researchers in these fields the opportunity to learn action models from real-world experimental data and compare them to their proposed action models. This way the researchers can test to what extent the assumptions underlying their action models are valid for autonomous manipulation robots, and to what extent the inferences performed by these formalisms are valid with respect to the behavior and the physical effects that robotic agents generate.

[21] `http://www.brain-map.org/`

The EASE consortium is a strong supporter of **open research**. In this spirit we wish to make the progress of EASE as transparent and reproducible as possible. EASE will record its long-term experiments and make them publicly available online on `www.open-ease.org` (Beetz *et al.*, 2015a). Through this platform, the experiment conditions, experiment data, knowledge bases, and performance will be openly available online. The data will include image streams, complete robot motion logs, and poses of (perceived) relevant objects, as well as the corresponding semantic information.

OPENEASE extensions by EASE The development of OPENEASE is not one of the main targets of EASE. The main development of OPENEASE, including infrastructure work and extensions unrelated to EASE subprojects, is expected to take place in a companion project. The knowledge gained from EASE regarding the representation of large datasets and methods for analyzing them are expected to be useful for OPENEASE however. Also, the addition of the large amounts of EASE everyday activity data and reasoning methods will help OPENEASE mature. Finally, individual EASE subprojects are expected to develop their own tailored OPENEASE presence where appropriate.

EASE intends to develop subproject-specific interfaces, knowledge representation concepts, reasoning, and analysis tools for OPENEASE. For example, Subproject H01 will conduct manipulation experiments with physically impossible objects in a virtual environment, which requires extensions to the knowledge representation as well as the visualization capabilities of OPENEASE. Subproject H03 will collect manipulation activity data, together with brain signals and think-aloud protocols. Again, OPENEASE needs to be extended to deal with the multiple modalities of experiment data and requires software tools to support the interactive analysis of these data.

1.2.5 Core concept 1: Narrative-enabled episodic memories (NEEMs)

Learning from past experiences

The first core concept of EASE is concerned with the structure and organization of knowledge, such that it facilitates knowledge acquisition and cognitive reasoning mechanisms. For humans, episodic memories play a key role in the knowledge acquisition process. Episodic memory is concerned with the recollection of activities and events that are embedded in experience, a particular time, place, and context: we can see a goal of the last soccer world championship with our *mind's eye* or tell somebody about a cross court tennis hit and they could recall how that would feel in the arm from personal experience. This history of past experiences that people are able to draw upon serves as a resource for better informed future decision making and learning. It provides the basis for mental re-experiencing and imagination of past events and hypothetical situations. The essential role episodic memory plays in everyday activity is, for example, illustrated by case studies of patients with hippocampal amnesia. Studies show that patients with severe episodic memory impairment but intact semantic memory also have difficulties imagining themselves in the future or imagining new experiences (Hassabis *et al.*, 2007). The ability for prospection and *mental time travel*, for imagining future events or needs based on past experiences is essential to our higher cognitive functioning and may depend largely on episodic memory (Suddendorf & Corballis, 2007; Vernon *et al.*, 2015) as well.

Extracting general knowledge (semantic memory) from episodic memories

Moreover, it is commonly believed that semantic memory is derived from collections of episodic memories. In other words, humans acquire general knowledge by first storing the episodes in which we are exposed to certain information. Then, through systems consolidation the episodic memories may be transformed to semantic memory, which contains general knowledge and facts that are no longer connected to a specific event or experience. Though it remains unknown how episodic and semantic memory in the brain interact exactly, there is no doubt that episodic learning and memory plays a pivotal role in human cognitive functioning.

Building models for actions and capabilities from NEEMs

Artificial agents can use experiences stored in an episodic memory-like system to build realistic models of their actions and capabilities by recording successes and failures, remember where things are, form expectations, and so on. Selected components of experiences can be used for supervised learning, experiences can be replayed to aid learning, and information from experiences can be re-evaluated in the light of later findings. In addition, a lot of commonsense knowledge concerning the regularities of everyday activity and naive physics knowledge is implicitly available in these experiences, which, when made explicit, can be used to aid the agents in mastering everyday activity.

Structuring knowledge using narratives

We propose to structure experiences and acquire knowledge in a fashion that is inspired by the human episodic memory system and closely related to humans' remarkable ability to process, interpret and communicate information in the form of stories or narratives (Anderson, 2015a). EASE proposes to collect knowledge in the form of episodic memories and enrich them with *narrative intelligence*. *Narrative intelligence* being the human ability to organize experience into narrative form (Blair & Meyer, 1997) in order to make sense of the world. Humans structure actions and activities and understand the rationale behind them by assimilating them to narratives.

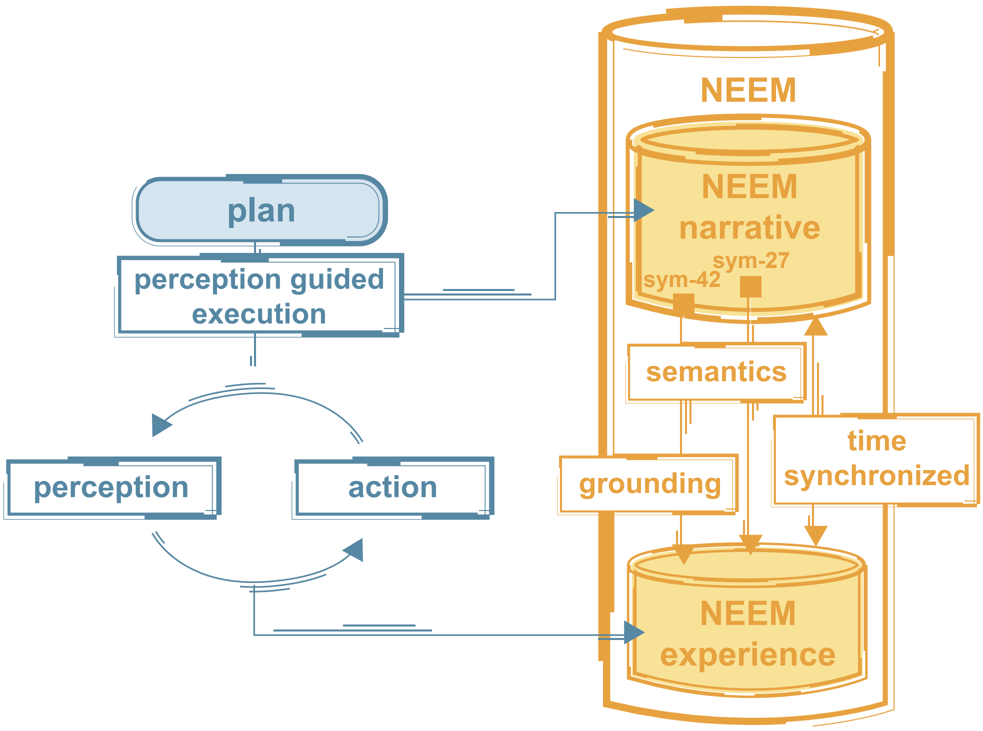


Figure 9: The figure shows how robotic agents can automatically generate NEEMs during plan execution. The perception-action loop of the robot generates subsymbolic data that constitutes the NEEM experience. At the same time, the plan interpreter generates symbolic annotations — the NEEM narrative — that describes the subsymbolic data in a time synchronized manner by grounding the experience and .

These structures in which experiences of manipulation activities are collected and enriched by narratives, are called **narrative-enabled episodic memories (NEEMs)**. Figure 9 shows how the NEEM system is structured and relates to the agent. The two constituent parts of NEEMs are described below.

1. The **NEEM experience** contains the raw data of an experience and can include, for instance, image streams that an agent perceives when performing the activity, detailed full body motions

of the acting agent, as well as sensations, such as the expended force or contact events. The experience data enables the agent to reconstruct the agent and world states during the episode. An agent's own NEEM experience is complete in the sense that any data structure that the agent used to decide on the course of activity, the parametrization of its actions, and other relevant information that has effect on the activity (such as detected failures) are stored as a named entity in the NEEM experience such they can be referred to by the NEEM narrative. The completeness implies that the value of any abstract relation that the agent uses to construct the NEEM narrative can be computed from the data structures of the NEEM experience.

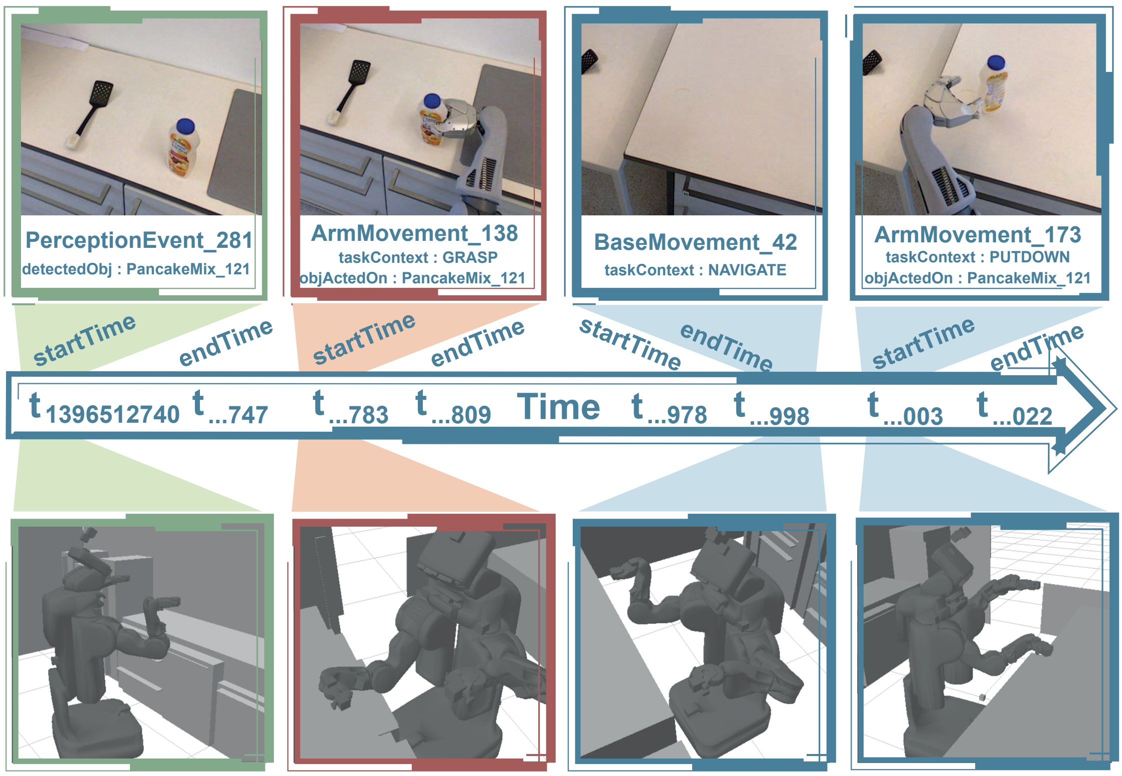
2. The **NEEM narrative** is a symbolic representation that structures an agent's activity (e.g., "set a table") into a hierarchy of subactivities ("add the cutlery"), actions ("fetch a fork"), and action phases ("grasping the fork"). It also represents the reasoning and information processing mechanisms that generated the activity in terms of concepts such as beliefs (where the robot saw a particular object), tasks (the intention to perform a particular action), context (the known conditions under which actions are executed), and behavior and effects that are generated by invoking and parametrizing control programs. Narratives complement the NEEM experiences. NEEM narratives can be automatically generated during plan execution by the plan interpreter by identifying, naming, linking, and annotating the data structures in the NEEM experience according to the semantics of the plan language (see Figure 9). The NEEM narrative turns the agent into a *cognitive* agent that knows what it did, why, how, how well, etc. and can anchor the abstract entities of its narrative into the continuous data structure of the NEEM experience.

NEEMs that are generated from observing the activities of other agents or from reading instructions can lack some input channels of experience and suffer from incomplete, vague, and noisy representations caused by perspective displacements, textual ambiguities, etc. For example, when an agent observes an activity performed by another agent, the observing agent does not have access to the intentions and beliefs of the acting agent, and when an agent reads instructions, it cannot see the motions that the respective action requires. In the latter case an agent can use NEEMs generated from its own activities in order to imagine how an action is being executed. The agents will have access to a large database of NEEMs from different (complementary) sources, through which they can derive the components of the knowledge they require. At the end of phase 1, an estimated 5,000 robot days worth of NEEMs will have been collected from execution, simulation and text.

One of the goals of EASE is to realize, test and extend the NEEM memory system and develop the software tools for querying, analyzing, and learning from NEEMs. This will provide robotic agents with the following essential information processing capabilities:

- **Replaying experience**. The robotic agent can use the NEEM system to retrieve subepisodes in which the robot performed certain actions and replay it. There are two basic retrieval operations: (1) retrieval of a situation, i.e. a snapshot of the experience at a given time instant, and (2) replay of the experience over time for a specified interval. For example, the robotic agent can retrieve episodes in which it poured pancake mix into the pan that resulted in a round pancake of a specified size. It can replay the trajectories of the bottle that was the source of the pancake mix, the forces it *felt* during the pouring action, or the stream of images it receives when monitoring the success of the action.

- **Answering queries**. Using the ability to replay experiences and by semantically indexing scenes through NEEMs, the robotic agent can selectively retrieve information from experiences that are relevant and use them in its decision-making process. For example, if the robot



assertion

```
occurs('PerceptionEvent_281', t2).
category('PerceptionEvent_281', 'ObjectDetected').
occurs('ArmMovement_138', [t6,t8]).
task_type('ArmMovement_138', 'Grasping').
object_acted_on('ArmMovement_138', 'PancakeMix_121').
category('PancakeMix_121', 'PancakeMix').
...
```

explanation

The facts above assert the occurrence of an event with the name `ArmMovement_138` of the type `Grasping` in the time interval `[t6,t8]`. The grasping action is performed on an object that was named `PancakeMix_121` by the perception system and classified as a pancake mix. This object hypothesis was generated by a perception event named `PerceptionEvent_281`.

Figure 10: Visualization of the representation of a NEEM subepisode during which a grasping action was performed and its links to subsymbolic data, including RGB-D images and continuous robot motion trajectories. The subsymbolic data and the symbolic knowledge structure create a timeline using universal timestamps. Time is structured into time intervals and time instants (intervals with duration 0). Changing relations are asserted using the predicate `holds(f,i)`, meaning the time a variable relation `f` holds in the interval `i`. This is used to assert the occurring events, `occurs(ev,i)`, and beliefs of the robot, `belief_at(f,i)`. The logical representation also includes unique names for data structures in the control system such as captured images and robot poses. The use of the logical representation for answering queries is described later in this section.

is setting the table and has to put white bowls onto the table, it can ask which perception algorithms in the past successfully detected white bowls and query images that were collected in the process of picking up a bowl. Results of an image query regarding a grasping action are shown in Figure 10. NEEM experiences and narratives can also be used to parametrizing manipulation actions. They can answer which parameters and events were involved in spilling a liquid, what the success/failure statistics of different ways of picking up similar objects are, how pouring pancake mix is different from pouring water, etc.

A common reasoning pattern for answering such queries is to first reconstruct the respective scene and then perform more complex reasoning tasks on it to derive new knowledge.

- **Learning from experience and deriving commonsense knowledge.** Supervised learning methods can be applied to the NEEMs for the robotic agents to learn concept definitions in terms of grounded features from semantically described scenes and episodes. Specifically, the agents can learn the action parametrizations that are expected to succeed in accomplishing vague action descriptions in generic plans, and thereby improve the mastery of these actions. For example, the robots can learn where to stand in order to pick up objects successfully, which grasps to apply, whether to use one or two hands, and so on (Fedrizzi, 2010).

  A vast amount of commonsense and naive physics knowledge is implicitly present in NEEMs. In order to make this knowledge explicitly available, EASE will consider the unstructured information management approach. The research resulting in the Watson system (Ferrucci *et al.*, 2010) has proven successful in learning to answer an open set of queries. It does so by hypothesizing many possible answers and assessing the likelihood of each hypothetical answer to be the correct one by collecting and applying statistics over knowledge patterns. EASE proposes a hypothesis along similar lines: that there is a way of acquiring commonsense knowledge through statistics generated from collections of NEEMs, without requiring the knowledge to be available explicitly.

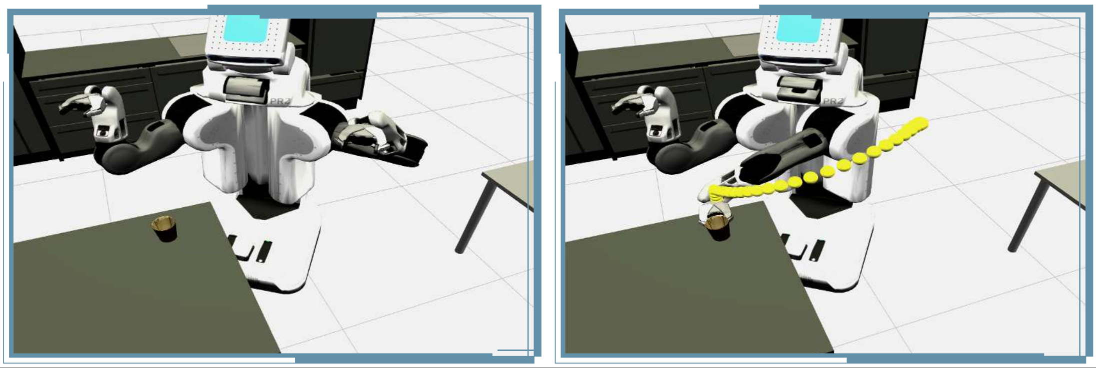
By using NEEMs, EASE intends to form repositories of established solutions for specific tasks, together with indications of the boundaries of their applicability. EASE intends to develop *narrative-enabled agents*, i.e. agents that are equipped with the information processing infrastructure to:

- interpret plans such that they generate behavior as well as store the experience as a NEEM;
- interpret and store activities that they observe, simulate or read about as NEEMs as well;
- maintain and manage large bodies of NEEMs from which narratives can be formed;
- reason about NEEMs for answering queries;
- learn from and transform NEEMs into general declarative bodies of commonsense and naive physics knowledge and;
- query the derived knowledge during the reasoning and execution processes to achieve mastery in everyday activity tasks.

Chronicle representation

Representational structure of NEEMs One convenient representational view on NEEMs is that of first-order time interval logic (McDermott, 1982, 1985; Allen, 1983), sometimes also referred to as a *chronicle representation* (Ghallab, 1996). *Time intervals* are specified through *time instants* at their start and end. *Events*, such as a reaching motion, *occur* over time intervals. Actions are considered to be events that are caused by an agent to achieve some goal. *Occasions hold* over a certain *time interval* and represent, for instance, the state of an object such as its location. Instantaneous events and states occur at a *time instant* t_i, which is a time interval with the duration 0.

As an example, consider the time chronicle representation of a fetch-and-place task depicted in Figure 10. The table at the bottom shows example assertions stated in the representation



<table>
<tr><th colspan="2">KnowRob (Prolog) queries</th></tr>
<tr><td><pre>entity(Tsk, [an, action,
  [type, grasping_something],
  [object_acted_on, [an, object,
      [type, cup]]
  ]]]),
occurs(Tsk, [TI_Begin,_]),
holds(knowrob:'pose'(
  pr2:'pr2_base_link', Pose),
  TI_Begin).</pre></td><td><pre>entity(Tsk, [an, action,
  [type, grasping_something],
  [task_success, true],
  [body_parts_used, BodyPart]
]),
occurs(Tsk, T_Int),
show(trajectory(BodyPart),
  T_Int).</pre></td></tr>
<tr><th colspan="2">reformulation of the queries in natural language</th></tr>
<tr><td>Let <code>Tsk</code> be a task of the robot in which it intended to grasp an object of type cup and let <code>TI_Begin</code> be the time instant where this task started. Then infer the pose of the robots base <code>pr2:'pr2_base_link'</code> in global map coordinates at time instant <code>TI_Begin</code>.</td><td>Let <code>Tsk</code> be a task of the robot in which it successfully grasped an object, which occurred in the time interval <code>T_Int</code>. Display the trajectory of the gripper that was used to pick up this object throughout the time interval <code>T_Int</code>.</td></tr>
</table>

Figure 11: Visual results of queries on logged robot experiences.

language, including actions and events that occur during the episode, the beliefs of the robot that hold at certain time instants, etc.

Grounding symbolic expressions in low-level experiences

The values of these assertions are not explicitly stated in the symbolic knowledge base of the robotic agent. Rather, they are computed through logical rules from the data structures that the plan interpreter generates when executing a plan. For example, the assertion

`belief_at(robot('pr2_base','Pose_423'),`$t7$`)`

is computed by retrieving the robot pose at the given time instance from the experience data of the episode. This approach has an important advantage over explicitly asserting symbolic expressions as abstract facts to the knowledge base. During plan execution, the plan interpreter uses the NEEM infrastructure and generates, connects, and labels the subsymbolic data structures narratives are built upon. Thus, the symbolic expressions are automatically grounded in the low-level experience data. This allows to include procedural attachments to look up knowledge from the NEEM experience, which results in a strong link between the symbolic knowledge representation infrastructure and the robot's perception-action domain. That enables us to easily access data structures, such as images, image features, motion forces, and robot poses through logical rules. As all the knowledge concerning plan execution is already contained in the NEEMs, there is no need to explicitly assert that information into a knowledge base: it can be automatically inferred through the procedural attachments instead.

Inferring knowledge from experience through procedural attachments

Such rules have two purposes. First, they translate subsymbolic data structures into symbolic relations. Second, they link the data structures to concepts in the ontological knowledge base of the robot. This way the robotic agent can apply its background knowledge contained in the encyclopedic knowledge base to the data structures produced by the plan interpreter.

Consistent, detailed representation of the robot's perspective

Maintaining a symbolic knowledge base that is physically grounded in the logged data structures of the control system has additional significant benefits. First, the symbolic knowledge base is always consistent with what the robot was seeing and doing. Second, the logged data structures are the most detailed representation about the state of the robot and the world from the robot's perspective. Different predicates can abstract these data individually according to their purpose. While one predicate might be asserted over the exact pose of the robot, another predicate might use an abstract location such as *in the kitchen*. In other words, the same data can be used to infer different symbolic knowledge. Finally, the interpretation of the events and the inference of the observable effects can be done using the plan interpreter rather than through symbolic reasoning over abstract models of actions, plans, and their effects. For example, instead of having defined symbolically a method for determining whether an object is reachable, one can use NEEMs to learn statistics over the states in which a reaching action tends to be successful.

NEEMs support predicates operating on different levels of abstraction

The first-order time interval logic representation allows us to ask sophisticated queries that combine information from logical facts with continuous and geometric data such as robot poses in three-dimensional space, as depicted in Figure 11.

Types of reasoning supported by NEEMs

In summary, the NEEM representation will be defined along a set of predicates and function symbols that is expressive enough to enable the robotic agent to reason about (1) the robot's hardware, its capabilities, its environment and the objects it manipulates; (2) what the robot saw, reasoned, and did, how it did it, why, and what effects it caused; (3) and the skills demonstrated by humans in training episodes that the robot can learn from.

We invite the reviewers to try out a prototype of such NEEM databases and test the queries shown in Figure 10 and Figure 11 (formalized in KnowRob (Tenorth & Beetz, 2013)) through the graphical, interactive web knowledge service provided at `http://open-ease.org` where a rich set of different manipulation epsiodes can be selected, and `https://data.open-ease.org/ease-review` where one of the manipulation episodes is auto-loaded for faster access.

Sources of NEEMs We have identified five sources of interest for NEEMs:

- **Experiences**, i.e. records of activities from the point of view of the performing agent. In this case the NEEMs include goals, sensory experiences, images, effect experiences, expectations etc. (Subprojects R01, R04, and R05).

- **Observed activities**, i.e. records of activities from observing the performing agent. In this case, the intentions and beliefs of the executing agent are not observable. The advantage is that the activities can be performed by humans, who are more competent in mastering everyday activities. Observed activities are therefore promising sources for imitation learning (Subprojects H01 and H03).

- **Language descriptions**, in particular instructions, from which narratives can be derived that often contain more abstract and general information about how to perform everyday activities, going beyond individual examples. Language descriptions enable learning by being told, but usually presume a basis of common knowledge (Subprojects H02 and P01).

- **Simulations**, i.e. records of activities in a simulated environment from the point of view of the performing agent. This enables agents to perform actions with different action parametrizations and thereby learn about naive physics relevant for particular actions. In particular, the agent can qualitatively learn about the causal relationships between action parametrizations and their physical effects. This also corresponds to the idea of *dreaming* to process new information (Subprojects R03 and P01).
- **Games with a purpose**: in this setting we generate knowledge acquisition tasks as game episodes that require competence in everyday activity in a virtual environment. Though this source has similarities to learning from observed activities, the setup and obtained information is very different. The unique opportunity that games offer is very fast, crowd-sourced, active learning: the robot can make up learning tasks and phrase them as if learning with a teacher (Subproject H02).

Different sources provide NEEMs with complementary types of information, which can be combined together in order to build up comprehensive everyday activity knowledge bases. The use of a common representational structure and a common ontology facilitates that.

Rationale behind NEEMs The reasons we believe that NEEMs will play a fundamental role in the acquisition of the knowledge needed for the mastery of everyday activities are as follows:

- In humans, episodic memories can form the basis of experience from which semantic memory (containing general knowledge and facts) can be derived. Moreover, episodic memory seems to be linked to the ability to imagine oneself in the future, or imagine new experiences (as discussed previously on page 34). This ability is essential in being able to predict future needs, experiences and events, and may therefore also play an important role in adapting behavior to the (predicted) situation.
- Research in cognitive sciences provides evidence that language and action share a common basis and might be based on a common grammar-like structure (Pastra & Aloimonos, 2011). EASE intends to take advantage of these relationships by hypothesizing that language (in particular, instructions and stories), actions, and episodic memories have the same structure and exploit this by choosing NEEMs as a fundamental knowledge representation structure.
- EASE hypothesizes that a lot of commonsense and naive physics knowledge is present in the NEEMs in the form of statistics, i.e. an approximate joint probability distribution. This theory can explain many observations:
  - "Why does it seem to be so easy for humans to acquire commonsense and naive physics knowledge?" Because they do so by learning statistics over NEEMs, rather than full knowledge of the world and all the laws of physics.
  - "Why do different people have so much overlap in their commonsense and naive physics knowledge?" Because everybody is performing, observing and talking about the same everyday activities.
  - "Why are humans so bad in explicitly stating their commonsense and naive physics knowledge rules and facts?" Because this knowledge is implicit and distributed over many statistics.
  - "Why are humans stereotypical in their everyday manipulation actions?" Because by acting stereotypically the entropy in the collected statistics decreases and the information content increases.

– “Why can commonsense reasoning facilitate such a broad range of reasoning techniques, including prediction, causal reasoning, diagnostic reasoning, intercausal reasoning, etc.?” The reason might be that if agents reason about a joint probability distribution then all these reasoning techniques can be considered as conditional probabilities that can be computed from the same joint probability distribution.

Using NEEMs brings several advantages over other data representation structures. We will discuss these next.

Advantages of NEEMs Main advantages of using NEEMs as a basis for knowledge bases of everyday activities are listed below.

Advantage 1: NEEMs can be constructed easily and quickly as a byproduct of plan interpretation. We can design a plan language and implement a plan interpreter such that a robotic agent can automatically generate and store NEEMs of the activities it performs without delaying task execution (Subproject R04). The NEEMs that such a plan interpreter can generate are:

- **comprehensive and complete**. NEEMs contain all data structures and information that was used in reasoning for and execution of a task. The data is semantically accessible in terms of objects, goals, actions, effects, behavior, etc. It also contains all low-level parameters particular to the execution. Thus in principle it contains all knowledge needed to execute the task again.
- **originating from multimodal input sources** including images, sensor streams (proprioceptive, touch), symbolic representations that result from the mental activity of plan interpretation, etc. Subsymbolic data structures are assigned unique names such that they can be referred to in the symbolic representations.
- **uniform in representation**: they use a common vocabulary of predicate and function names as well as a common ontology of concepts. Thus we can have a common representation across modalities: experiences, observations, reading, etc.
- **grounded in perception and action**: symbolic object descriptions are linked to the image regions that were used to generate and refine the symbolic descriptions. Similarly, symbolic action descriptions are linked to the stream of control signals generated for their execution and the corresponding feedback sensor data streams.

Building knowledge bases using NEEMs

Advantage 2: NEEMs enable investigation of novel and promising means for acquiring everyday activity knowledge bases. Collecting large sets of NEEMs consisting of experiences and narratives enables the robot to acquire knowledge bases of individual everyday activity experiences. The individual experiences comprehensively represent episodes at symbolic as well as subsymbolic levels. These levels are interlinked through semantic relations such as the goals, beliefs, actions, effects, behavior and intentions of the agent while carrying out the activity. This knowledge base of everyday activity experiences can be considered an artificial episodic memory system. In humans (as discussed on page 34), episodic memory plays a distinctive role in acquiring new knowledge.

Let us illustrate the potential of NEEMs for the acquisition of everyday activity knowledge bases using the learning of affordances as an example. Having collected a large set of everyday activity episodes, the robotic agent can learn concepts such as seats of chairs as the objects that people sit on. To do so, the agent can obtain training data for learning object affordances by asking NEEM queries such as the following:

- which actions were performed with such an object?

- which actions were performed successfully with these objects?
- how are actions with this object performed so as to be successful?
- what are typical failures in performing this action with this object?
- what are the causes of these failures?

Learning concepts such as affordances from experience can be considered a form of *knowledge compilation*. NEEMs in themselves are inadequate for answering general queries. To extract general knowledge, different learning methods can be applied to the experience episodes: data analytics, Unstructured Information Management, and Deep Learning are all used in the successful acquisition of knowledge bases that have the breadth and depth to enable agents to master everyday activities.

Advantage 3: NEEM narratives can be used as indexing schemes for the NEEM experiences as a result of the way the abstract symbolic representations are linked with the concrete subsymbolic ones in NEEMs. Research in the cognitive sciences indicates that superior human reasoning capabilities are supposedly realized through subsymbolic inference mechanisms such as the imagistic reasoning, cognitive inference through simulation, and learning by re-experiencing the past (Hesslow, 2012). In artificial agents, NEEMs

NEEM narratives as indexing schemes for NEEM experiences

- allow to activate/access subsymbolic knowledge (we can ask how it *felt* to pick up a heavy pot),
- form the basis for learning simulators of activities and parts thereof,
- can be considered as virtual knowledge bases of everyday activity knowledge,
- and allow application of data analytics methods that are successfully used in *web mining* to summarize news, detect super- and subclasses of concepts, and find new instances of relationships.

Advantage 4: NEEMs might serve as a basis for explaining some aspects of the nature of commonsense and naive physics knowledge in humans. The aspects that are addressed by NEEMs have already been discussed in the Rationale behind NEEMs section on page 41. The hypothesis is that the usage of symbolic representations as a means for retrieving subsymbolic data and the fact that data is not being represented explicitly, might be a reasonable explanation of why people possess commonsense knowledge but have problems in stating it explicitly. Thus, through researching NEEMs and how they can be constructed and applied, we may formulate new theories regarding commonsense and naive physics knowledge in humans.

Advantage 5: NEEM prototypes are promising. The usage of NEEM-like recordings and some of their advantages have been investigated and demonstrated (Winkler *et al.*, 2014), showing that we can have some confidence in the feasibility of the intended NEEM constructs. The scope and usability of the NEEMs developed in EASE will far exceed these first attempts in all aspects, such as the various sources of NEEMs, their content, their use-cases, their structure, etc.

Relations/synergies to selected theories in Cognitive Sciences The investigations of NEEMs to be conducted in EASE relate to and have synergies with various theories in Cognitive Science, Cognitive Psychology, and Cognitive Neuroscience. EASE is inspired by research from these fields and although EASE does not intend to directly extend these theories, we expect that the outputs of EASE can substantially contribute to questions and theories regarding the mechanisms behind acquisition and use of everyday activity knowledge, and the mastering of everyday activities.

Below we will briefly look at a selection of theories in cognitive fields that are particularly interesting for the research program of EASE.

Episodic memory and memory models in Cognitive Science The problem of knowledge acquisition and representation is impressively solved by the human memory system, which is intensively studied in Cognitive Science and Cognitive Neuroscience.

There are different theories and models explaining different observations, but there is no well-established unified theory of the Human Memory System. As Tulving (1995) stated: “Research in Cognitive Psychology and Neuropsychology of memory has produced a wealth of data. ... However, our success has been somewhat less remarkable in interpreting and making sense of this abundance of data. There is less agreement among practitioners as to what the findings and facts tell us about the larger picture of memory.” The foundation of the research on the organization and functional processes in the human memory system were largely informed by case studies of patients with localized brain lesions. For example, the famous case of patient H.M., whose hippocampi were removed bilaterally, resulting in a loss of the capability to form long-term memories of events experienced after the surgery while still remembering those events that happened before the surgery.

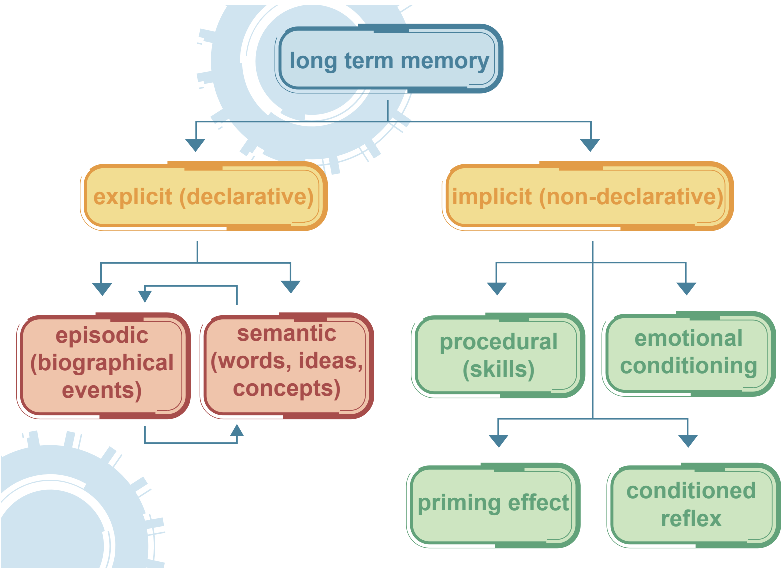


Figure 12: Tulving’s taxonomy of human memory systems.

A widely adopted structure of the brain into functional components, which is also most relevant for the EASE research program, is depicted in Figure 12. It distinguishes between the declarative and non-declarative components of the human memory. The declarative components are those that can be explicitly queried for answers while the non-declarative ones store implicit skills such as riding a bike. You can ride a bike without being able to explicitly explaining how you do it. You can activate the skill without noticeably attending to it. In contrast, the declarative memory can be explicitly queried and returns answers that can mostly be stated in language. The declarative memory itself is considered to consist of the episodic and the semantic memory. The role of the episodic memory is concerned with *remembering* experienced events, such as the final game of the last football world championship, while the semantic memory is concerned with *knowing* facts, such as the name of the player who scored the decisive goal.

Episodic memory (Tulving, 1972) enables humans to re-experience event sequences from their recent or remote past and play them out in in their minds as if they were reliving the experience. It stores significant events in one’s life, including images, sounds, smells, emotions. The memories are consciously and declaratively recollected, involve the specificity of time and place, and are rich in vivid detail. The episodic memory system also provides the information basis for learning other kinds of models such as prediction models for actions, knowledge about objects, environments, tasks etc.

EASE knowledge architecture is inspired by human episodic memory

EASE point of view: In EASE we adopt the hypothesis that the acquisition of everyday activity knowledge including commonsense and naive physics knowledge can greatly benefit from a knowledge representation and processing architecture inspired by the episodic memory system. The research of EASE will focus on information processing models that provide capabilities that correspond to some of the functionalities of the episodic and semantic memories in the human memory system. We expect to invest particular research efforts in the investigation of computational methods for compressing and representing NEEMs, learning from them, and efficiently reasoning with them.

Simulation theory of cognition The second field of Cognitive Science that suggests strong synergies with the EASE research program is the simulation theory of cognition. The simulation theory of cognitive function by Hesslow (2012) explains how an organism could simulate interaction with the external world. It is argued that such simulation explains the appearance of an inner or *mental* world in humans. The theory is based on three hypotheses. First, an action can be simulated by activating motor structures in the frontal lobe roughly as they would be activated during an overt action, except that the final motor output is suppressed. Second, perception of an external stimulus can be simulated by internally elicited activation of sensory cortex as it would have been activated during normal perception of that stimulus. Third, there is a simple anticipation mechanism such that early stages of both overt and covert actions can elicit perceptual simulation of their normal consequences before the action has been performed. The simulation approach incorporates the cognitive concepts such as working memory, imagination, thinking, cognitive maps and the concept of a goal and explains findings from experiments that address these phenomena. It also explains the relation between motor and cognitive functions, and provides a simple and plausible account of how evolution has added cognitive functions to more primary functions and behaviors.

The hypotheses are supported by a wealth of empirical data. For example, motor imagery studies show that many of the same neural structures are involved in imagining an action and executing an action (Jeannerod, 2001; Schnitzler *et al.*, 1997; Decety, 1996). Similarly, visual imagery studies show activation of similar structures whether an image is being seen or imagined (Kosslyn *et al.*, 1997). The same goes for auditory imagery (Kraemer *et al.*, 2005). In fact, there may not be that much difference between imagining an action or perception and executing that action or receiving that percept.

The use of simulation as a reasoning tool is also put forward for language understanding. For example, the neural theory of language (Feldman & Narayanan, 2004, 2011) proposes to implement the understanding of utterances by mentally simulating their content. In human language understanding, mental simulation can exploit the neural structures used for action, perception, and memory. Linguistic information, such as constructions, frames, embodied schemata, metaphors, mental spaces, is used for parametrizing the simulation.

EASE point of view: Ideas of the simulation theory of cognition will be investigated in the EASE Subprojects P01 and R03. P01 investigates the understanding of natural-language instructions under the assumption that it requires being able to mentally execute and predict the consequences of an intended parametrized action. Subproject R03 investigates simulation-based reasoning and its application to prediction.

Mind's eye and Mental Imagery The phrase *mind's eye*, often also called mental imagery, refers to the human ability for visualization: one's ability to *see* things with the mind even in the absence of the respective external stimuli. When these things are actions and activities this idea is very similar to the simulation theory of cognition but it also includes imagining and reasoning about static scenes.

EASE point of view: In EASE the aspects of mastering everyday activity that relate to the concept of the mind's eye are investigated in the Subproject R03. R03 investigates methods to assert the belief state of an agent to a physics simulator and employ off-screen rendering in order to reason about what can be seen and to infer physical interpretations of perceived scenes.

Frames and scripts EASE also draws inspiration from the computational models of frames (Minsky, 1986) and scripts (Schank & Abelson, 1977). Minsky (1986) motivates the idea of frames as a means to provide structure and background knowledge for scene interpretation: "When one encounters a new situation (or makes a substantial change in one's view of the present problem) one selects from memory a structure called a *frame*. This is a remembered framework to be adapted to fit reality by changing details as necessary. <...> A frame is a data-structure for representing a stereotyped situation, like being in a certain kind of living room, or going to a child's birthday party. Attached to each frame are several kinds of information. Some of this information is about how to use the frame. Some is about what one can expect to happen next. Some is about what to do if these expectations are not confirmed." Scripts are extensions of frames specifically targeted at language and story understanding problems.

Frames and scripts are representational structures that are designed for exploiting the stereotypical structure of scenes and actions/activities in order to reason about them more effectively and efficiently.

EASE point of view: Frames and scripts inspire the representational structure that is used in EASE to organize the knowledge, i.e. the NEEM data structure.

Relations/synergies with Engineering/Computer Science

Like in the Cognitive Sciences, narratives and NEEMs are also promising candidates for synergies with a number of leading-edge developments in Computer Science, Artificial Intelligence, and Engineering. Since these technologies are expected to generate substantial breakthroughs over the next years, using novel, open-source technologies as a resource will be essential for the realization of the EASE research program.

Data Science *Big data* are data collections that cannot be adequately processed with conventional data processing technologies because they often satisfy a combination of the following characteristics:

- The **volume** or **quantity** of the data far exceeds that of traditional database applications
- The data exhibits large **variation** of the form in which they are stored. Often data engineers talk about them as unstructured information, data for which the syntactic structure does not mirror the semantic meaning of the data. While traditional data formats such as database tables, spread sheets, and forms are mostly very structured, typical *big data* formats include images, videos, audio files, free text formats such as entries in social media, and complete books.
- Big data are often generated and change with high **velocity**.

EASE point of view: Collections of NEEMs can also be considered *big data*. *Big data* techniques could be adapted for NEEM databases and newly developed techniques could also be relevant for other domains.

Data analytics and Unstructured Information Management A milestone system demonstrating the potential of data science was the Watson system of IBM , which won the popular US TV quiz show Jeopardy! by outperforming the best human experts in trivia question answering. The Watson system automatically read high volume web information resources including Wikipedia, dictionaries, and specialized information sources such as movie databases in order to acquire a comprehensive knowledge base enabling open domain question answering.

Scaling to open domains

The knowledge acquisition and information processing methods that were employed in the Watson system and that arguably played a key role in the scaling of the question answering capability towards huge and open knowledge domains substantially differed from the mainstream AI approach.

For one, it was found to not be necessary to transform the complete raw text into knowledge. It suffices to extract knowledge pieces and assert them. Because of the redundancy in the raw information sources, the linked knowledge pieces provide enough information to answer a large range of queries on a statistical basis.

Secondly, the knowledge bases that are extracted from the raw text information sources do not need to be consistent. It is more important that the existing knowledge bases entail the answer for all conceivable queries than that all answers derived from the knowledge bases are meaningful. Consistency checks can be incorporated into a later stage of compiling the answer, rather than in the knowledge base construction processes. Indeed, Watson generates hundreds of hypothetical answers for a single question, which are subsequently analyzed for consistency and ranked according to the confidence: it is much easier to ensure the consistency and meaningfulness of an answer rather than that of a complete knowledge base that is meant to entail the answers to all conceivable questions.

Knowledge bases can be inconsistent and redundant, only the final answer has to be reliable

Ensembles of experts

Another approach that substantially contributed to the performance of the Watson system was the employment of ensembles of experts. Rather than trying to provide a general reasoning method for answering all questions, the Watson system uses large sets of expert methods that are able to effectively find answers under restricting assumptions and use them in the context of a hypothesize-and-test control strategy.

EASE point of view: In EASE we consider collections of NEEMs *big data* from which actionable[22] knowledge can be queried. To do so, we will store and manage NEEMs using UIMA (Unstructured Information Management Architecture), an open-source software framework supporting the realization of Watson-like knowledge systems. We consider Data Science tools for the storage, management, and information mining from *big data* to be valuable tools for learning commonsense and naive physics knowledge from collections of NEEMs. Another approach is the use of Deep Learning in order to automatically generate subsymbolic representations that compress the stereotypical behavior patterns generated in everyday activity well.

Text Mining With the increasing availability of large bodies of text, there have also been many advances in processing (natural language) texts and extracting information from them. For example, the SNAP library offers tools for analyzing large social and information networks (Leskovec & Sosič, 2014). It makes it possible to summarize key information from these large sources. Given the structural similarities between NEEM (narratives) and stories, we can adapt the powerful tools developed for handling text to NEEMs.

EASE point of view: the NEEM structure enables us to consider them as stories. We expect that it will be feasible to apply *text combination and summarization* methods to NEEMs in order to construct and abstract summaries for collections of memories. The well-researched and developed linguistic methods for *text retrieval* might also be adapted for the use on NEEMs.

[22] *actionable* in this context referring to data that is made available in such a way that it can be used for and is sufficient for making decisions on how to act. `http://searchengineland.com/big-data-is-not-big-data-unless-it-gives-you-actionable-insight-167225`

This will allow us to extract meaningful parts of an experience by asking a question, for example, “what were all the episodes of successfully picking up a heavy object?”.

Graphics simulation and rendering In the field of Computer Graphics, 3D animation, and Game Engineering, we see rapid development of more powerful, realistic, and efficient methods for physical simulation and rendering. Electronics companies develop special purpose computer chips that provide strong hardware support and parallelization techniques for the relevant computation methods. As a result, we expect that powerful and fast simulation will become an available and well-developed computation technology, which can be used in EASE for collecting NEEMs and simulation-based reasoning methods.

EASE point of view: The rapid development of simulation and rendering technology will be used by EASE to obtain valuable knowledge from simulations and GwaP. By also storing these experiences as NEEMs, we will be able to acquire complementary information about objects, actions, and the effects of actions.

EASE research concerning NEEMs Research on NEEMs will focus on the following topics, which are detailed in their respective subproject descriptions in Section 1.2.7:

- **Recording NEEMs.** The research questions to be addressed for this EASE research topic are: “What are episodes?”, “"What are the data that make up NEEMs?”, “How to record NEEMs from different modalities such that they can be easily combined?”, etc. The recording of NEEMs will be investigated in the Subprojects H01, H02, H03, R01, R04, and R05.
- **Storing and managing NEEMs.** The questions to be addressed in this topic are: "How to store the experience of hundreds and even thousands of activity episodes compactly?", "How to build knowledge bases that are tailored for answering specific queries in specific contexts?", "How to compress or forget NEEMs?", "How to learn expectations to generalize over sets of NEEMs to store only most valuable knowledge?", etc. The storage and management of collections of NEEMs will be investigated in the Subprojects H03, R01, and R05.
- **Using NEEMs.** The research questions in this topic are: "How to learn environment models from NEEMs?", "How to learn capability models of robots?", "What about objects, their affordances, robot’s behavior, and expected effects?", "Which reasoning and cognitive capabilities can be learned from NEEMs?", and so on. The use of NEEMs and knowledge learned from them will be the research targets of the Subprojects R01, R04, and R05.
- **Semantics of NEEMs.** The semantics of actions and plans will be investigated in the EASE Subprojects P01 and P04.
- **Plans and the NEEMs they generate.** The design of plans that can generate NEEMs will be investigated in the Subprojects P01, P04, and R04.

1.2.6 Core concept 2: Pragmatic everyday activity manifolds (PEAMs)

We consider mastering an activity as demonstrating human-like competence in real situations in real time. Although NEEMs show promise as effective and efficient means for equipping robots with large bodies of everyday activity knowledge, by themselves they cannot support the decision making speed that humans exhibit in everyday activities.

Redefining tasks using PEAMs

Therefore, the second core concept of EASE– **pragmatic everyday activity manifolds (PEAMs)**– is concerned with how the queries needed for competent execution can be answered quickly during task execution. A PEAM is a reformulation of a reasoning problem that must be solved in order to master a specific class of everyday activities (e.g. setting a table), simplified using assumptions about the problem structure. The system is to gather information about the distributions of instances of the problem over the set of NEEMs and the expected properties of solutions based on previous findings. This information, including which problem subsets are never needed for the activity, which particular subsets are often needed, and which additional assumptions can be made, is used to create PEAMs.

We propose PEAMs as a solution to the following apparent paradox. On one hand, human decision-making for a mastered skill is staggeringly fluent, robust and adaptable. Consider for example a professional dishwasher in a restaurant; the person picks up dirty dishes such that the pile remains stable. She selectively grasps the individual dishes while already taking into consideration how she intends to clean them and place them afterwards. All this happens at a steady speed, despite requiring sophisticated scene understanding, physics reasoning, motion planning, grasp planning, and manipulation planning at different stages of execution.

On the other hand, despite tremendous progress in AI technologies[23] such as automated planning, machine learning, computer vision, probabilistic inference and representation and reasoning, the question of how artificial agents could possibly perform scene understanding, physics reasoning, motion planning, grasp planning, manipulation planning at a level required for the mastery of everyday activities remains largely unanswered. In fact, we can prove that the computational problems associated to these tasks are unsolvable (Bertero *et al.*, 1987), undecidable (Erol *et al.*, 1995; Sellen, 1996), or computationally intractable (Nebel, 1994; Bylander, 1991; Renz & Nebel, 1999) at best.

Exploiting the structure in everyday activities to boost decision making processes

This apparent paradox suggests that the brain has found *better ways of stating and decomposing the computational problems* into more feasible ones. As Horswill (1996a) put it: “We have to find and exploit the loopholes in life. We have to find and identify the structure in everyday problem-solving that enables us to reformulate unsolvable and intractable problems as computationally simpler problems without losing their validity and applicability with respect to the subset of problems that indeed has to be solved as part of a task”.

Indeed, recent breakthrough successes such as the Google car autonomously driving through California, the Watson system winning the Jeopardy! quiz show, and the Siri agent on the iPhone, have been made possible through impressive progress at a system level rather than at the level of specific problem solving algorithms.

This view of problem-solving is different from the way reasoning in artificial intelligence and robotics is typically considered. In Prolog[24], for example, we have a reasoner that is supposed to find answers to all the possible problems that can be stated as a valid Prolog query; in motion planning, often a single algorithm is supposed to find solutions to all motion planning problems that can be stated in its input language; algorithms for AI planning must find solutions to all problems that a judge of a planning competition might want to challenge the competitors with.

Typically, the use of general methods to solve a broad range of problems results in search spaces growing exponentially with search depth, making the problems computationally hard. However, there is an approach to achieve computational tractability, already known since the early days of AI. It is perhaps most succinctly expressed by Lenat & Feigenbaum (1991) in these quotes: “more knowledge implies less search” and “in the knowledge lies the power”.

Decomposition of difficult computational problems

Accordingly, the mastery of everyday activity should not require agents to employ general reasoning methods that can solve all conceivable problems of a particular class. Such reasoning methods would be too slow to be practical. Instead, we propose PEAMs as ways to decompose, structure, and reformulate computational problems and transfer them into other representational spaces to render the problems more feasible and efficient. This is illustrated in Figure 13. The

[23] This progress is documented and presented in very selective and high-impact conferences and journals including but not limited to the *International Conference on Automated Planning and Scheduling*, *International Conference on Machine Learning, Journal of Machine Learning Research*, *International Conference of Computer Vision, IEEE Conference on Computer Vision and Pattern Recognition, International Journal of Computer Vision*, *International Conference on Principles of Knowledge Representation and Reasoning*, *Conference on Uncertainty in Artificial Intelligence, and Conference on Neural Information Processing Systems*.

[24] a general-purpose logic programming language

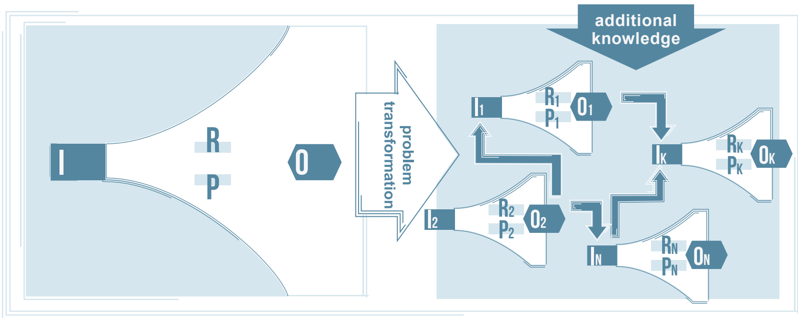


Figure 13: A computationally hard inference task and one of its PEAMs. Left: A computationally hard inference task where I is the initial problem state, R the representation, P the solving method and O the solution. The task implies a huge exponentially growing search space. Right: An approximate solution to the inference task that decomposes the task into smaller subproblems with specialized representations and inference methods. All the relevant problem space is covered by these subproblems, although they perhaps do not cover the entire possible problem space.

specialized representations and processes are designed to achieve efficiency by exploiting the structure of the subproblems and assumptions about the conditions under which they are to be solved. This is shown on the right of Figure 13. We refer to the structures and regularities of the reasoning problems and their solutions in everyday activity as *manifolds*, in analogy to the mathematical concept of manifolds referring to low-dimensional local representations. However, we will use it in a much broader sense.

Generality and tractability within everyday activity domains

In the history of AI research, different strategies have been employed to achieve this, e.g., approximation, heuristics, independence assumptions, contextual disambiguation, and dimensionality reduction. In this spirit, EASE will develop reasoning mechanisms that target both generality (from the point of view of everyday activities rather than that of computational problem classes) and tractability (do not cause significant execution delays). We call such methods **PEAM reasoners**. A PEAM reasoner might include special-purpose problem-solving strategies, algorithms, reasoners, and knowledge bases. A PEAM reasoner achieves its desired performance by exploiting the problem constraints and the knowledge provided by the respective PEAM.

PEAMs are methods to describe the subclasses of inference problems that are relevant to mastering everyday activities, so as to enable the definition and use of specialized efficient solvers for these problems.

In EASE we will start by investigating important categories of PEAMs. We will realize and analyze PEAM reasoners as research objectives of EASE subprojects. Later we will increasingly study methods for agents to automatically detect and analyze PEAMs and transform general reasoners into PEAM reasoners.

Rationale behind PEAMs To motivate the use of PEAMs let us consider the queries that generic plans might ask to decide on the course of action and parametrize behavior; these are shown in the column on the left. The second column lists the general class of computational problems that answering that query requires solving, while the third column lists the computational complexity of the respective problems.

Query	inference problem	hardness
is there a container that can hold 1 liter water on the table?	computer vision problem: what is the scene depicted in a given sensor image	ill-posed (Bertero *et al.*, 1987)
how should I reach for the coffee pot on the table?	motion planning problem	PSPACE-complete (Canny, 1988)
can I pick up the plate?	qualification problem	semi-decidable for predicate logic (Gödel, 1931)
what will happen when I crack the egg on the edge of the bowl?	ramification problem	semi-decidable for predicate logic (Gödel, 1931)
are the items on the breakfast table arranged in the right way?	inference in spatial calculi	intractable for sufficient expressiveness (Renz & Nebel, 1999)

The common approach in knowledge representation and reasoning is to start with the definition of a formal language that is designed to encode the problems and knowledge necessary for solving them. Levesque & Brachman (1987) investigate the issue that problems can be more or less difficult to solve depending on the representation language that is chosen for stating the problem and the means for solving it. More specifically, they show that the difficulty of reasoning problems can dramatically increase with the increase of the expressive power of the language used. They call this correlation the "fundamental tradeoff between the expressiveness of a representational language and its computational tractability."

The approach most commonly employed by researchers for assessing expressiveness is to define the languages used for problem-solving syntactically. A characteristic example of this approach is Bylander's seminal article on the computational complexity of propositional (STRIPS) action planning Bylander (1991). STRIPS represents actions as triples of form $\langle$*name, preconds, postconds*$\rangle$, where *name* is the name of the action, *preconds* a set of propositions that are required to hold for the action to be executable, and *postconds* a set of propositions that hold immediately after the action is successfully executed. Postconditions can be either of type "$+$" (a proposition is set to true) or "$-$" (a proposition is set to false). A planning problem is then specified through an initial state, formulated as a set of propositions asserted to initially hold, a set of action representations, and a goal specified as a set of proposition that are to hold after the plan is executed. Plan generation is the determination of a sequence of actions such that (1) the preconditions of each action holds in the state in which the action is to be executed, (2) the precondition of the first action in the sequence is satisfied by the initial state, and (3) the goal is satisfied after the action sequence has been completely executed.

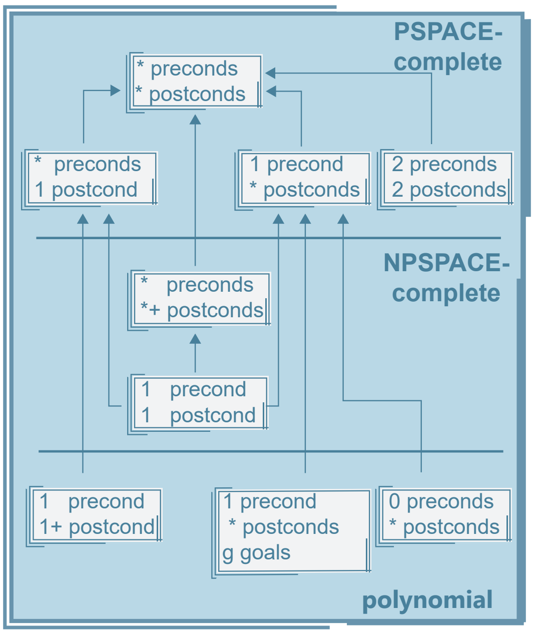


Figure 14: Complexity results for the plan existence problem in propositional STRIPS planning (Bylander, 1991).

Bylander (1991) investigates what is the impact of restricting STRIPS problem formulations on the computational complexity of algorithms needed to solve them. The respective complexity results are summarized in Figure 14. STRIPS restrictions that Bylander uses are formulated in terms of the number of preconditions for each action (0, 1, "$*$" to symbolize unrestricted etc.), the number and type of postconditions (1 means at most 1 postcondition, 1$+$ means at most one postcondition, which must be of "$+$" type), and in restricting the number of goals to a constant upper bound g. The complexity classes reported for the various restrictions range between P (Polynomial-time; largely considered the class of feasible problems), through NP (problems where candidate solutions are easy to check, but are considered difficult to find) and up to PSPACE-complete (which in practical terms means the time needed for an exact solution will be large, and heuristics or approximations are necessary, but that the memory requirements to solve the problem are not prohibitive). It can be seen that the difficulty of STRIPS problems varies greatly, and not very intuitively, with the strength of the restrictions. Problems with at most 1 precondition and 1 "+" postcondition for each action are "easy" (Polynomial-time), and they remain so even if we allow an unlimited number of "+" postconditions, but the difficulty jumps to PSPACE as soon as we allow 2 preconditions and postconditions for each action.

Similar investigations of the complexity of reasoning tasks have been conducted on reasoning domains including spatial reasoning, scheduling, temporal reasoning, constraint-based reasoning, and diagnostic reasoning.

Syntactical characterization of a problem often does not capture its difficulty

While there is a correlation between the syntactic expressiveness and the complexity of the respective reasoning problems, this correlation is often not as strong as one would wish. The fact that a problem is expressed in a more complex language and can not be expressed in the simpler language is in many cases not strong evidence that the reasoning problem is more complex. In other words, syntactic characterizations of reasoning problems might not be very informative predictors for the computational complexity. Thus there might be better characterizations to classify reasoning problems into difficult and easier ones and these better characterizations might also lead to structures that can be exploited to reduce the reasoning complexity. A second disadvantage of syntactic problem characterizations is that the increase of expressiveness might not substantially increase the set of problems that can be solved in the relevant application domain. For example, researchers have yet to determine the set of relevant planning problems that can be solved when allowing 2 preconditions instead of one.

Analyzing problem structure and its domain to specialize reasoners

An alternative approach is to start with a thorough analysis of the structure of the reasoning problems and conditions/assumptions that can be used to reduce their computational complexity. This includes the previous approach (results based on limiting the expressive power of the language used to state problems), as well as other avenues for problem simplification. One example, detailed in Section 1.2.6: Operational definition of plans, is to use constraints on how plans can be constructed to guarantee that they are easy to reason about. Another example is problem domain analysis, as proposed by Long & Fox (1998) and Fox & Long (2001). In this work, Long and Fox demonstrate how to detect and recognize specializations of planning tasks (such as navigation subproblems) and solve them with faster specialized solvers.

Based on this analysis one can then specialize the reasoning problem to take advantage of the problem structure and that the problems are to be solved only under restrictive conditions. As Bylander (1991) states in the preface of his article: "If the relationship between intelligence and computation is taken seriously, then intelligence cannot be explained by intractable theories because no intelligent creature has the time to perform intractable computations. Nor can intractable theories provide any guarantees about the performance of engineered systems. Presumably, robots don't have the time to perform intractable computations either."

The investigation of PEAMs constitutes an important research focus of EASE. The analysis and reformulation of reasoning problems to use a collection of more specialized reasoning methods that allow for the exploitation of the structures and regularities in the task domain will improve the performance on these tasks immensely.

Using findings from Cognitive and Computer Sciences as a foundation for task reformulation

Relations/synergies to Cognitive Science and Engineering/Computer Science The process of formulating and decomposing everyday activities, and identifying the most promising "manifolds" is also inspired and informed by research in the cognitive sciences. Many shortcuts to performing what in essence are complex control problems can be observed in humans and other animals. For example, cognitive psychology of motor control offers a number of insights into why human motor control in everyday manipulation performs so well and efficiently (Rosenbaum *et al.*, 1969; Shadmehr & Mussa-Ivaldi, 1994; Arbib, 2006). Arechavaleta *et al.* (2006a,b) study the stereotypical nature of human walking trajectories and characterize them with simple curve functions. Another example is the use of simulation of movements during their execution in the so-called mirror system (Arbib, 2006; Rizzolatti & Luppino, 2001; Rizzolatti, 1998). Although there are still many debates about the roles of mirror neurons in different cognitive functions, it is clear that restricting yourself to movements that are at the same time simulated can have computational advantages as it allows for slower control cycles, for the cancellation of uninformative sensor data, for prediction, and various other computational problems of motor control.

EASE does not restrict itself to specializations of problems that are inspired by the Cognitive Sciences. Research in computational sciences, e.g. Artificial Intelligence and Algorithm Theory, provide a fertile foundation. For example Horswill (1995) has shown that general perception problems, such as the localization of a robot for collision-free navigation in an office building, can be verifiably achieved with simple vision methods given assertions including that the ground is planar, all obstacles rest on the ground, and the environment is largely rectangular. Kontchakov *et al.* (2010a) investigated syntactic restrictions of logical languages that are nearly sufficient for the tasks but are computationally better behaved. Beetz (2002b) restricts the means by which plans are generated rather than restricting the languages syntactically. In his approach, he characterizes the plan language of a robot operationally as the set of atomic plan components together with the plans that can be generated by composing and revising them. Because in this setting all plans that a robot has to reason about are generated by the robot itself, one can investigate whether it is possible to avoid generating plans that are hard to reason about and still be able to generate all the behavior patterns one needs. This way, robots are enabled to reason about plans that allow for the specification of concurrent, event-driven behavior with context-specific decision making.

In the remainder of this section we will discuss three categories of PEAMs: first, manifolds of specific everyday activities, second, manifolds of low-level behavior and action, and third, manifolds contained in abstract, high-level problem-solving methods.

1. PEAMs for specific everyday activities To motivate the first category of PEAMs we consider the opportunities that everyday activity plans provide for reformulating inference tasks in order to exploit the structure of the context they are to be solved in. To illustrate this, we will consider the perception tasks required for setting the table as an example. Note that the same considerations apply to other computationally expensive tasks, such as reasoning or task and motion planning.

For setting the table, the most frequent and critical perception tasks involve finding the items to be put on the table and looking at the table to find places for the items to be arranged correctly. In the course of setting a table, a robot will have to look for cutlery in the drawer, clean cups and plates in the cupboard, milk in the fridge, etc. Thus, in the context of everyday activity, the robot can form strong expectations about the distribution of perception tasks it is to perform. It can predict which objects are to be detected, the places where the objects can be found, and what the context conditions for the perception tasks are, such as expected clutter and the lighting conditions. When looking for cups and plates, the robot can expect them to be stacked in the cupboard and for finding knives and spoons it might be sufficient to detect the right tray and pick any of the items from it.

Specializing plans

These expectations can be used by robots to autonomously specialize their fetch-and-place plans for setting the table in a particular environment. Such robots would be able to perform better than those that do not exploit this type of information. As long as task distributions are fixed, the perception routines would be faster and more robust. Thus, in the context of everyday activity, robotic agents can learn perception plans through task and environment specialization (Horswill,

1995) by specializing to the task distribution of perception tasks. If additionally they monitor the execution of the specialized plans and apply more general methods when the specialized ones fail, the robot would also not lose their general problem-solving capabilities.

Anticipatory perception

Knowledge about which specific perception tasks will have to be performed does not only aid task-specific refinement of perception mechanisms, it also aids anticipatory perception. A robot that knows it will soon have to look at the table from the front side can already inspect the scene from other view points meanwhile in order to inform future perception tasks. By taking advantage of multiple views, robots will be able to better deal with occlusions and cluttered scenes. The robots could also perceive the scene on the table beforehand in order to hypothesize a model of the scene that can be validated later when the perception task is to be performed. This helps to minimize the computational resources needed by the perception process because in many cases the validation of expectations is of lower computational complexity than the generation of correct results from scratch (Blum & Kannan, 1989).

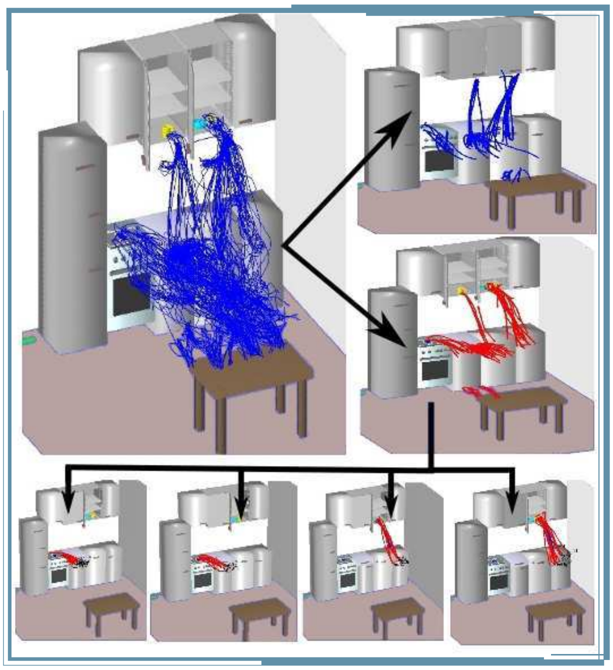


Figure 15: Stereotypical hand movements for uninstructed table setting can function as low-level PEAMs.

Another opportunity for forming PEAMs are specialized algorithms tailored for the perceptual characteristics of the expected objects in a scene rather than a general object categorization method such as classifiers trained on general purpose image databases (Deng *et al.*, 2009). The same problem structure can be exploited, for example, by robotic agents that learn probability distributions of the co-occurrence of perceptual features and objects in the respective scenes and thereby boost the performance of perception through the application of first-order probabilistic reasoning techniques (Nyga *et al.*, 2014).

Yet other possibilities include the use of mental imagination of the belief state of the robot in order to prime the perception system towards expected scenes. According to the *mind's eye* paradigm, a robotic agent can assert its belief state regarding a scene to be perceived into a simulation environment and apply off-screen rendering to predict expected images (Kosslyn & Rabin, 1999).

Pragmatic everyday activity manifolds (PEAMs) of the perception tasks for table setting could be formed through reformulations and anticipation of these tasks. Specialized algorithms decompose particular perception tasks in different ways and use different assumptions, representations, and methods to reduce the delays that perception would cause for plan execution. They are investigated in Subproject R02.

PEAMs are not only useful for performing perception tasks. Similar mechanisms can be applied to the reasoning, motion planning and grasp planning for table setting tasks. To make reasoning more effective, the robot could use query-specific knowledge bases, partial evaluation to answer subqueries as soon as the necessary information becomes available, or hypothesize answers using heuristic methods and test them afterwards, to mention only a few examples of possible PEAMs. The PEAM formation and use for these tasks are investigated in the Subprojects R04 and R05.

2. Low-level PEAMs The second category of PEAMs are those concerning low-level behavior and action. As examples of low-level PEAMs we consider stereotypical reaching motions, the low-dimensional embedding of full-body motions for table setting tasks, and the use of deep learning from experience data. Each of these will be described in more detail below.

Reaching trajectories as low-level PEAMs Human reaching is an excellent example of a low-level pragmatic manifold. In order to flexibly reach for objects in dynamic scenes, most robots employ motion planning algorithms. These algorithms take the initial pose of the robot and the desired pose as their inputs. The algorithms then compute a collision free sequence of poses that transforms the initial pose into the desired one. When running the motion planning algorithm on a distribution of reaching tasks that are required for table setting, the algorithms would produce collections of trajectories that may be optimized with respect to some cost function but would most likely exhibit a high entropy over the trajectories.

In contrast, the reaching trajectories of humans are very stereotypical. Consider the reaching trajectories of different people during table setting tasks depicted in Figure 15. The continuous trajectories in the upper left are the trajectories of the right hand during observed, complete table setting episodes. The smaller pictures show clusters of trajectory segments for reaching towards objects. The segments depicted in blue are reaching motions for handles and other furniture pieces, while the red ones are those for reaching for objects of daily use such as cups and plates (Nyga *et al.*, 2011).

Exploiting stereotypicality

As depicted in Figure 15, the trajectories for the different reaching tasks are very similar, stereotypical and efficient, even though the table setting task was performed without prior instruction by three different people of different size. Clearly, such stereotypical reaching movements might not always achieve optimal performance, but they have huge advantages in terms of computational properties and amortized performance. Since stereotypical behavior exhibits less entropy, people can more quickly learn new reaching patterns, better diagnose exceptional behavior, and more easily read the intentions of others, greatly improving nonverbal communication.

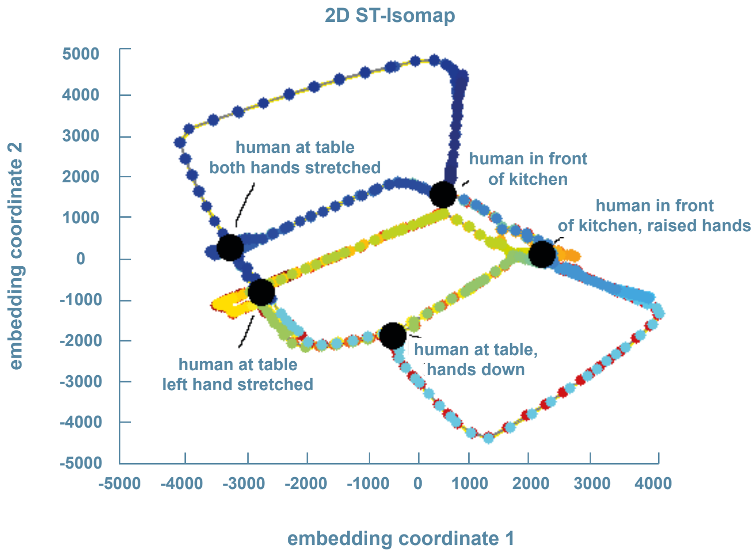


Figure 16: Low-dimensional embedding of a human table setting activity into a three-dimensional subspace.

The structure of reaching is well-studied in the cognitive psychology of human action. The literature mentions a number of constraints that human reaching motions satisfy. For example, the reaching motion is linear in the coordinate frame of the eye, satisfies a bell-shaped velocity profile, and reaching trajectories are optimized with respect to a weighted combination of minimum torque, maximum comfort, and end-state comfort. The reaching motions can also be considered to minimize the Bayesian risk of failure, to support accurate visual contact prediction, and so on.

Thus, instead of searching for optimal motion plans in the space of all possible movements (LaValle, 2006a; Latombe, 1991), we can more quickly search for appropriate plans in the subspace of the stereotypical curve functions (Ziebart *et al.*, 2009) — which, based on cognitive psychology research, might be a PEAM of human reaching. The use of stereotypical movements does not have to restrict generality. Given a manipulation context, robots will first retrieve and test stereotypical movements, and fall back to general motion planning techniques if no satisfying stereotypical reach is found or it is predicted to fail to achieve the desired goal.

Low-level activity models based on pose trajectories Instead of considering the hand trajectories for human manipulation actions, we can also investigate the pragmatic manifolds of complete human pose trajectories. Figure 16 shows an embedding of full-body pose sequences into a low-dimensional subspace using spatio-temporal non-linear dimension reduction using ST-ISOMAPs (Jenkins & Matarić, 2004b). The 3-dimensional embedding is sufficient to approximately replay the full-body pose sequences of the table setting episodes.

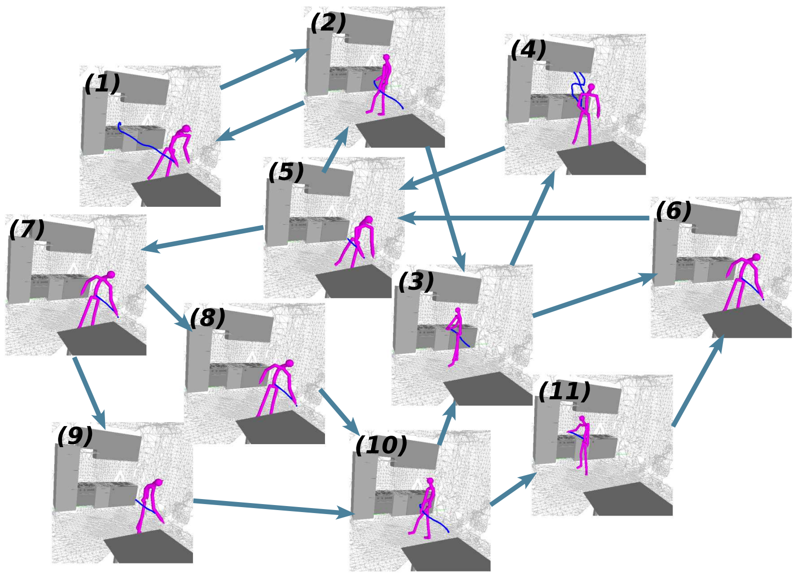


Figure 17: Automaton model learned from the low-dimensional embedding depicted in Figure 16. Encodes the full-body motion patterns and possible transitions between them.

Figure 17 shows the set of actions encoded as motion patterns and the possible transitions between them, which this low-dimensional embedding reveals. Thus, the motion is defined as a finite state automaton (Finite State Machine (FSM)) where each state corresponds to a motion pattern. Motion patterns are identified by projecting the initial high-dimensional motion sequence to a subspace of much lower dimension using nonlinear spatio-temporal dimensionality reduction. The automaton itself is learned from the low-dimensional embedding of observed or experienced activities. Once a motion FSM is obtained, we can both recognize known states from new observations as well as predict probable transitions between action states. Using such compact automata models that describe the dynamics of the captured motion, robotic agents can successfully infer the ongoing subactions leading to the overall goal of setting a table. The automaton model together with the low-dimensional embedding is another PEAM example.

Deep Learning The last example of low-level PEAMs that we want to discuss is the learning of adequate representations of high-dimensional data by uncovering hidden structure through deep learning methods (Bengio *et al.*, 2013). Such methods automatically learn, often at multiple levels, suitable intermediate representations that facilitate more effective learning. In these approaches the higher-level representations are learned in terms of lower-level ones. The objective is to abstract away from irrelevant variations, or noise, in the data, and capture the relevant information. Thus, when applying deep learning to the full-body pose trajectories of the table setting activity, one can expect the deep learning algorithms to develop representations for characteristic poses and motion patterns. An important characteristic of deep learning is that it can learn monolithic mappings from raw input data to solutions based on massive training data.

3. High-level PEAMs The third category of PEAMs are those concerned with high-level activities and concepts. We will consider grammars of actions and activities, plan languages, ontologies, and activity strategies as examples of pragmatic manifolds here.

Grammars of action Research has found that language and action share a common neural basis in the Broca area of the human brain. Broca's area deals primarily with grammatical aspects of language. It also deals with language comprehension and production, as well as activity generation and recognition. These findings have motivated some researchers, including Pastra & Aloimonos (2011), to propose grammar-based methods for activity understanding and generation.

Based on such principles, Guerra-Filho & Aloimonos (2007) have proposed a powerful system for interpreting and understanding the structure of action in unconstrained activity demonstration from video. They view this as a step towards developing a “praxicon, a computational resource that associates the lexicon of words and concepts with the corresponding motor and visual representations, enriched with information on co-occurrence patterns among the concepts for forming higher-level complex concepts [...] [The praxicon] offers support for the new idea of achieving artificial intelligence by measuring, structuring, parsing, and analyzing human behavior.”. In a similar vein, Nyga & Beetz (2015) have proposed a system for generating manipulation actions from natural-language instructions.

In EASE we will investigate grammars as a pragmatic manifold for the structure of human action and activity. Such methods will be primarily researched in the EASE Subprojects P01, R01, and R04.

Operational definition of plans Robotic agents capable of mastering everyday activity need reliable and fast algorithms for the construction of plans, the diagnosis of plan failures, and revision of sub-plans during their execution. Since these computational problems are unsolvable for arbitrary, concurrent, reactive plans, EASE will not target algorithms that work for arbitrary plans. Instead we will use algorithms that make assumptions to simplify the computational problems. An attractive alternative to making assumptions about worlds and robots, as done by other planning algorithms, is making assumptions about plans. This is attractive because the planner constructs and revises the plans and can thereby enforce the assumptions hold.

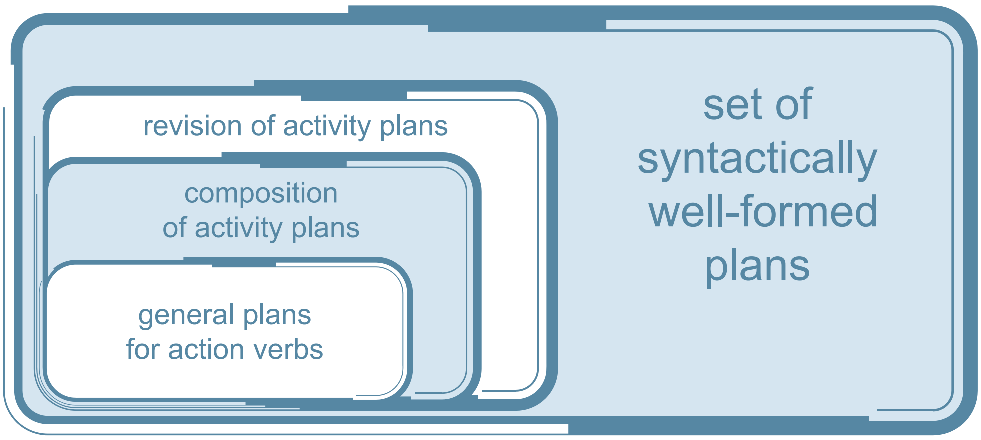


Figure 18: The set of syntactically possible plans and the set of plans that can be generated by a transformational planner.

A transformational planner generates only a small subset of valid plans: this set consists of the routine plans and their revisions (see Figure 18). Such an operational definition of the plan language is preferable over a syntactic definition because it suggests a way to enforce that the plans in the language have a certain property Q. Suppose we had, for a given property Q a method that maps any plan and any plan revision rule into a semantically equivalent plan that has or preserves the property Q. We could then use such a method to rewrite a transformational planning system with a plan library $\mathcal{P}$ and a set of plan revision rules $\mathcal{R}$ into one with a plan library $\mathcal{P}'$ and revision rules $\mathcal{R}'$, which can generate the same behaviors of the robot and forestall the same kinds of behavior flaws but reasons only about plans that satisfy the property *Q*. If *Q* simplifies the computations listed above, we can obtain simpler computational problems without limiting the scope of robot behaviors we can specify.

Beetz (1996) demonstrates how such an operational definition of plans can be used to guarantee that all plans are easy to reason about. To this end, Beetz (1996) proposes a plan language in which each plan is *transparent*, *restartable*, and *coherent*. Transparency enables the planner to understand important aspects of the plan by syntactic matching and pattern-directed retrieval. Restartability enables the robot to integrate a plan revision into a subplan by simply repeating the execution of the subplan. Coherence simplifies the construction of routine plans because coherent plans can run concurrently without additional synchronizations and still achieve a coherent problem-solving behavior.

The operational definition of the plan language used for performing everyday activities can be considered a PEAM.

Other instances of PEAMs Instead of simplifying the language in which problems can be formulated, one can also relax the required solution quality. Simon (1956) proposed waiving the requirement for optimal solutions and accepting performance that is “good enough”. The hypothesize-and-test control strategy does not require correct solutions because it tests each generated candidate solution (Blum & Kannan, 1989). For many computational problems, hypothesizing and testing can improve resource efficiency because the complexity of testing a solution is often much lower than the complexity of computing a correct solution.

In the study of topological manifolds, when e.g. mapping surfaces of three-dimensional objects onto two-dimensional coordinate charts, some local surfaces are covered multiple times while others are missing. We expect that the same thing will happen for the mapping of general reasoning problems of everyday activities into computationally better-behaving problems. This approach will result in reasoning tasks that fall into more than one subproblem, and reasoning problems for which no adequate easier computational problem can be found. We will apply methods such as ensembles of experts and methods that compute confidence ratings for their results to deal with these issues. Ferrucci *et al.* (2010) have demonstrated in the Watson system that these methods have great potential for scaling towards real world reasoning complexity.

EASE research concerning PEAMs The EASE subprojects relate to the principle of PEAMs in different ways. Some subprojects try to uncover possible manifolds in human everyday activities by studying and analyzing NEEM s of problem-solving episodes (in particular H01, H02, and H03). Others use PEAMs to make their own computational tasks feasible (especially P02, P03, and P04). Again others develop information processing methods that can exploit PEAMs in order to allow the robot to perform fast execution-time decision making (such as R01, R02, and R03).

EASE aims at discovering the *manifolds* underlying the perceptual and reasoning problems involved in everyday manipulation and at exploiting these *manifolds* to simplify the tasks appropriately and make them computationally more efficient.

EASE researchers have contributed to and used manifolds as a means for making complex information processing problems feasible in various ways.

PEAMs in perception

In the area of perception, Schill and Zetzsche have investigated biological vision systems that have evolved efficient designs of their perceptual mechanisms. These vision systems can extract a maximum amount of information from their environment using attention mechanisms that maximize the information gain of autonomous agents (Schill *et al.*, 2009). Schill and Zetzsche transferred their results to artificial systems that represent uncertainty through Dempster-Shafer belief measures. By reasoning about uncertainty and using the information gain as the guiding manifold, the system can sequentially select informative image regions, identify the local structure in these regions, and use this information for drawing efficient conclusions about objects in the scene. Frese and his colleagues use manifolds implied by the laws of physics to simplify perceptual tasks. For example, they have developed very fast, accurate, and robust probabilistic ball-tracking algorithms that enable an autonomous robot to catch balls thrown towards it (Birbach & Frese, 2013). They exploit the fact that ball detections can be restricted to the manifold of physically possible ball trajectories, and thereby increase the accuracy, robustness, and efficiency of their algorithms in the face of sensor data uncertainty. Zachmann and his colleagues exploit the coupling between the motions of individual fingers in order to estimate the 27-dimensional hand pose in a lower dimensional embedding (Mohr & Zachmann, 2010a,b). They have hierarchically decomposed the manifold of all poses in image space by a hierarchy of area-based templates, resulting in an extremely fast method. Zachmann (1998, 2002); Zachmann & Weller (2006) also approach the problem of 3D proximity computations among virtual objects consisting of millions of polygons by exploiting manifolds that approximate the 3D objects with successively finer levels of detail. One kind of such an approximation is based on a special kind of bounding volume that can by itself approximate convex hulls arbitrarily closely. Another kind of approximation is based on polydisperse spheres and allows for extremely fast approximations of the intersection volume of virtual objects (Weller & Zachmann, 2010, 2011). Yet another way of utilizing the concept of manifolds is their approach to define surfaces over point clouds (Klein & Zachmann, 2004b,a).

PEAMs in robot control

In the area of grasping, Ritter and his colleagues (Li *et al.*, 2015) explore manifolds for manual interaction. They gradually increase manual competence by exploring manual interaction spaces for many different kinds of objects, investigating different strategies for such active exploration in realistic settings. Control approaches presented by Albu-Schäffer *et al.* (2007) use torque, position and impedance control to address different manipulation tasks, as for example opening a door or wiping a table. Due to the robustness of these controllers with respect to uncertainty, manifolds for different manipulation tasks can be considered, situations identified where the precise manipulation can be offloaded to the reactive controllers.

PEAMs in knowledge representation and reasoning

In the area of knowledge representation and reasoning Baader *et al.* (2005) have proposed the description logic EL++ as an expressive yet tractable fragment of larger OWL languages, which has initiated a paradigm shift to less expressive, but computationally effective ontology languages. EL++ was later standardized as OWL2 EL in the W3C recommendation OWL2 (Motik, 2008). EL++ and related languages such as DL-Lite can be seen as a syntactic manifold inside OWL2. Since 2005, this manifold has been pushed further in several directions. One reason for the good computational behavior of EL++ is that it resides in the Horn fragment of first-order logic. This initiated a careful investigation of the limits of the 'Horn-ness' of an ontology language, which can itself be viewed as a manifold. Lutz has contributed widely to this research, analyzing the computational complexity and expressive power of this family of languages (Krisnadhi & Lutz, 2007; Lutz, 2008; Eiter *et al.*, 2009) and establishing intimate links to constraint satisfaction problems that explain the capabilities and limits of this manifold (Lutz & Wolter, 2012). Another reason for the good computational behavior of EL++ is that it admits a certain type of goal-directed calculus, called 'consequence-based', that was first proposed by Lutz and coauthors along with the EL++ language itself. The applicability of this kind of calculus could also be seen as a manifold and has recently been investigated in detail by researchers from Oxford, leading to new calculi for practical ontologies formulated in expressive ontology languages, and to results about parametric complexity that intend to explain the surprisingly good computational characteristics of real-world ontologies.

Große and Drechsler have worked in the field of formal verification and investigated techniques based on symbolic execution as suitable for correctness checking of software and hardware (Le *et al.*, 2013). Native execution and parallelization has been exploited to significantly improve the scalability of symbolic execution in this context (Herdt *et al.*, 2016). However, domain specific optimizations as well as suitable abstractions are needed to enable scalable verification of plan properties.

1.2.7 Research plan

In this section we outline the full EASE research plan, organized in three 4-year phases, and in particular the first 4-year phase. EASE research takes the cognition-enabled control framework as a starting point. Each phase is to use and extend 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 19.

Here we will provide a short summary for each of the three 4-year phases before describing the research areas and projects of the first phase in detail. Phase 2 and 3 are increasingly loosely defined as they are further in the future.

- **Phase 1: Understanding by building (2017-2021)**: The focus of the first research phase is on understanding everyday activity and building a common knowledge representation across the different research areas. EASE will investigate how humans master everyday activities and, based on this knowledge, we will design, realize, and analyze information processing models that enable robotic agents to successfully perform long-term everyday activity in a simplified household setting.

  This phase includes research in the collection, representation and compression of experiences and how to extract generally applicable knowledge from them, as well as establishing a database of NEEMs. In addition, PEAMs and their role in mastering everyday activity are studied in various subprojects.

  From the start, EASE will have access to complete, robotic systems that the subprojects can use to incorporate/test the components they research. Rather than researching aspects separately and later trying to fit them together in one coherent system, we intend to take a much more interactive approach. A complete robotic system will be evolved using research components, and the components will be tested and refined using the available system.

  After the first two years, we expect to have collections of NEEMs of categories of everyday activities that are commonly investigated by the different EASE subprojects. They include NEEMs from robot and human activity as well as ones generated by reading instructions for activities, simulations, and playing Games with a Purpose (GwaP) about these activities. To achieve maximal synergies between the subprojects, the NEEMs will be represented in OPENEASE and linked to a common ontology used by all partners (Tenorth & Beetz, 2015).

- **Phase 2: Common information processing framework (2021-2025)**: EASE will focus its research on combining the research from the first phase into a common framework. The framework will integrate individual research results into a common model that facilitates key learning and reasoning approaches such as broad commonsense and naive physics reasoning, cognition as simulation, language foundations of cognitive control, and prediction-based robot control. We expect to use a collection of various representation methods that may be redundant (together with methods to resolve possible inconsistencies) and ensembles of specialized reasoning methods.

  In the second phase, the concepts of NEEMs and PEAMs will be defined more rigorously and the effects of these representations on the reasoning capabilities of the robotic agents studied more thoroughly. We will investigate the principles underlying the representation and reasoning mechanisms employed in the control of the robotic agents and their interplay with each other deeper. Challenges include the design, implementation, and analysis of systems that allow reasoning problems to be solved by combining different methods that might arrive at different conclusions. For example, a particular spatial reasoning problem could be solved through logic-based methods or by mentally simulating a plan execution and extracting the answer from the simulation results. As both reasoning methods work on a different granularity of representation inconsistencies between the reasoning results are possible and need to be dealt with appropriately.

We will continue to generate hypotheses about how humans perform everyday activity and test these in experiments.

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

  In this phase we will investigate how the structure of everyday activity facilitates the cooperation between different agents. For example, it supports using powerful mechanisms such as implicit communication: Take two persons building an Ikea shelf as an example. The assisting person can hand over the right tools at the right time without hindering the primary action, often without the need of a request for it. Another example would be a scenario in which a person serves food to a guest, interpreting that the guest is not proceeding with the expected activity and looking around as cues that something is missing. EASE will investigate learned models of activities and the knowledge abstracted from them to understand and replicate such competent cooperation and implicit coordination patterns.

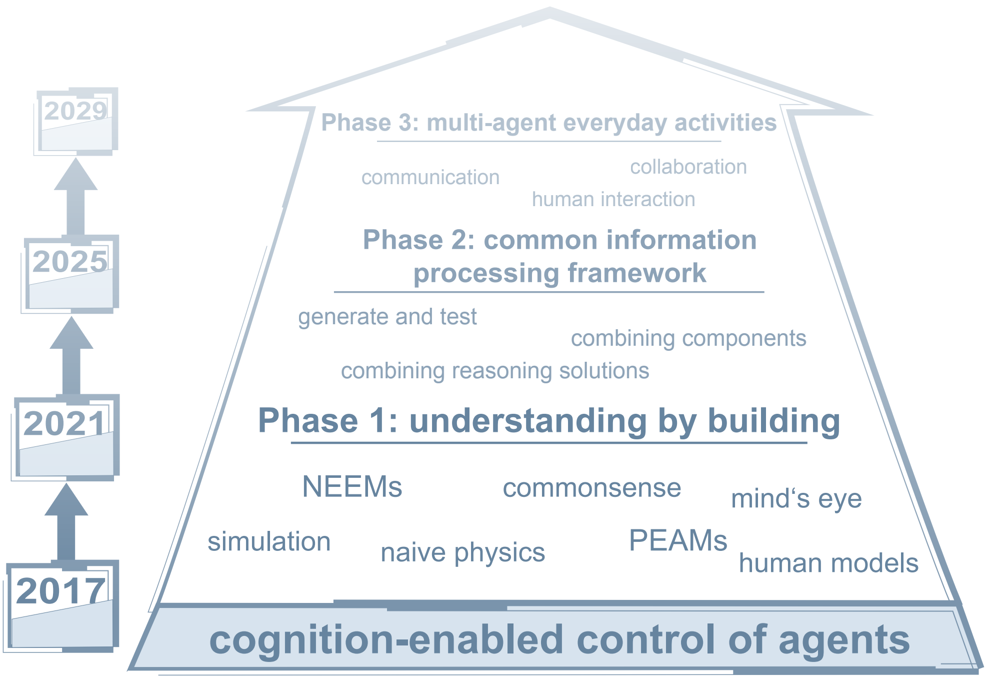


Figure 19: Long-term perspective of the EASE research program.

Structure of the EASE research program

This section describes the concrete EASE research program. First, we will define the research areas and give a broad overview. Figure 20 shows how the EASE CRC is structured into three research areas and how these areas cooperate with each other:

1. **Research Area H** (Descriptive models of human everyday activity, depicted in yellow) will design novel methods to reveal the knowledge and processes underlying human mastery of everyday activity. It will build descriptive models of these capabilities. It will also try to detect constraints and PEAMs that can be exploited to simplify the information processing needs.

2. **Research Area P** (Principles of information processing for everyday activity, depicted in red) will investigate information processing principles of everyday activity. This includes a language perspective of action, logical foundations and formal reasoning aspects, and methods for probabilistic reasoning and learning of aspects of everyday activity. It uses information from Research Area H to take on relevant, computationally hard information processing tasks and make the solution thereof more efficient.

3. **Research Area R** (Generative models for mastering everyday activity and their embodiment, depicted in green) will build and examine comprehensive generative models for mastering everyday activity and will embody them into a robotic agent to test its capabilities. Useful heuristics and preference models are tested and incorporated into the decision-making and control process. The findings from Research Area R can be fed back to Research Areas P and H to inform the representations, models, and methods investigated and will likely raise new questions that require intensive collaboration.

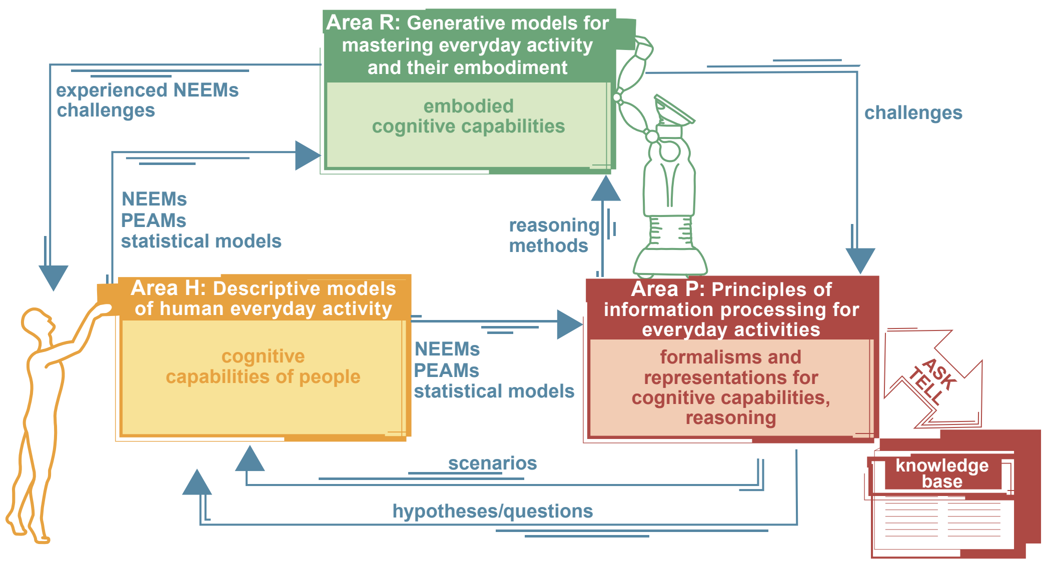


Figure 20: The interplay of the three main research areas in EASE. The information processing principles investigated in Research Area P (red) will be applied to robotic agents in Research Area R (green) and the resulting performance is analyzed. The formation of the core concepts is facilitated through Research area H (yellow), which aims at the acquisition of descriptive models of how humans master their everyday activities.

Figure 21 shows how the research areas and individual subprojects contribute to the information processing model described in Section 1.2.3 and Figure 6. “generation of NEEMs” and “acquisition of generalized knowledge” focus on building a NEEM-enabled knowledge system.

"analysis of information processes and control tasks" and "inference mechanisms for competent activity" focus on PEAM-enabled task optimization. Finally, "generative models for everyday activity" focuses on cognition-enabled control using plan-based task execution.

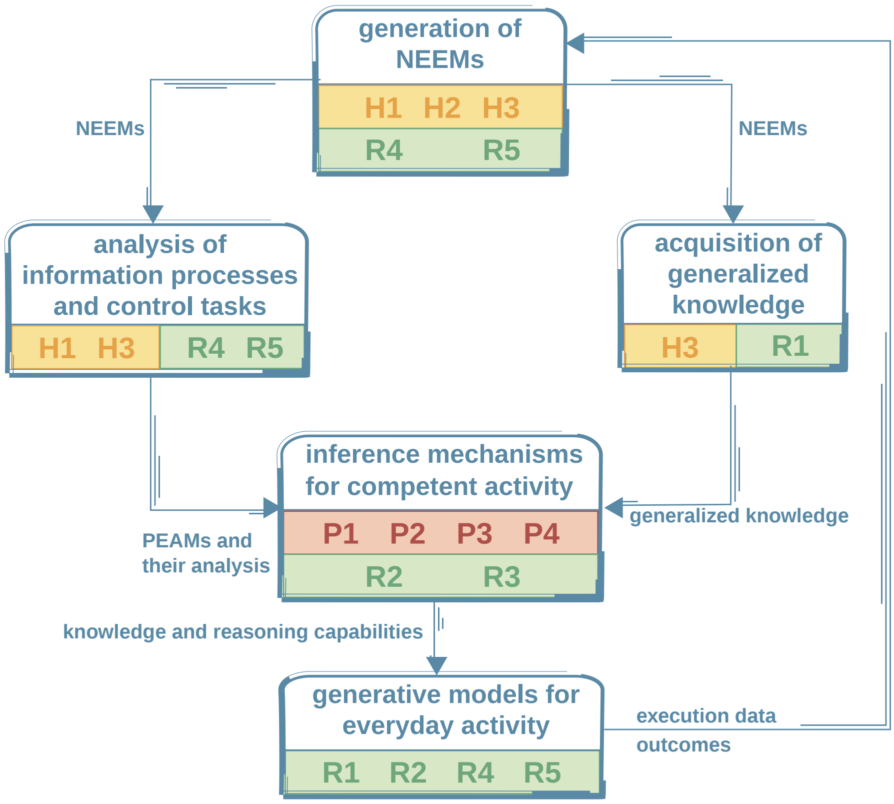


Figure 21: Research processes and their relations to research areas and subprojects

A description of the subprojects contained within each research area is given in Table 1. The particular research goals and objectives of the research areas and subprojects will be detailed in the following sections. Note that the goals are long-term challenges to be addressed for the 12-year duration of EASE, while the objectives are to be achieved in the first 4-year phase of EASE.

Research Area H: Descriptive models of human everyday activity investigates how humans accomplish their everyday tasks, represents the respective data as NEEMs, builds models of activities, and identifies possible PEAMs for tackling tasks.	
Code	Project description
H01	**Acquiring activity models by situating people in virtual environments** performs controlled experiments with humans performing manipulation tasks in specially designed virtual environments to investigate everyday activity strategies in unfamiliar or unexpected situations.
H02	**Mining and explicating instructions for everyday activities** investigates methods for obtaining knowledge about everyday activities through reading instructions from the web. It disambiguates and completes the instructions using simulation and Games with a Purpose (GwaPs).
H03	**Natural activity statistics** collect, annotate, analyze, and interpret complex human everyday activities by a combination of statistically driven bottom-up and guided top-down methods in order to detect PEAMs, on which generative models of human activities will be learned.

Research Area P: Principles of information processing for everyday activity researches the representational foundations of NEEMs and robot plans, and investigates the potential of using PEAMs to make reasoning tasks feasible within the time constraints of the task.	
Code	Project description
P01	**Embodied semantics for the language of action and change: Combining analysis, reasoning and simulation** investigates simulation-based semantics for action steps (based on the respective action verb) that constitute the atomic steps of narratives and creates logical formalizations of compound narratives.
P02	**Rightsizing ontologies** investigates the fundamental trade-off between expressiveness and tractability in ontological reasoning by identifying PEAMs for reasoning that have just the right expressiveness.
P03	**Spatial reasoning in everyday activity** performs an in-depth investigation of the spatio-temporal aspects of reasoning about everyday activity in particular and investigates PEAMs in this domain.
P04	**Formalizations and properties of plans** investigates formalizations and axiomatizations of the plans used by the robot to perform everyday activities. It develops methods for the verification of safety constraints of the generated robot plans such that heuristic planning and learning methods can be used without compromising safety.

Research Area R: Generative models for mastering everyday activity and their embodiment takes the knowledge about humans performing everyday activities (Research Area H), representation and reasoning mechanisms investigated in Research Area P, and experience-based learning to build comprehensive information processing models that enable robotic agents to master everyday activities. It also includes the construction of NEEM databases from execution logs, which will be useful not only for constructing the robot control framework and system, but also informs Research Areas H and P.	
Code	Project description
R01	**NEEM-based embodied knowledge framework** designs and realizes the subcomponent of the information processing and control system that acquires and manages NEEMs, and abstracts the information contained in NEEMs into generalized knowledge.
R02	**Multi-cue perception supported by background knowledge** investigates the efficient and robust accomplishment of selected challenging perception problems in the context of everyday activity that require common knowledge and the exploitation of PEAMs.
R03	**Embodied simulation-enabled reasoning** investigates embodied reasoning methods that use simulation-based prediction. This subproject also aims to develop faster-than-realtime simulators.
R04	**Specializing and optimizing generic robot plans** realizes control systems for robots for performing the household chores in the main scenario, including preparing simple meals, setting the table, cleaning up, shopping, and storing purchased items. Scientifically the subproject investigates how plans can be optimized through specialization by exploiting the PEAMs of the respective everyday activity.
R05	**Episodic memory for everyday manual activities** investigates how collections of NEEMs of everyday manual actions can be obtained for robots and how they can be used to bridge between semantic, procedural, and perceptual memories in order to support competent and dexterous manipulation in everyday contexts.

Table 1: Short description of the research areas and subprojects within these areas.

Research Area H: Descriptive models of human everyday activity

Humans are capable of performing everyday activities near-optimally while at the same time being capable of timely, flexible responses to whatever happens (Anderson, 1995). So far we do not have a good understanding of how we can design computational mechanisms that can achieve such efficiency and are at the same time as adaptive as human activity.

The goal of Research Area H, which will be pursued throughout the duration of EASE, is concerned with a better understanding of the mechanisms needed for these performance characteristics. To this end we will investigate information processing models of humans, with a focus on giving answers to the research questions listed on page 2. EASE will research, realize, and analyze methods for systematically collecting information from humans performing everyday activity. The data acquired through these methods will form the basis for the information processing models. The data will be transformed into NEEMs as a common representation and studied using the reasoning principles discovered by Research Area P.

The goal of Research Area H is to understand why and how people can perform vague instructions for everyday manipulation tasks so competently and investigate hypotheses about the form and role of PEAMs in competent human everyday manipulation activities.

The role of Research Area H in EASE is threefold.

1. Research Area H will acquire data about humans mastering everyday activities from the following sources. Firstly, we will conduct full-body observation of humans performing activities (such as table setting and cleaning up) in a kitchen environment in virtual reality. This setup allows us to analyze the processes and strategies underlying behavior by challenging them with impossible situations. The activities can be combined with think-aloud protocols as well as bio-signal data streams of the participant. The second source are natural-language instructions on how to perform everyday tasks, obtained from texts. Thirdly, data will be generated through interactive Games with a Purpose (GwaPs), by placing players into situations in which they have to apply knowledge needed for the mastery of everyday activities.
2. Research Area H will represent the acquired human data in the common NEEM representation and make them available in OPENEASE.
3. NEEMs will be interpreted and abstracted into layered models of everyday activity and the researchers will hypothesize and investigate possible PEAMs underlying the human competence in everyday activity. NEEMs will be used to learn semantic knowledge about activities and design (new) experiments with human participants. One of the main outcomes of the research in Area H will be to propose PEAMs that can be used in the other research areas to make computationally hard decision making problems easier to solve.

For the first 4-year phase, the goal of Research Area H is decomposed into two tightly integrated objectives with a particular focus on the detection and analysis of potential PEAMs:

Objective H.a: Acquiring and managing multimodal, semantically annotated, high-volume data sets of humans performing vaguely formulated everyday manipulation tasks

One of the goals of Research Area H is to understand how people perform vaguely formulated everyday manipulation tasks. We decompose this question into three interacting subquestions:

- What is the generated observable behavior for these vague instructions?
- How do humans communicate their everyday activities in natural language?

- What happens when we challenge human competence in experimental settings using virtual scenarios and computer games?

To find the answers to these questions, EASE will use three primary modalities for acquiring data of human everyday activity: manipulation in virtual reality, computer Games with a Purpose (GwaP), and machine reading of text instructions for everyday manipulation activities relevant to the main scenario.

Objective H.b: Learning descriptive and causal models of everyday manipulation activities

Human strategies in everyday manipulation tasks are often very stereotypical (Arechavaleta *et al.*, 2006a,b) and optimized (Todorov, 2004). Examples of such stereotypicality are found in reaching motions that are optimized w.r.t. end state comfort, minimum torque, and Bayesian risk (Körding & Wolpert, 2004b,a). The low entropy of such behaviors allows for faster learning and better monitoring and diagnosis, thus for better optimization. The objective is to discover the principles underlying this behavior and research PEAMs that can be based on these principles.

To do so, we will investigate the representation, generalization, and refinement of abstract models of the data collected under Objective H.a and investigate PEAMs as possible means of achieving high-performance behavior. We will research descriptive, normative and predictive models. To obtain such models, the researchers will investigate the application of structure mapping, analogical reasoning, and statistical and probabilistic learning on the NEEMs acquired as part of Objective H.a.

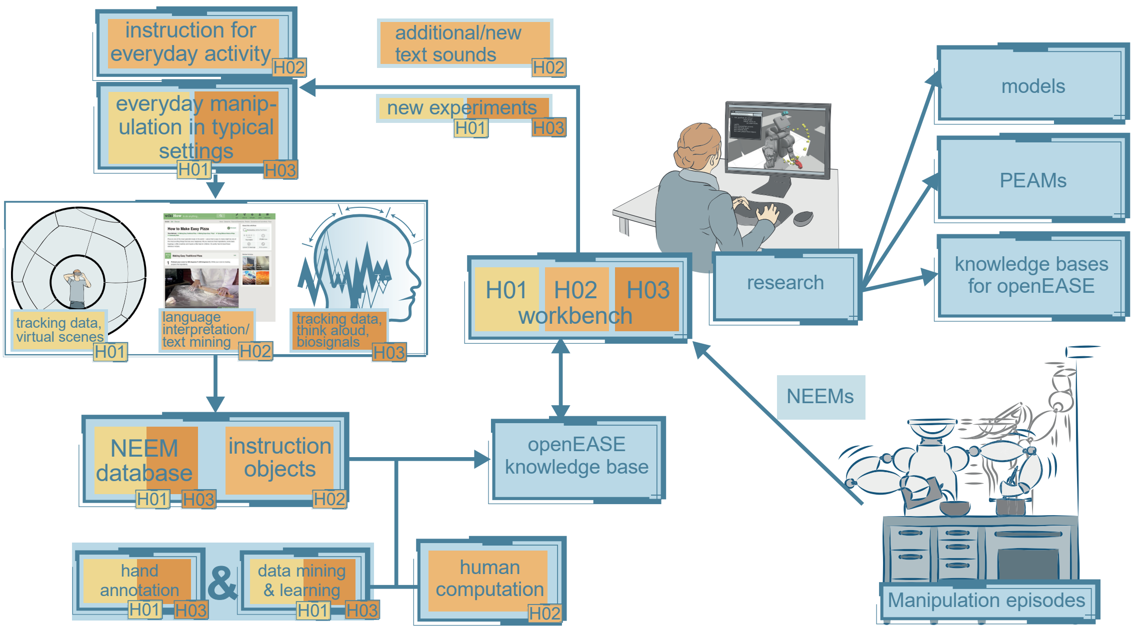


Figure 22: General organization of subprojects in Research Area H.

Structure of subprojects in the Research Area H The subprojects in Research Area H share a common organization, which is depicted in Figure 22. Researchers conduct experiments with humans performing everyday activities and log the data into big data databases. Automated interpretation routines and semi-automated methods that (partly) rely on humans annotating the low-level experimental data will transform the data into NEEMs, which will form the basis of the respective OPENEASE knowledge bases. The researchers will then process the data using visual analytics, statistics, and learning tools. The relevant interpretation and learning tools developed and used by the researchers of the subprojects will be encapsulated to be operational in OPENEASE. Using extended and tailored OPENEASE web interfaces, the researchers generate models of human everyday activity, hypotheses of PEAMs, and knowledge bases as primary outcomes of the research area.

Research Area H is composed of three subprojects. Subproject H01 will inform cognitive models of everyday activity by observing humans performing manipulation tasks in a virtual reality environment. Subproject H02 will be concerned with the acquisition of abstract and disembodied knowledge by reading text sources from the Web and using games with a purpose. Subproject H03 will contribute additional modalities to be included into NEEMs. These complementary data sources will include bio-signal data streams and think-aloud protocols of activities. The goal will be the development and investigation of bottom-up and top-down models that can explain how humans accomplish the activities contained in the recorded experimental data.

Subproject H01: Acquiring activity models by situating people in virtual environments

The role of Subproject H01 is to explicate reasoning and decision making mechanisms for grasping and tool use. These mechanisms are so automated and implicit that people are not consciously aware of them. The basic idea of Subproject H01 is to break the automatisms of everyday activity by confronting human agents with **physically impossible objects and scenes** in virtual reality. These virtual reality scenarios can contain objects and physical relations/effects that cannot exist in reality.

Using impossible objects and scenes in VR to investigate unconscious human reasoning processes

A virtusphere is used to display the virtual environment realistically. It is a 3 meter hollow sphere, placed on a special platform that allows the sphere to rotate freely in any direction according to the user's steps. A user is able to walk and run inside the sphere, viewing the virtual environment through the head-mounted display. The sensors collect and send data to the computer in real time and the user's movement is replicated within the virtual environment.

Subproject H01 conducts experiments in the virtusphere where humans who are instructed to perform everyday activities with impossible objects and in impossible situations are observed. To do so, methods are developed that allow for realistic, real time simulation of hand-object interactions. The full body pose data, the state of the virtual environment, and the interactions between the two are all stored in big data databases with semantic indexing structures. This is realized by transforming the data into NEEMs and making them available through OPENEASE.

The OPENEASE knowledge bases are used to investigate models of everyday activities, gain scientific insights and suggest PEAMs. The main research results will be models of how humans grasp objects and use tools. The methods used will include Hierarchical Hidden Markov models, learned models over trajectories, and heat map representations. H01 will also produce algorithms for simulating manipulation actions based on the learned (generative) models. The categories of activity models as well as the learning algorithms will be integrated into the KNOWROB knowledge representation system (see Section 1.2.10) within OPENEASE. Synergies are facilitated by aligning the scenarios of the experiments in Subproject H03 with the instructions mined in Subproject H02, as well as the activities that the robotic agents have to accomplish in Subproject R04.

Subproject H02: Mining and explicating instructions for everyday activities

The role of Subproject H02 is to acquire knowledge about everyday activities through machine reading of natural language instructions made *by* humans *for* humans. Natural language instructions are valuable knowledge sources for mastering everyday activities. They do not only describe the sequence of actions to be performed, but also give hints and information about what might go wrong and how to avoid such mistakes. Additionally, EASE aims for robots that can execute underspecified instructions such as the ones made by humans.

Methods to gain knowledge from natural language instructions and GwaP

For example, the instruction “flip the pancake” does not mention the tool to be used, the destination of the pancake after flipping, and other pieces of information that an artificial agent requires to perform the action successfully. Subproject H02 investigates the under-specifications that are common in natural language instructions made by humans for humans and develops methods for explicating and augmenting missing information.

Where explication and augmentation cannot be performed through natural language technology, Subproject H02 employs human computation: the system automatically generates tasks for humans that help it fill knowledge gaps. To this end, Subproject H02 will generate interactive game (GwaP) episodes and interpret human solutions in order to acquire the missing knowledge.

The instructions are chosen such that the subproject can achieve maximal synergies with the other EASE subprojects. The output of the instruction interpretation and learning step will be NEEM narratives. The NEEM narratives will be represented in the EASE knowledge format and made accessible to other EASE subprojects through the openEASE infrastructure.

Combining top-down and bottom-up approaches to analyze human everyday activity

Subproject H03: Natural activity statistics The role of Subproject H03 is the investigation of typical everyday activities in order to acquire synchronized activity models at different levels of abstraction. To this end, the research activity is organized as a combination of bottom-up and top-down research.

In the bottom-up branch, full-body pose trajectories together with bio-signal data streams and think-aloud protocols are collected. The combination of these data will constitute a unique and comprehensive data set about human everyday activity, which will be processed through *data mining, probabilistic model learning* and *deep learning* techniques.

The top-down methods will investigate probabilistic inference hierarchies, in particular ones based on Dempster-Shafer’s theory of belief functions, in combination with semantic descriptors including named entities that are aligned with the low-level data.

The bottom-up and top-down branches will jointly acquire structured representations of human activities at different levels of abstractions through automated learning and interpretation methods.

The data as well as the learned models will be made accessible to the other EASE subprojects through openEASE. The activities to be investigated will again be coordinated with the other EASE subprojects in order to achieve maximal synergies.

Research Area P: Principles of information processing for everyday activity

The knowledge about everyday activity from Area H needs to be represented such that it is useful to artificial systems. Research Area P will investigate the representational foundations, reasoning techniques and formalizations of NEEMs, and common knowledge and plans for mastering everyday activity. The aim is to design representation and reasoning mechanisms that capture the intuitions behind human activity. If we map a reasoning problem into the respective formal representation, apply the automated reasoning method to the problem, and consider the semantics of the formally derived result then this ideally should be comparable to the commonsense reasoning of humans.

Research Area P will advance our understanding and ability to work with the knowledge underlying intelligent everyday behavior. It will feed directly into various control mechanisms for robotic agents as investigated in the context of Research Area R and provide feedback for the research activities in Research Area H.

The goal of Research Area P is to understand the representation and reasoning foundations of information processing methods that enable robotic agents to master everyday activities.

Processes at different levels of control typically benefit most from different kinds of representations. The representations an agent holds can differ along several dimensions:

- Abstract representations (linguistic, plans, narrative) vs. concrete representations (simulation).
- Expressive formal languages for representation (and offline reasoning) vs. systematically and well-defined simpler approximations for more effective (online) reasoning.
- Explicit problem representations at different levels of complexity held by cognitive agents vs. cognitive *outsourcing* to external structures in the environment.

Typically, one has to choose a trade-off along these dimensions and in the past, intelligent agency has largely been investigated at a specific and fixed level of the representational spectrum. In contrast, EASE aims for representations and reasoning methods that allow for context- and task-specific selection of representation.

Research Area P provides both the theoretical underpinnings and methodological advances necessary for supporting flexibility of this kind and the application of these principles to competent activity. The area is targeted at the key representational concepts required: in particular, a better formal understanding of NEEMs, background knowledge, spatio-temporal representations in manipulation, and plans that can produce flexible and robust everyday behavior.

Goal P is decomposed into three objectives to be achieved in the first phase:

Objective P.a: Effective representation of experience in NEEMs and using it to learn and reason about consequences of actions in embodied systems

Using explicit assertions to capture the commonsense and naive-physics knowledge needed for competent everyday activity has proven to be tedious and difficult. One reason this may prove so difficult is that building an abstract model of this knowledge in the form of assertions implies that the model has to be correct for all possible cases it abstracts from. In contrast, EASE will investigate the storage and use of this knowledge in terms of experiences, in particular as NEEMs. Thus, instead of asserting all knowledge that might possibly be relevant even though most situations will never occur, the knowledge is extracted from a sample of real cases. Given that the everyday activities will be repeated very often, the sample will contain most of the relevant situations.

We will develop (1) novel logic-based representation and reasoning methods for NEEM narratives, and (2) a novel simulation semantics for action verbs used in NEEM narratives.

Objective P.b: Flexibly using the spectrum of formal languages for knowledge representation with varying expressivity

Different representations of the same situations, actions, etc. are optimal for different types of queries. Our objective is to overcome the strict separation between expressive versus lightweight representation and reasoning methods. We aim to construct representational structures that support the flexible selection of expressivity, ontology approximations, and problem representations for everyday reasoning that reflect the flexibilities observed in human problem solving. This will be achieved by providing and comparing a collection of different representational mechanisms, and developing suitable abstractions and approximations among these different representations. With the latter, it will be possible to move between different representations in trade-offs between expressiveness, efficiency, and representation by an agent versus offloading to the environment. Thus, one could for example use a more expressive representation for modeling purposes and later move to a more efficient one for reasoning purposes, or rely on combinations of internal cognitive spatial representations and properties of the environment to outsource selected aspects of the problem structure.

The research will be focused on (1) the representation of background knowledge using ontological representations and (2) spatio-temporal reasoning for manipulation activities. In particular, EASE will focus on fast inference methods for these applications by investigating PEAMs of the respective reasoning problems. Attention will especially be paid to how the representation of problems itself can contribute to finding solutions. Again in the spatio-temporal domain, problem representations will be explored that trade off internal and external information with explicit pre-structuring of the problem. EASE considers the ways in which humans structure their approaches to diverse problems as an additional solid indicator of how to best achieve such capabilities in artificial systems.

We will give due consideration to the logical foundations as an essential part of our investigations of information processing principles. As Nilsson (1991) put it: “Anyone who attempts to develop theoretical apparatus relevant to systems that use and manipulate declaratively represented knowledge, and does so without taking into account the prior theoretical results of logicians on these topics, risks (at best) having to repeat some of the work done by the brightest minds of the twentieth century and (at worse) getting it wrong.”

Objective P.c: Formalizing plans for mastering everyday activities and investigating the properties of these plans and their preservation under plan revisions

This research objective concerns itself with how we can ensure that plans for everyday activities satisfy certain conditions with respect to safety, robustness, and goal achievability using plan property verification.

Plans that are predetermined sequences of actions cannot achieve the flexibility, robustness, and efficiency required for mastering everyday activity. They have to specify how the robotic agent is to respond to sensory and unexpected events in order to successfully accomplish its tasks. To competently deal with atypical situations, robots also have to be able to revise their course of action. For this purpose, EASE will investigate formalizations of plans and properties that facilitate fast and adaptive planning.

Structure of subprojects in the Research Area P Research Area P focuses on the information processing principles for mastering everyday activities. Here we aim at understanding how technical systems and robotic agents can achieve the generality, flexibility and robustness of human performance in everyday activities. Subproject P01 will investigate different representation and reasoning approaches for NEEMs. Subproject P02 will investigate efficient reasoning with encyclopedic knowledge by studying description-logic representations that are expressive enough for everyday reasoning purposes, but restricted enough to allow for resource efficient reasoning. Subproject P03 will analyze spatio-temporal structures in representations of everyday activities. Finally, Subproject P04 will investigate formal models of plans for everyday activity and methods for verifying important properties of the plan-based controllers.

Mental execution/simulation of (under-specified) natural language instructions

Subproject P01: Embodied semantics for the language of action and change: Combining analysis, reasoning and simulation The role of Subproject P01 is to further investigate the relation between language (in particular, instructions), action, and simulation. The perspective behind P01 is that understanding instructions entails being able to (mentally) perform the instructions. This view can be contrasted with alternative approaches that aim at language applications, such as question answering based on texts, or machine translation of texts. A key difference is that in order to (mentally) perform instructions, knowledge gaps in the written sources have to be filled and the meaning of words and expressions have to be fully disambiguated. When executing instructions, the agent has to commit to one specific motion trajectory, grasp one particular object, select one particular grasp with a specific positioning of the fingers, and hold objects in specific poses. This commitment to action instances goes far beyond what is typically necessary for linguistic applications.

Subproject P01 investigates the foundations of agents that employ action simulation in order to reason about the instructions they receive. In particular, this subproject aims at developing a simulation-relevant semantics for natural-language instructions by combining approaches targeting simulations with more formal models of the linguistic semantics of actions and their contextualization. The approach to combine formal models with simulation-based semantics of language is internationally unique. P01 also investigates the potential role of simulation as a PEAM of the reasoning mechanisms for language interpretation.

Subproject P01 will enable the realization of robotic agents that answer queries about the effects, the behavior, and risks in the execution of instructions. A possible implementation basis for the simulation techniques will be those researched in Subproject R03, possibly extended with more abstract simulation mechanisms such as ST-Isomaps that can be learned from the NEEMs. Cooperations with the Subprojects R01 and R04 will investigate the possibilities to use interpreted instructions directly to control the physical robotic agents.

Answering queries about effects, behavior and risks associated with instructions

Subproject P02: Rightsizing ontologies The role of Subproject P02 is to investigate the efficient reasoning with expressive ontological knowledge bases that is needed to capture everyday activity knowledge. Ontological knowledge bases play a fundamental role in EASE. First, they are important representation and reasoning tools to cope with huge and open knowledge bases. Second, they are the basis for connecting data from NEEMs with the background knowledge of the robotic agents: by asserting that a data piece in the NEEM is an instance of a particular concept in the ontological knowledge base, the robotic agent can perform more informed reasoning and analysis on the data pieces. Third, the ontological knowledge base provides the infrastructure that is needed to (interdisciplinarily) combine knowledge and findings among the EASE subprojects. This is for example needed to analyze and compare how everyday activities are carried out in human and robot experiments.

Finding good solutions to expressiveness/efficiency tradeoff in ontologies and reasoning techniques

Subproject P02 researches the trade-off between expressive ontologies and efficient reasoning. To this end, it investigates the foundations of novel reasoning techniques that will make it possible to take full advantage of expressive ontologies in offline tasks while using (approximations of) the same ontologies online. These approximations can be considered a form of PEAM. Subproject P02 will also provide adaptations and approximations of concrete, existing ontologies for the purposes of EASE. It will work together with the Subprojects R01 and R04 to achieve this. The plan is to have an offline reasoner integrated in OPENEASE and an online reasoner as an expert reasoning method in the NEEM-based embodied knowledge framework to be provided by Subproject R01.

Subproject P03: Spatial reasoning in everyday activity Many everyday activities, such as finding objects in scenes or placing objects in certain arrangements, require spatial reasoning. The role of Subproject P03 is to investigate the qualitative spatial reasoning capabilities required for mastering of everyday activity and making them available. This project also transfers competence and results of the successfully completed CRC TR Spatial Cognition into EASE.

Investigating and combining two approaches for spatial reasoning

The research will be conducted by two PIs, who investigate the research topic from two viewpoints. PI Bhatt will investigate formal logic-based reasoning methods and develop models, algorithms and tools for reasoning about space as extensions to the constraint logic programming based CLP(QS) declarative spatial reasoning system. PI Schultheis starts from a human cognition perspective as formalized in computational cognitive models and researches cognitive principles underlying human proficiency in spatial reasoning in everyday activities. One focus will be on so-called *strong spatial cognition*, the replacement of computational effort from the central processor by direct manipulation. The two approaches are complementary and can have substantial synergies. They can be realized as different reasoning experts in the NEEM-based embodied knowledge framework to be provided by Subproject R01.

CLP(QS) is already integrated in OPENEASE and the KNOWROB knowledge representation and processing system, which is to be employed by EASE. This means that CLP(QS) can be loaded as additional Prolog modules into KNOWROB and be used by querying CLP(QS) specific Prolog predicates. We also plan to conduct experiments with robotic agents using the spatial cognition models and make the data and results available through OPENEASE.

Providing guarantees for behavior generated from heuristic methods

Subproject P04: Formalizations and properties of plans The role of Subproject P04 is to provide EASE with a robot plan verification environment that is able to check temporal properties of plans and helps to enforce disciplines of safe and robust plan development. It will also research the learning and automatic synthesis of correct plans from temporal specifications.

In EASE we want to investigate PEAMs to render the reasoning and decision making processes required for the mastery of everyday activity efficient. This will often be accomplished through heuristic methods that cannot be guaranteed to return correct solutions. To facilitate and accommodate the use of heuristic methods, Subproject P04 will develop and investigate methods for guaranteeing that robot plans satisfy essential behavior conditions. An example of such a condition is that in order to avoid damage, the robot should never turn or navigate away with an arm extended into a fridge or cupboard.

The plans that are investigated are plans implemented in CRAM (Cognitive Robot Abstract Machine (Beetz *et al.*, 2010a)), the plan language that will form the basis of the plans in the Subprojects R01 and R04. An important result of the Subproject P04 will be a formalization of the CRAM language that can be used for validation. Selected properties of subplans or simplified versions of the plans developed in Subproject R04 are to be validated such that decision making methods employing PEAMs can be ensured to generate safe robot behavior.

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

An advantage of Computer Science and Artificial Intelligence approaches is the ability to build computational models that test theories and hypotheses as a whole and observe the effects empirically. CS and AI methods are used to design, implement, and apply information processing principles to autonomous control and investigate how changes in the information processing mechanisms affect the capability of mastering everyday activity.

The goal of Research Area R is the investigation of a control framework including perception, learning, and reasoning mechanisms that enables robotic agents to master human-scale everyday manipulation tasks.

The goal of Area R is to investigate the information processing infrastructure necessary for robotic agents that can master human-scale everyday manipulation activities. This system will enable robotic agents to take vague task descriptions and use information about the task, situational context, and object context to perform the task appropriately. Inference of the appropriate action parametrizations has to be done without delaying plan execution. The information processing infrastructure will improve with experience (in the form of NEEMs) and by exploiting the routine and mundane character of everyday activity by making use of the PEAMs investigated in the Research Areas H and P.

The goal will be aimed for by achieving the following the following three interacting objectives:

Objective R.a: Investigating and constructing an embodied, NEEM-based knowledge system

This objective is concerned with how the comprehensive body of knowledge required for the mastery of everyday activity can be collected, abstracted into generalized (commonsense) knowledge and made actionable in the perception action loops of robot agency.

The embodied, NEEM-based knowledge system is to provide robotic agents with a comprehensive body of commonsense and naive physics knowledge. The core of the knowledge system will be a knowledge base that represents a large collection of NEEMs of experienced, acquired, and communicated episodes of everyday manipulation activities. This system will be integrated into the cognition-enabled control framework (Beetz *et al.*, 2012) and will specifically target the mundane, routine, and knowledge-intensive character of everyday activity. It uses different components developed in EASE incorporating NEEMs and PEAMs into the control framework.

Objective R.b: Developing perception and inference techniques for perception-based reasoning, grounded reasoning about object(-part)s, and simulation-based reasoning, and using these in ensembles

Mastering everyday activities requires perception and reasoning abilities that go far beyond the current state-of-the-art. For example, the robots will have to perceive stacks of plates, envision how changes in the way an action is executed influence its physical effects, predict flexible and robust execution, and automatically change the course of action when facing complications. This goes beyond the state-of-the-art in that the robots require the use of knowledge to perceive what is seen, need much more fine-grained and realistic action models, must predict perception-guided activity, and have to understand and automatically revise plans with sophisticated control structures.

Objective R.b will implement reasoning mechanisms for decision making using an ensemble of specialized reasoning methods that are integrated into a knowledge-processing framework especially designed for autonomous robot control (Beetz *et al.*, 2012; Tenorth & Beetz, 2013). Simulation-based reasoning methods will receive particular attention because of their potential for easily producing detailed answers that are difficult to derive using other reasoning methods.

Objective R.c: Realizing a plan-based control framework system for mastering complex everyday manipulation tasks and instantiating it for mastering household chores

This objective has several purposes: We intend to show the midterm feasibility of the visions put forward in this proposal and, more importantly, aim at a better understanding of how we can realize control systems with cognitive capabilities by building integrated experimental systems. We call them experimental systems rather than demonstration platforms because they will be used to test various hypotheses about computational models of cognition-enabled control, the ways computational problems should be phrased and decomposed, and the power and usefulness of the Cognitive Robot Abstract Machine (CRAM) investigated as part of Goal R. We consider a scenario that includes tasks such as setting tables, cleaning up, going shopping, and preparing meals. In addition, the robot should be able to automatically acquire new task skills, deal with novel objects, improvise, and learn preferences for task execution.

Structure of subprojects in the Research Area R Research Area R takes a more comprehensive view of the information processing models, as well as a more functional view on individual perception, reasoning, and plan-based inference techniques. In this area, we will study the embodied, generative information processing models for mastering everyday activity. Subproject R01 will investigate aspects of knowledge acquisition, representation, and reasoning that are needed for the knowledge intensive mastery of everyday activity. Subproject R02 will study cognition-enabled perception mechanisms for perceiving objects in everyday contexts. Subproject R03 will research different embodied reasoning methods including simulation-based reasoning, grounded objects-parts reasoning, and plan transformation for planning and learning. Subproject R04 will investigate integrated robotic agents performing everyday activities in human living environments, covering various aspects of housework. To this end, Subproject R04 aims at integrating individual results of the research areas into a NEEM and PEAM-enabled plan-based control framework. The integrated robotic agents will be based on software tools developed in this area as well as the other research areas. The analysis of the scenarios that involve interactions with humans will also serve as experimental data for the investigation of models of human everyday activities. Finally, Subproject R05 investigates a principled framework for the building and use of NEEMs and PEAMs for rapid learning of dexterous hand manipulation capabilities.

The software tools designed, implemented and investigated in Research Area R will serve as the implementational basis for the realization of the robotic agent for mastering housework. Research Area H will compare the performance of everyday manipulation activities by the robotic agents realized in Research Area R with the everyday activities performed by humans in similar settings, creating a feedback loop in EASE's research endeavor.

Human-inspired, hybrid knowledge system

Subproject R01: NEEM-based embodied knowledge framework The role of Subproject R01 is to provide EASE with an information processing framework for the collection, storage, management, and retrieval of NEEMs and the use of NEEMs to learn comprehensive bodies of everyday activity knowledge from them. This information processing will be:

- integrated and embodied into the existing cognition-enabled control framework (CRAM) in order to realize robotic agents that can collect NEEMs and learn commonsense knowledge bases from them (this integration is done under the lead of Subproject R04); and
- used stand-alone in combination with OPENEASE to mine knowledge from NEEMs collected from the observation of human everyday activities, and other sources.

The research focus is the design, realization, and investigation of a hybrid *symbolic/big data* knowledge system, inspired by models of the human memory system, that can construct commonsense knowledge bases from NEEMs. The knowledge bases will be constructed through collections of special-purpose learning methods. The learning process will be organized according to the unstructured information processing paradigm.

Subproject R01 teams up with R04 to realize the concrete control systems for the EASE scenarios using the NEEM-enabled control framework provided by R01. The reasoning methods and PEAMs investigated and developed in the subprojects of Research Areas R and P will be integrated into the framework. The collection of NEEMs and knowledge bases realized in R01 will be provided to the other subprojects.

Building expert perception methods from PEAMs, integrating into ROBOSHERLOCK

Subproject R02: Multi-cue perception supported by background knowledge Subproject R02 investigates the competent perception of household objects in typical household scenes. Examples of such objects and scenes are packed crockery and cutlery items in cupboards and dishwashers, made from textureless, reflective or even transparent material. To this end, Subproject R02 will research expert perception methods for these perception tasks and integrate them into the existing ROBOSHERLOCK perception framework (Beetz *et al.*, 2015b). ROBOSHERLOCK will then provide necessary perception functionality to the robotic systems, such as object detection and segmentation, texture- and form-based object recognition and reconstruction methods as well as knowledge-enabled reasoning for perception.

In order to accomplish the challenging perception tasks successfully, Subproject R02 studies and exploits PEAMs. It will enable robotic agents to learn the perception tasks they are to accomplish and the regularities of the scenes they have to accomplish the perception tasks in. The resulting perception module will be integrated as an expert perception method into the ROBOSHERLOCK perception system (which is integrated into the rest of the system by Subprojects R01 and R04). Subproject R02 will also draw from the results of Subproject H01, which studies the perception capabilities of humans in their everyday activities.

Subproject R03: Embodied simulation-enabled reasoning Subproject R03 investigates a particular reasoning method: simulation-based reasoning. Simulation-based reasoning is expected to be one of the essential inference techniques in EASE. Following the discussion in Section 1.2.5: Simulation theory of cognition, simulations could be a very powerful PEAM for reasoning about everyday activities; mentally simulating an activity and looking at it with the mind's eye is in some instances much easier than deriving the same results through other reasoning methods. Also, simulation-based reasoning is a powerful source for generating NEEMs.

Using simulation to reason about tasks

The basic idea is to translate reasoning tasks for robot control in the real world into simulation tasks, after which a physics- and rendering-based simulation module logs the low-level simulation into data structures from which the answers to the reasoning tasks can be determined. For example, when seeing a tray that the robotic agent has to lift, the robot asserts the objects it detects on the tray, their shape, pose, estimated weight, friction, supporting objects, etc. as models in simulation scenes and carries out a pick up action in the physics simulation. It can then use the simulation results in order to form expectations about whether or not objects will fall down when lifting the tray in that manner and decide whether some objects should be rearranged before picking up the tray.

The simulation-based reasoning methods from Subproject R03 will also be candidate methods for the investigations of the simulation-based semantics of action verbs researched in Subproject P01. Simulation-based reasoning will be fully integrated into the control framework and the specific control system researched in R04. The log files of the physics simulations will be transformed into NEEMs and provide the basis for learning commonsense and naive-physics knowledge.

Subproject R04: Specializing and optimizing generic robot plans The role of Subproject R04 is to investigate how robotic agents can improve the behavioral performance that their plans generate. This will be achieved through PEAMs by specializing plans and in particular the methods for answering the queries and accomplishing perception tasks. To this end, this area analyzes NEEM collections to identify PEAMs and exploits PEAMs to improve perception and reasoning. This is done by applying the methods for PEAMs that are contributed by other EASE subprojects.

Improving robot plans by exploiting available PEAMs

Key research challenges tackled by Subproject R04 are the design of plans that can be transformed to exploit PEAMs and methods for the transformation of plans and the specialization of reasoners and perception experts that can exploit PEAMs.

In addition, subprojects R04 and R01 will be responsible for providing the working control system for realizing the main EASE scenario described in Section 1.2.2. This control system includes the perception system ROBOSHERLOCK, the robot knowledge processing system KNOWROB, and the low-level manipulation control system that can generate competent grasps and motions from symbolic action descriptions (described in more detail in Section 1.2.10). Finally, R04 will also generate the NEEMs from robotic agents accomplishing the scenario tasks, which will be used in other EASE subprojects.

Integrating components of complete control system

Subproject R05: Episodic memory for everyday manual activities During the human brain's evolution, its information processing capacity had to increase drastically in order to meet the challenge of successfully performing more sophisticated and complex object manipulation tasks, in particular ones that require reasoned decision making. A number of researchers argue that

Representing and learning complex object manipulation using NEEMs and PEAMs these improved capabilities are coupled with other evolutionary developments such as representations of actions (Rizzolatti *et al.*, 2001) and the co-development of action and language (Arbib, 2006). Consequently, hand manipulation tasks in the context of everyday activity are at the core of EASE's research program.

Subproject R05 investigates the NEEMs and PEAMs in object manipulation with the hand. It will develop and investigate a framework that combines symbolic representations of knowledge about manipulation tasks and subsymbolic representations of procedural knowledge about low level manipulator control, by bridging them with the help of episodic memories collected from a robotic agent's own experience and human demonstrations it observes.

To this end, we let people execute object manipulation tasks in a physics simulation-based virtual environment setting and track their hand poses with a high accuracy, high-speed marker-based tracking system. The observation system generates NEEMs from these observations that allow us to relate the hand manipulation strategies to the context including the object to be manipulated, the task to be performed, and the situational context of the manipulation scene. The observed manipulation strategies are generalized and transferred to a two-hand upper-body robotic agent that applies the learned strategies to its own manipulation tasks. This generalization and transfer of observed manipulation skills are heavily based on PEAMs. A number of studies in the Cognitive Neurosciences showed that, notwithstanding the complexity of the human hand, a few variables are able to account for most of the variance in the patterns of human hands configuration and movement (Bicchi & Kumar, 2000).

Therefore this subproject will also look at "episode mirroring": methods to transfer manipulation knowledge between agents with structurally similar body schemas (from human to robot as indicated above, and between different robots).

Cooperation between research areas and subprojects As can already be inferred from Figure 20 and 21 at the beginning of this section, there is considerable cooperation and interaction between the different research areas and subprojects. The organizational structure of EASE and its research plan inherently calls for cooperation between the subprojects. We will outline these interactions in more detail below. In addition, EASE proposes several measures for strengthening cooperation. These measures include the establishment of cross-area research topics and integration workshops, a series of common colloquia, means for successful research training, the establishment and intensification of international cooperations, support through cooperative research projects, and the training of young researchers.

Subproject interactions Table 2 sketches how the proposed subprojects combine to realize EASE's overall vision and how the individual subprojects cooperate in the overall research plan.

The letters show which research outputs are produced by the projects listed in the rows and consumed by the projects in the columns. Four categories of output are defined: (1) data from observations and experiments (**D**), (2) general representations and models of everyday activities (**R**), (3) methods and algorithms (**M**), and (4) implemented software components (**S**). As can be seen from the table, the projects in Research Area H primarily produce data of humans performing everyday activities and models/representations. The projects in Research Area P produce models, as well as algorithms and implemented software. Research Area R is mainly concerned with the realization of information processing models for everyday activities in robotic agents. It produces software components and methods. Additionally, it produces data from the robot agents using the models and executing everyday activity. Subprojects R01, which investigates the embodied knowledge system, and R04, which will also most closely work on the integrated systems, deserve particular attention. Subproject R01 provides the implementational framework for all projects in Research Area R and a methodology-oriented view on reasoning problems investigated in Research Area P. Finally, we see that most projects directly contribute to the realization of the integrated robotic agents that R04 is responsible for. In return, R04 feeds

H: mainly produces data from humans and models

P: produces models, algorithms and software

R: produces robot data, software and integrates a complete system for robot agent

	▶H01	▶H02	▶H03	▶P01	▶P02	▶P03	▶P04	▶R01	▶R02	▶R03	▶R04	▶R05
H01▶		D	D,R			D	D		D	D,R	R	D,R
H02▶	D		D	D,R						R	R	
H03▶	R	D				D		R		R	D,R	R
P01▶		R					R	R	R		R	
P02▶				R				M	R			
P03▶				S				S				
P04▶								S			R,S	
R01▶					R				S		S	S
R02▶						S					S	S
R03▶				S				S			S	S
R04▶						D	D					S
R05▶	D		D					S			S	

Table 2: Matrix of collaborations among the EASE subprojects. The rows list the research output produced by that respective project. The columns list the projects that consume these outputs. The output categories are **data** from observations/experiments (**D**), **representations** and **models** of everyday activity (**R**), **methods** and **algorithms** (**M**), and **software** components (**S**).

back the collected NEEMs from execution into the other EASE subprojects in order to provide experimental data (in cooperation with project H03) and problem statements for guiding research and for increasing its impact.

Cross-area research topics   EASE further promotes research in the individual research areas to be synergistic with other research areas by introducing cross-area research topics as an orthogonal dimension of project organization. The following cross-area research topics will be pursued from the start of the CRC:

- **Narrative-enabled episodic memories (NEEMs)**, coordinated by PI Bhatt, will discuss and integrate different perspectives on the concept of NEEMs investigated in EASE. Different areas and subprojects are likely to have a different perspective on the requirements for NEEMs and how to use them. This is to stimulate the sharing and use of results between areas and cooperation across subprojects.
- **Pragmatic everyday activity manifolds (PEAMs)**, coordinated by PI Zetzsche, will investigate different kinds of pragmatic manifolds that can potentially simplify the computational complexity of mastering everyday activity.
- **Machine Learning and statistical/Bayesian methods**, coordinated by PI Cheng, will discuss recent progress in the area of Machine Learning including deep learning, Probabilistic Reasoning and Learning, and how this can be applied in the different areas of EASE.

The researchers within these cross-area topic groups will meet regularly and exchange information about their ongoing work, possible synergies, possibilities for software integration, and interesting papers and research trends they observe. In particular, in the areas PEAMs and NEEMs, the participants are expected to document the conceptual and implementational aspects in technical reports. We also encourage the topic groups to invite leading researchers outside of EASE to join these meetings.

Integration workshops   EASE will also organize yearly integration workshops to foster integration and exploit synergies. Here the goal is to promote close cooperation within EASE on research and software. Members of the CRC are invited to team up to tackle a system research topic. For example, they are encouraged to do joint experiments in the intersection of the CRC on a joint code basis and publish a joint research paper. The CRC will support these workshops through a social event, pizza deliveries to the lab, and by sponsoring a conference visit by the organizer once the respective paper gets accepted.

The rationale behind these integration workshops is that researchers in different projects can at least partially work together on a common code basis and thereby exploit long-term synergies. Integration workshops might also stimulate joint project proposals. Integration is further facilitated by the principal investigators committing to use common, open-source software frameworks including ROS (Robot Operating System), PCL (Point Cloud Library), and OpenCV (computer vision library).

What is not on EASE's research program In order to retain feasibility and focus, we consciously exclude certain research topics that are relevant to EASE, but would broaden the CRC too much to be effective. For this reason, the following topics will not be within the research focus of EASE. They will be covered either in prominent collaborative projects and/or through national and international cooperations with research partners. These topics include:

> *I'm as proud of what we don't do as of what we do.*
> *— Steve Jobs*

Hardware. EASE does not dedicate resources for research on robot hardware and low-level interfaces. New innovations in sensing and robot hardware will be integrated into the EASE robot platforms by the technical support project F (System Integration). The integration of better sensing and actuation technology is a support activity for the promotion of EASE's primary research efforts.

Behavior-based and developmental robotics. The investigation of behavior-based and developmental information processing models (Cangelosi & Schlesinger, 2015) will not be covered in EASE. Because of its unique strength in symbolic representation and reasoning methods, EASE will focus research on other types of information processing models. Specific ideas from behavior-based and developmental agency will be employed in the context of grounding and embodying reasoning, in particular within the Subproject R03.

Experimental psychology of perception and action in constrained settings. The EASE main scenario was made to reflect natural conditions. Experimental psychology studies in highly constrained settings lie outside the core focus. EASE intends to complement research concerned with specific psychological hypotheses under highly constrained settings with a whole-system approach and complex activities in natural settings. The experimental psychology view will be integrated into EASE's research activities through the organization of an international workshop, such as a Dagstuhl seminar, aiming at different psychological views of everyday activity, and by inviting and cooperating with international experts. In addition, we will invite leading researchers in experimental psychology of human activity for research stays in Bremen. Current and past collaborations include Bernhard Hommel in the RoboHow project, Joachim Hermsdörfer in CogWatch, and Heiner Deubel and Michael Zehetleitner in the context of CoTeSys. Werner Schneider at CITEC also complements the experimental psychology expertise of EASE.

Outdoor robotics. EASE will limit its scope to robot applications in indoor human living and working environments. While outdoor scenarios can certainly contribute much to realistic everyday settings, they are beyond the core focus of EASE. Research on the topic may be conducted through cooperative projects, such as EU IP SHERPA and a number of projects conducted by the DFKI Innovation Center Robotics.

Later phases of the program To keep goals clear and attainable, we find it important to clearly define the agenda for each phase of EASE. Therefore areas that are considered important may perhaps only gain attention later on, when their success can be heavily increased through the EASE output from earlier stages. **Human-robot cooperation and interaction** is one such area. Interacting with humans is a very important aspect of mastering daily activity. Not only do we expect these interactions to benefit from the routine nature of everyday activity, the solid understanding of everyday activity and common information processing framework acquired in earlier stages of EASE are a good basis for effective communication with other agents about these everyday activities. In order to understand the intentions of humans and make one's intentions understood naturally in a certain task, it is highly desirable for robots to have the relevant knowledge that people commonly implicitly assume other human to have, available. Therefore, interactions with humans will be the focus of the last stage of EASE. EASE can also profit from research results of the Excellence Cluster CITEC, with which EASE would maintain an active collaboration and which has human-robot cooperation and interaction as one of its research foci.

1.2.8 Infrastructure for collaboration

EASE will be a collaborative, multi-disciplinary research center that is software, data and knowledge intensive. The subprojects in Research Area H collect, interpret, and analyze data of humans performing everyday activity using various tools designed to deal with various data such as human language, muscle and brain activity, body motions, texts, data from game engines, and data statistics. The subprojects of Research Area R realize robotic agents that perform activities studied in Research Area H, and generate experience data from the plans they execute, including control signals, perceived objects and scenes, action effects. The subprojects in Research Area P investigate abstract, often logic-based or probabilistic models of the activities studied in Research Areas H and R.

The research areas and even the subprojects within the same research area will likely use specialized software packages written in different programming languages and different data formats to conduct their research tasks. Many research outputs in EASE depend on comparing and transferring data and models across research areas however. This requires the integration of data, data formats, knowledge bases, procedures, and software libraries from different communities and making them semantically accessible using a common conceptual apparatus.

The research in EASE is data and knowledge intensive. Data and knowledge are to be collected, shared, and used in all different research areas. Part of the everyday activities that are used to study human everyday activity (Research Area H), are also tasks to be performed by the robotic agents (Research Area R) and include representation and reasoning problems that are investigated in Research Area P. To facilitate common research based on research data generated in different EASE projects, we will use a common data and knowledge infrastructure and use a common platform to provide semantic access to all the research data.

For these reasons, the software infrastructure of EASE, including the architectural constraints, data models, and representation languages, will play an important role in its success. The strategy adopted by the EASE CRC is to put little constraints on the software in the individual projects but require encapsulations in agreed-upon open-source software libraries and provide access through adopted interlinguas.

Common software The agreed-upon standardized implementation platforms are the open-source robot middleware library ROS (Robot Operating System), OpenCV as the supported open-source machine vision library, and PCL (Point Cloud Library) for RGB-D image processing and vision. If subprojects use or develop other software components, these components are to be encapsulated to provide interfaces compatible with those of the above software libraries. The PIs agree upon holding EASE-wide integration workshops in which the required software components are encapsulated to be usable using the libraries mentioned above.

With regard to data, EASE will make use of noSQL (not only SQL). NoSQL databases support the high volume storage and access of unstructured data as in relational database systems and can support document structures, graph structures, tuples, and key value stores. The open-source Mongo[25] database system has been successful in storing of and working with various data for OPENEASE tests. In EASE these database systems will be mainly used for the storage and retrieval of log data from robots and observation data from human experiments.

Common knowledge At the knowledge representation and reasoning level, the agreed-upon software platforms are SWI-Prolog and UIMA (Unstructured Information Management Architecture). SWI-Prolog will be the interlingua between different representations and reasoning methods employed in EASE. SWI-Prolog also provides an extension for OWL (Web Ontology Language), which allows for the alignment of predicates with agreed upon ontologies.

Reasoning methods as procedural attachments

One way of integration is to include reasoning methods as procedural attachments of predicates that compute the truth values of instantiated formulas using specialized algorithms. With other words, any type of procedure or method can be wrapped inside predicates to interface with the common knowledge representation scheme. The resulting knowledge base can be defined in terms of data and robot control structures rather than only being defined in terms of other knowledge. The procedural attachment integration enables individual projects to integrate other forms and representations as long as the methods can be invoked as a Prolog predicate that is instantiated with the result of the invoked inference process.

UIMA for integration

A second way of integrating methods from the individual projects is using UIMA (Unstructured Information Management Architecture). UIMA, the middleware that was used for the Watson system, is the only standardized framework that supports the development, implementation, composition and deployment of multi-modal analytics for the analysis of unstructured information and making the information semantically accessible. EASE data will include unstructured data such as natural language texts, sensor and pose data streams, and images. The Apache open-source implementation of UIMA that will be used in EASE includes APIs and tools for encapsulating reasoners that combine knowledge bases and inference methods as analysis components.

EASE uses ontologies that are explicitly contained in the EASE knowledge bases to model concepts. EASE will make use of multiple ontologies, starting with two main ontologies. Other ontologies can be added to the knowledge base on demand, but might have to be encapsulated into modules of the UIMA architecture if concept definitions are incompatible with the main ontologies.

WordNet ontology to assisst natural language understanding

First, the WordNet ontology (Fellbaum, 1998), which is a dictionary knowledge base created and maintained by linguists.

It contains the words one typically finds in dictionaries and includes all their alternate meanings (called synsets). The synsets are arranged as a tree-structured taxonomy that captures the relations between different words and word senses. The ontology is designed for natural-language applications and achieves a high coverage of English words and their possible meanings, but does not contain much background knowledge in any other form. Using the taxonomy, researchers have defined different measures of the semantic distance of two word senses based on the relative positions of the word senses in the WordNet taxonomy, such as the Wu & Palmer (WUP) distance (Wu & Palmer, 1994). While the relational knowledge about the word senses is very limited, it can still prove very useful for an agent acting in an open world. If a robot encounters a new word or concept that it has no knowledge of, it can use the WordNet ontology to relate the new word to concepts it has deep knowledge of (for example, the semantically closest concepts in the knowledge base).

KnowRob ontology to describe robotic concepts

Second, the KNOWROB ontology, which extends the core ontology of OPENCYC with robot and manipulation specific concepts (Tenorth & Beetz, 2013). It is a knowledge base designed for robots.

[25] `https://www.mongodb.com/`

It is much smaller than WordNet, but it contains a much richer relational structure and knowledge of concepts. The knowledge stored about objects goes beyond textual knowledge. For example it can include images or CAD models of objects, sensor data, methods to compute relations and attribute values. Concepts can also facilitate the anchoring of the representational structures in the control system. For example, concepts might include perceptual descriptions of object classes that the robot perception system can use in order to detect, localize and model instances of the object class. It could also specify the motion constraints that have to be satisfied in order to manipulate the respective objects successfully, such as the constraint of holding open containers horizontally. Some concepts also have associated methods that dynamically compute their instances from the data structures of the robot control system. For example, the 3D pose of the robot might be associated with a method for computing it as the global maximum of the probability distribution over robot poses computed by a probabilistic localization algorithm.

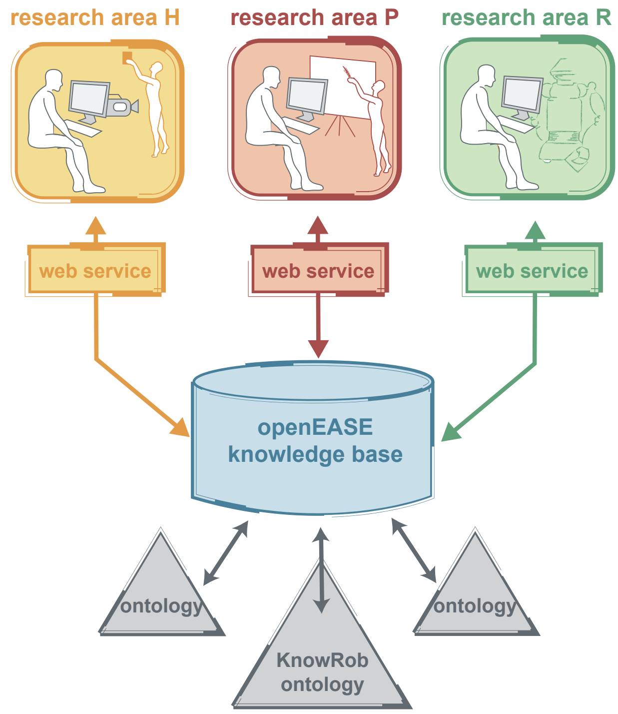


Figure 23: OPENEASE as the common knowledge service used by the EASE CRC.

Access any data using one language

Prolog predicates can be used to access both the concepts in the KNOWROB ontology and subsymbolic data from "big data" databases. These rules also make it possible to make observation data from human activities and experience data from robot activity simultaneously accessible and connect the data through a common ontology. A more detailed description regarding the implementation of knowledge representation and processing in EASE is given in Section 1.2.10.

The EASE knowledge base can be structured into sub knowledge bases; collections of concepts and facts typically pertaining to one particular knowledge topic. These sub knowledge bases can contain specialized reasoning capabilities, such as the spatial reasoning capabilities developed in Subproject P03. A sub knowledge base could even be associated with a particular query in a particular plan. If it targets a specific query, it can exploit all PEAMs of the query and the query context in order to make answering fast and accurate.

EASE does not require all knowledge bases to be consistent. Indeed, acquiring consistent knowledge bases from the robot experiences generated through uncertain perception and manipulation capabilities is in many cases simply not possible. Instead, we require consistency only for the answers returned by the reasoning services. Such mechanisms have been successfully applied in systems like Watson. This played a key factor in it being able to successfully scale question answering capabilities to huge and open query domains. One of the key methods here is to generate many answer hypotheses and assess them afterwards. Of course, if answers are available from reasoning mechanisms that are logically proven to be correct, they take precedence over results generated by other methods.

Unique through comprehensiveness and accessibility of available data and methods

Cooperation using OPENEASE EASE intends to use a common, overarching platform to facilitate cooperation between subprojects and with partners world-wide. OPENEASE is to be the common database, knowledge base and management system of the EASE CRC (Figure 23).

OPENEASE is a powerful software research tool for EASE due to its uniqueness with respect to (1) the comprehensiveness with which real execution data of modern autonomous manipulation robots are logged, stored and made openly accessible to researchers; (2) the representational infrastructure that is provided to make very inhomogeneous experience data from different robots and even human manipulation episodes semantically accessible in a uniform and standardized concept vocabulary; and (3) the suite of software tools that enable researchers and robots to interpret, analyze, visualize, and learn from the experience data.

Diversity of data available in OPENEASE

The data generated and used by EASE can be stored in OPENEASE with a common ontology. The representational infrastructure provided by OPENEASE covers the data and knowledge needs of EASE. For example, it already contains diverse knowledge such as logs of robotic agents performing human-scale manipulation tasks and motion data of humans setting the table. An exception is the biosignal data generated by Subproject H03. The work by EASE on extending OPENEASE will be limited to enabling the support of new data types and associated methods and visualizations where necessary. The web-based software workbenches can be individualized to fit the needs of the particular subprojects.

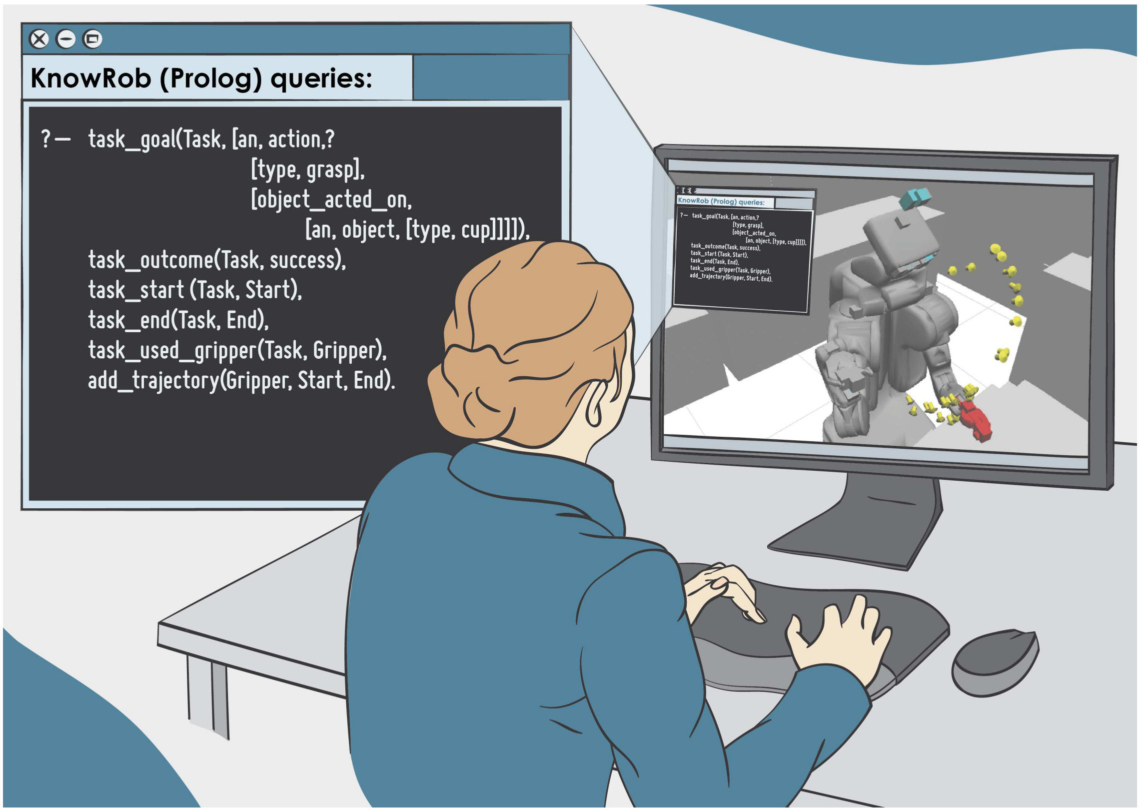


Figure 24: Visual results of a query on logged robot experiences showing the robot and the trajectory of the gripper during a certain task: let Task be a task of the robot in which it intended to grasp an object of type cup, Start and End be the time instances where this task started and ended respectively, and Gripper the gripper of the robot that was used in the task. Display the trajectory of the gripper between time instances Start and End.

OPENEASE data accessible with Prolog predicates

The knowledge bases and (subsymbolic) databases in OPENEASE can be queried symbolically using a Prolog interface. An example query and visualization thereof is shown in Figure 24. Researchers can use OPENEASE through an interactive, graphical web interface to access and visualize the knowledge and data contained. Robots can also connect to OPENEASE directly via a webservice API. This enables them to use OPENEASE's knowledge to provide semantic meaning to their sensor data and data structures. EASE can extend this system to let robots upload their own data structures and execution log files, declare new Prolog rules, and thereby add new knowledge to the platform.

The data and knowledge bases in OPENEASE are linked to concepts in the KNOWROB and WordNet ontologies, as mentioned in the previous subsection. These ontologies define taxonomies of concepts and background knowledge for the individual concepts. Common ontologies for the individual knowledge bases are of key importance to the EASE research project, as common definitions of concepts relevant to different subprojects, system functionalities, and agents (humans and robots) achieve consistent concept use throughout the project and an easy transfer of results across subprojects.

Powerful reasoning tools

OPENEASE contains reasoning tools for first-order deductive reasoning, probabilistic reasoning, procedural reasoning, simulation-based reasoning, temporal reasoning, CAD-based reasoning, and geometric reasoning. These can be extended and used by EASE to equip its robot with various reasoning capabilities. Moreover, OPENEASE includes a host of other, (open-source) data analysis and machine learning tools to facilitate empirical investigations, such as the Weka machine learning library (Hall *et al.*, 2009), the Caffe deep learning toolbox (Jia *et al.*, 2014), and the statistics software package R (Hothorn & Everitt, 2014).

The data, knowledge, reasoning, and learning tools can be used by researchers to analyze data and to generate hypotheses or answer them. The common platform makes it more straightforward to use data provided by other subprojects and work/test new tools together. This is also true for the realization of robots performing the actions; OPENEASE can be accessed directly by robots. They could use the knowledge bases and reasoning methods developed and provided by EASE for performing the everyday activity using OPENEASE as a remote knowledge representation and processing service.

Finally, OPENEASE will make it easier to share results with the international research community. Experimental data and results can be published on it, allowing partners to use the web interface to conduct research using EASE data. Moreover, we can use the visualization methods provided to generate videos, statistical diagrams, etc. to illustrate the work done in EASE.

Four EU research projects already using OPENEASE

Reviews of four European research projects (ROBOHOW, ACAT, SAPHARI, and SHERPA) have been made available to the reviewers through the OPENEASE web interface. The EASE principal investigators commit to the use of OPENEASE as the common platform for research cooperation in EASE.

A multi-day tutorial and workshop for the use of OPENEASE for EASE research will be conducted within the first six months.

Integration between in the EASE subprojects is not only a research policy induced by the principal investigators. The researchers of the CRC have substantial incentive to integrate and make their research accessible in OPENEASE, because OPENEASE will

- allow researchers to use the data of others with little effort and thereby make cooperation much easier.
- enable the researchers to illustrate their results visually, for example by automatically generating videos and high-quality graphics from experiments.
- support both semantic processing and quantitative analyses. It supports the use of software toolboxes such as statistics package R[26] (R Core Team, 2016) within OPENEASE.
- increase visibility of researchers and support dissemination, as a central platform that publishes research activities and experimental data online.

[26] The R Project for statistical computing, `//www.r-project.org/`

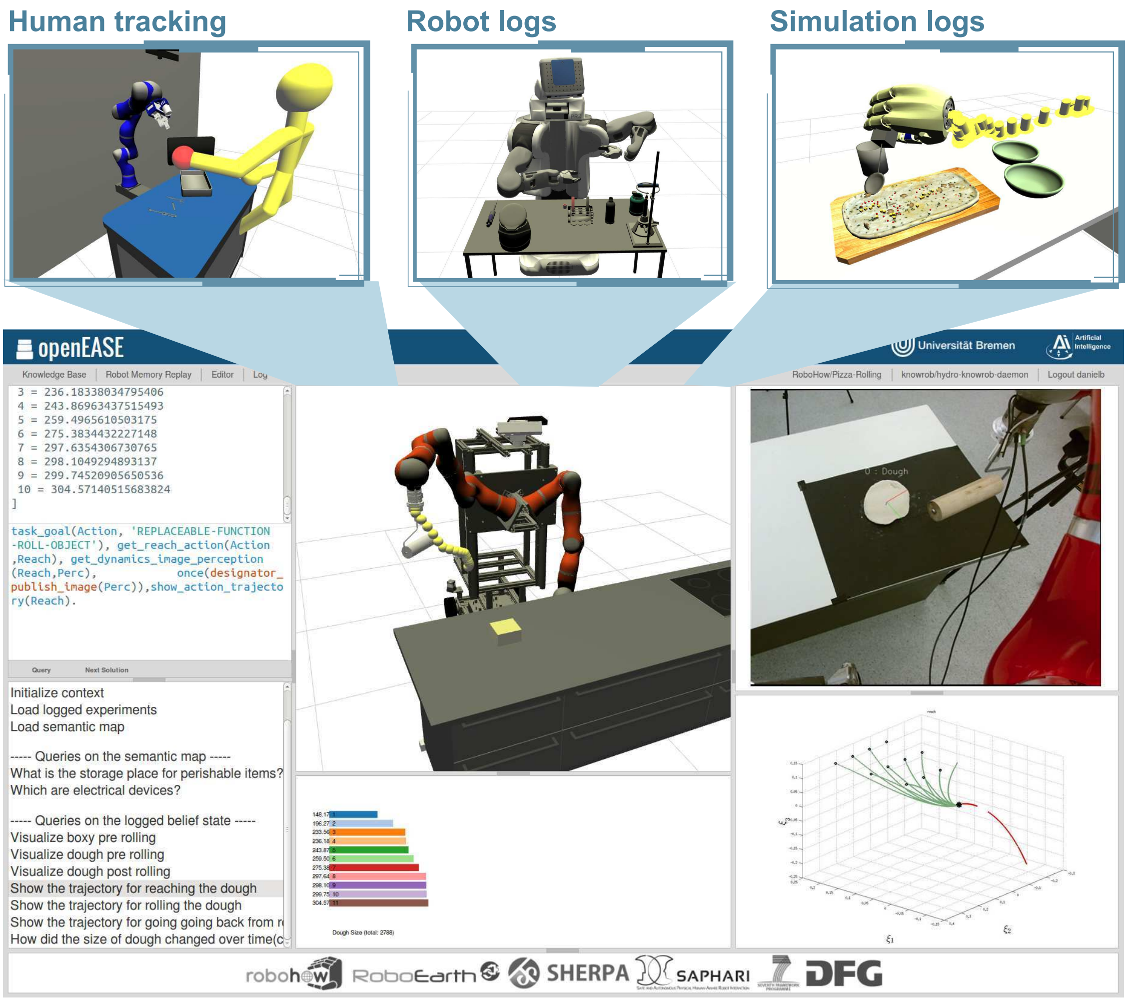


Figure 25: OPENEASE interactive web interface for an experiment involving rolling out pizza dough. The top left pane shows the command area, where queries can be entered and results are returned. The bottom left pane shows a library of default queries. The top middle pane shows a visualization of the robot, trajectories, and object position during a (sub)action. The bottom middle pane shows some statistics regarding the action. The top right pane shows the image the robot saw at a specific timepoint. Finally, the bottom right pane contains a visualization of (learned) trajectories associated with this subaction. This interface works for data from many sources, including human tracking, robot logs and simulations.

A first version of OPENEASE is already operational and provides online tutorials to get started. The graphical user interface, with command, visualization, example queries, statistics, images and detail panes, is shown in Figure 25.

The current version of OPENEASE can be tested on `http://open-ease.org`

1.2.9 Collaborating institutes

Cooperation with CITEC Besides the contributions to Subproject R05, Prof. Ritter, the director of the excellence cluster CITEC (Cognitive Interaction Technology) at the University of Bielefeld (UBI), as an EASE principal investigator will open up promising cooperation opportunities with CITEC. Specifically, EASE can directly profit from results from three highly relevant CITEC research areas: *Motion Intelligence*, *Memory and Learning*, and *Situated Communication*. EASE research results and plans will be communicated early in CITEC research meetings and we will proactively generate research cooperations between the two Collaborative Research Centers. We also plan focused meetings of subgroups in specialized research topics that lie in the research focus of both institutions. These forms of cooperation are time and cost efficient as public transportation between Bremen and Bielefeld is excellent. The research faculty of CITEC can also complement the expertise in EASE. For example, CITEC employs leading researchers in the areas of Cognitive Psychology of action (Schneider), cognitive architectures of human movements (Schack), ontology-based interpretation of natural language and processing big unstructured data (Cimiano), social cognitive systems (Kopp), and others.

Exchange of results between EASE and CITEC

Vice versa, we expect that EASE research to complement and contribute to CITEC in important ways. First, the main EASE scenario of an autonomous mobile robot performing human-scale manipulation tasks that are stated abstractly is highly relevant for CITEC research but so far there are no plans for realizing such a scenario within CITEC. Having access to the fully realized scenario will open up new research opportunities for the CITEC research areas listed above. Also, the research faculty in Bremen complements the one at CITEC by bringing in in-depth competence in additional, highly relevant research areas such as symbolic Artificial Intelligence for agent control, Spatial Cognition, and Robot Perception.

The efforts to build a solid basis for cooperation are part of EASE Subproject R05 and explicitly stated in the work plan (work package 1). Here, we spend the efforts to integrate the software frameworks used in CITEC and the University of Bremen (UB) in a coherent system that runs the EASE high-level control on top of the CITEC software framework.

Integration with CITEC is part of Subproject R05

The purpose is to have a reference implementation running at CITEC that includes plan-based control mechanisms, the knowledge representation and processing infrastructure (KNOWROB), and graphical access through OPENEASE. Upon interest, this reference implementation can then be transferred to other integrated agent systems.

The cooperation will be strengthened by the doctoral students in the subproject having the second principal investigator as a co-advisor, the doctoral students working together in the EASE software integration weeks, extended (one week long alternating visits once a year), and additional day long visits of the PIs and researchers. The doctoral students at CITEC will become a member of the CITEC graduate school.

Cooperation with DLR Besides adding competence by providing the expertise and technology of some of the most advanced motion control techniques that are necessary for the competent execution of everyday manipulation, the DLR institute Robotics and Mechatronics (DLR-RM) will be a strong collaboration partner for EASE. To this end, DLR-RM and UB have founded a common virtual laboratory with the name Perceptive Autonomous Agents Laboratory (PAAL). The PAAL Laboratory focuses on joint research in the areas of knowledge enabled robot manipulation, robot perception, and safe physical human robot cooperation. The working of the cooperation can already been seen in a number of joint publications covering all cooperation areas (Leidner *et al.*, 2015; Birbach *et al.*, 2015; Beetz *et al.*, 2015c).

DLR-RM and UB founded Perceptive Autonomous Agents Laboratory

The forms of collaboration includes the development of joint research project proposals, DLR-RM research scientists teaching at UB, supervision of doctoral students at DLR-RM, and support for the exchange and visits between the institutes.

Ongoing integration of motion and plan-based control systems

The main software components for motion control (DLR-RM) and the cognition-enabled plan-based control system have already been integrated as part of the EU FP7 project SAPHARI and received excellent project reviews. In addition, we have developed a preliminary version for semantically annotating experience logs generated by the DLR-RM control system such that they can be represented with KNOWROB and processed with OPENEASE.

Extensive collaboration

DLR-RM will contribute the following technologies and expertise to EASE: impedance-based robot motion control, object manipulation capabilities, acquisition of environment models, deep learning based robot perception. EASE will contribute to DLR-RM research in the areas knowledge-enabled robot control, knowledge processing for robots, fast object tracking, and learning from robot experience.

The doctoral student at DLR-RM is expected to become a member of the EASE IRTG or the TUM graduate school of science and engineering. The doctoral student will participate in the EASE integration workshops and the researchers will have yearly alternating one week research stays at the partner institution.

Cooperation with TUM The cooperation with TUM is based on bidirectional cooperations with individual research groups that we cooperate or intend to cooperate with. The groups include the Institute for Cognitive Systems, Sensor Based Robotic Systems and Intelligent Assistance Systems, Robotics and Embedded Systems and the Institute of Automatic Control Engineering. The cooperations add additional competences to EASE in the research areas of Intelligent Cognitive Systems, Robot Perception, Robot Control, Imitation Learning and Human-Robot Interaction.

1.2.10 Realization of robotic agents

EASE intends to bring together people, methods, and theories from different fields to advance our understanding of information processing models underlying the mastery of everyday activity. To validate the results as a coherent whole, robotic agents that can perform the human-scale manipulation tasks contained in the main scenario (Section 1.2.2) over extended periods of time are necessary.

Mastery of everyday activities imposes high requirements on robotic agent hardware

The mastery of everyday manipulation activities such as cooking, implies hard requirements for the robot hardware and software. Steps that require delicate manipulation are especially difficult, because the movements need to apply small but purposeful forces, and react in real time according to the interaction between the tool and the manipulated material. Examples of such actions are: peeling or cutting vegetables, grabbing a single slice of ham out of a package, or pressing a button to switch on the mixer.

To make those difficult actions possible, it is important to provide very advanced robot hardware, control algorithms and high-end robotic perception. The better the available infrastructure, the broader the possible variation in manipulation actions.

Using existing, available advanced robotic hardware and extending existing software libraries

The main role of hardware in EASE is to enable EASE researchers to gather robot experience data and test developed methods and systems. Therefore, EASE is based upon existing, leading edge platforms and extends existing software libraries. Achieving the ambitious goals of EASE while having to start from the ground up would be undesirable, unrealistic, and unattainable. The EASE consortium has access to existing high-end robotic platforms and other equipment that makes it possible to conduct experiments and gather long-term robot activity data from the beginning. Here we will describe some of the robotic platforms and integrated control systems available to show that the work on EASE research goals can start almost immediately.

During the project, the know-how of the members of EASE in building and maintaining leading edge mobile manipulation platforms and libraries will be used to design and build common robotic platforms for EASE. These platforms will improve upon and substitute the existing platforms. Three identical platforms are planned to be constructed to ensure continuous, long-term operation.

Jump starting EASE Realizing the main scenario requires access to autonomous robot manipulation platforms that satisfy the hardware requirements of the scenario. It also requires leading edge software libraries that support subsystems for object and scene perception, knowledge representation, plan-based control, and motion control. These components have to be connected to the sensors and actuators of the robot and tightly integrated with each other. The EASE consortium already has such systems to built upon. This enables us to jump start EASE research and focus on the core research questions and goals.

Robotic platforms From the start of the project, two high-end mobile manipulation platforms with humanoid upper bodies will be available for studying the main scenario. These manipulation platforms are a PR2 robot from Willow Garage and “Boxy”, an in-house designed robot that integrates leading-edge manipulation mechatronics and sensors (Figure 26). Note that here we only discuss the hardware available for studying the main scenario in the common EASE laboratory. These platforms are by far not the only ones to be used for EASE research.

Both robots use the ROS middleware and share a significant amount of software infrastructure. The low-level driver layer, which communicates to the hardware, is tailored to each robot but offers similar interfaces for higher-level systems.

The PR2 is a standard platform for autonomous mobile manipulation that is used by many of the internationally leading researchers in robotics. This will allow EASE to cooperate with these laboratories with minimal efforts.

Figure 26: Mobile manipulation platform Boxy available for EASE from the beginning of the project.

The second platform, “Boxy”, is mobile manipulation platform with a high-precision omnidirectional drive. It has two KUKA LWR-4+ lightweight robot arms, one of today’s most advanced robot manipulators. It has a variety of end-effectors, from simple parallel grippers to humanoid 5-fingered force-controlled DLR/HIT hands, which can be exchanged depending on the task. The control software for the arms is provided by DLR-RM in a cooperation with UB. The robot manipulators can be equipped with 6-DOF force/torque sensors (KMS-40 sensors from Weiss Robotics) at the wrist, which can measure the interaction forces between the tool on the robot and the manipulated objects at high frequency and low noise. Boxy is equipped with a high-precision robotic platform on Mecanum wheels, with BLDC servo motors, planetary gears, and a direct drive design. It can carry up to 400kg, is fully holonomic, and controlled at a high frequency using an EtherCAT bus and ROS middleware.

Beside high-end navigation and manipulation the robot is also equipped with an imaging sensor suite mounted on a robot arm that functions as the neck. To facilitate development of perception systems across different robots, the Institute for Artificial Intelligence (IAI) laboratory at UB uses a standard sensor setup that can be installed in all the robots that are used. Perception algorithms can easily be used in several robots if they share the perception hardware. Currently, the multimodal robot sensing package includes a Kinect-2 RGB-D sensor, a thermal camera, and a high resolution color camera. All the cameras are calibrated intrinsically and also calibrated to each other (extrinsics), making it possible to relate points from one modality to another: for example, to have a point cloud that has color and temperature information for each point. The hand-eye calibration also has an absolute position error under 2mm: from a position of points detected using a calibrated Kinect-2 sensor to reaching that same position using our calibrated KUKA LWR-4 arm.

Many existing tracking, scanning, and supporting equipment, and associated know-how can immediately be used

Other equipment The robots will be deployed in a simple apartment setup consisting of a kitchen and a dining room area. The state-of-the-art facilities will include power and computational infrastructure to conduct long-term robot experiments and capture human motion. The robot laboratory at IAI can accommodate up to 15 workspaces and has good networking infrastructure, multimedia projectors, 120cm presentation monitors, and powerful computer workstations. Some of the available tracking and sensory equipment will be highlighted below.

The experiment space includes a state-of-the-art infrared optical tracking system (OptiTrack) with 12 cameras installed on the ceiling of the robot laboratory. It tracks objects that have been marked with infrared markers at a sampling frequency of 125Hz and is complemented with a full-body suit for tracking human movements in the same environment as the robot. It can be used to complement on-board perception of the robots, or be used while the on-board perception system is being developed.

Another available method is the Xsens human motion tracking system, which tracks full-body motions using an array of inertial measurement units (Xsens MVN) on a wearable suit. Its high quality makes it applicable even for biomechanical diagnostics. This system can capture the movements of a person under difficult lighting conditions, like outdoors, where vision-based systems would be problematic.

The VirtuSphere is an oversized, nearly three meters high spheroid hamster wheel for humans. The sphere is mounted on rollers which allow it to rotate freely in any direction, thus allowing for experiments in which the test subjects can move around in the virtual world by simply walking. A special head-mounted display worn by test subjects tracks their motions and displays images of the dynamic virtual environment. The sphere itself is also equipped with an array of sensors that track the test subject's motions while being inside the sphere. It enables us to study aspects of cognition like spatial perception, reasoning and action inside a virtual space.

There is also ample 3D-Model building equipment, including two hand-held state-of-the-art 3D scanners (AmeTek GoScan), which are used for industrial 3D-model creation, reverse engineering, and verification. Complete with servo-controlled scanning turning table and professional photographic lightning equipment. This enables us to generate 3D models that include color information (meshes with texture) in under ten minutes for typical objects with accuracy and resolution better than 0.25mm. The models created this way can be used by the robots to detect and find the position of the objects in the environment using the sensors on-board.

Other available supporting equipment and tools in EASE include prototype and robot parts manufacturing facilities, electrical diagnostic tools, and a electronics workshop.

Integrated control systems for robotic agents The success of research projects that investigates generative models of agency depend on software that can perceive their environments and produce competent actions. For robotic agents to physically perform manipulation activities as complex as required by EASE's scenario is particularly challenging. To the best of our knowledge only very few research groups world-wide are currently able to develop such systems. Below we will explain in detail how EASE provides an adequate control system to built upon from the very beginning.

The robot control system developed in EASE are unique world-wide 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

Robot control framework EASE takes the *cognition-enabled control* (Beetz *et al.*, 2012) paradigm as starting point for investigating embodied generative control models for mastering everyday activity. The paradigm was designed for and successfully applied to autonomous achievement of complex manipulation tasks by mobile manipulation robots. Examples include making pizza (rolling out dough, spreading sauce, topping with cheese), making pancakes (pouring batter, flipping pancake), (Beetz *et al.*, 2016), conducting chemical experiments (opening/closing tubes, pipetting, loading/unloading a centrifuge) (Lisca *et al.*, 2015), and sorting surgical instruments in cooperation with humans (Beetz *et al.*, 2015c).

Cognition-enabled control framework employed in EASE proved itself previously in a multitude of contexts

Cognition-enabled control is a control paradigm characterized by the embodiment of AI techniques in a physical robot system. It hinges on the combining *reactive behavior specifications* represented as *semantically interpretable plans* with *inference mechanisms* that enable flexible decision making (Figure 27). It provides the basic organizational principles for an information processing infrastructure aimed at improving the task performance in terms of robustness, flexibility, adaptivity, and efficiency. It does so through the application of cognitive mechanisms such as model acquisition, reasoning, planning, and learning from experience.

The paradigm has its roots in transformational planning of reactive behavior (McDermott, 1992b; Beetz, 2000, 2002a, 2001) and was developed into the main control framework for the autonomous robots within the German cluster of excellence CoTeSys (Beetz *et al.*, 2007). Cognition-enabled robot control addresses how to specify flexible and context-directed robot plans that at the same time allow the robot to reason about what it is doing. It does so using concurrent reactive plans, written in a high-level language, that employ inference mechanisms for selecting the appropriate actions and parametrizations. These inference methods are complemented by mechanisms that reason about and manipulate the concurrent reactive plans during their execution. Cognition-enabled control is used by and extended in several EU projects within the cognitive systems program, including RoboHow, RoboEarth, ACAT, SAPHARI, and SHERPA.

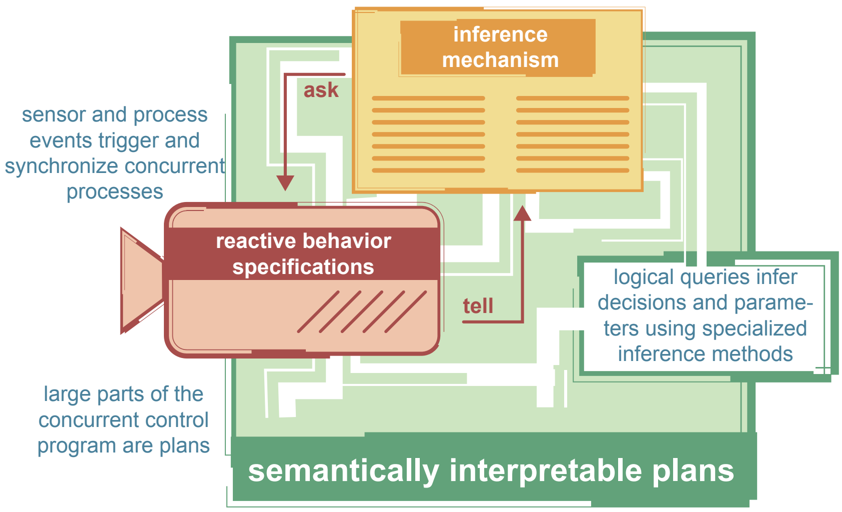


Figure 27: Control architecture realizing cognition-enabled control.

An implementation of this paradigm is found in CRAM (Cognitive Robot Abstract Machine (Beetz *et al.*, 2010a)), a largely open-source software toolbox [27]. 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).

While a variety of everyday tasks have been implemented on robots using CRAM, it is **incomplete**: the reasoning has only been applied to selected subproblems. The framework does not specify how to overcome the computational complexity of all the reasoning tasks necessary.

EASE uses CRAM to work on different aspects of interest and extends cognition-enabled control

[27] http://www.cram-project.org

How robots can acquire the body of commonsense knowledge that is needed for appropriately executing actions in a variety of contexts also remains unclear. The findings of EASE can form a natural, powerful extension of this framework. Moreover, by using this framework and the existing software of CRAM, EASE researchers can work on different aspects of interest while being able to integrate and test the results in a complete system.

> *If you give someone Fortran, he has Fortran. If you give someone Lisp, he has any language he pleases.*
> *— Guy L. Steele*

CRAM writes robot control programs in the CRAM Plan Language (CPL), implemented in Lisp. Plans are not only executable program code, but can also be modified and reasoned about by the robot itself. They describe the desired behavior in terms of a hierarchy of goals, rather than a fixed sequence of actions that need to be performed. This leads to increased flexibility, e.g. in case goals are serendipitously achieved without explicitly performing an action.

CPL provides sophisticated control structures for concurrent reactive plan execution, for example *in-parallel-do* for running multiple threads, *try-in-parallel* for trying several alternatives at once until one succeeds, or *with-constraining-plan* for running another plan in parallel, for example to supervise the execution and to interrupt the main plan if necessary. Besides these control structures, a central element of the CPL language are *designators* which are (possibly incomplete, redundant or even wrong) descriptions of objects, locations or actions. By not requiring them to be correct and consistent, the robot can reason about these descriptions and actively decide how to handle them. The current information can be combined with background knowledge (about the action or plan) in a knowledge base.

ROBOSHERLOCK for perception

Robot perception Robot perception in EASE will be realized using ROBOSHERLOCK[28]. It is an open-source project, which means that researchers are able to integrate their own robot perception capabilities into the framework and thereby extend its functionality as well as robustness. ROBOSHERLOCK is further developed and investigated in the DFG project "ROBOSHERLOCK" (2015–2018). Here we will briefly highlight the relevant features, for more details see Beetz *et al.* (2015b) (Best Service Robotics Award ICRA 2015). In ROBOSHERLOCK, perception and interpretation of realistic scenes is formulated as an Unstructured Information Management (UIM) problem. ROBOSHERLOCK combines knowledge with perception and supports knowledge-enabled reasoning about objects and scenes, and uses these results together with knowledge to generate perception pipelines automatically.

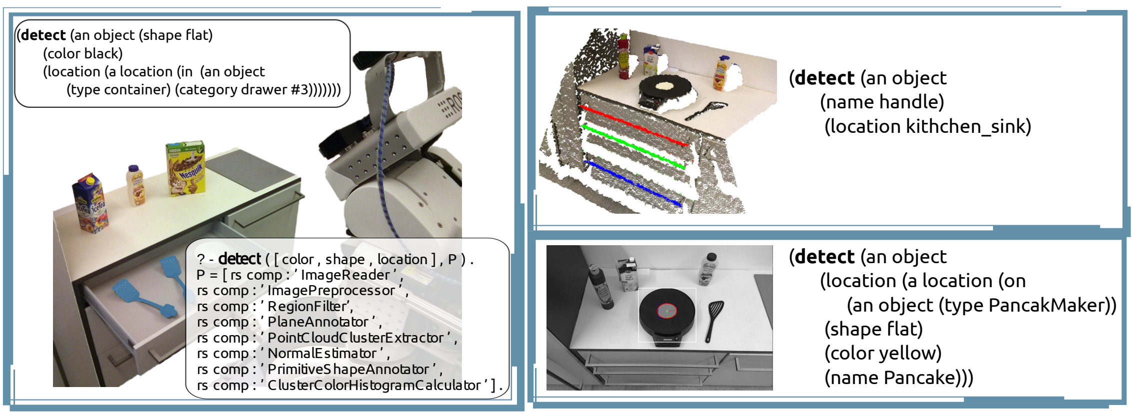


Figure 28: Queries formulated in the meta-language offered by ROBOSHERLOCK with their respective results highlighted in the processed image or point cloud. Custom perception pipeline planned for a query is shown on the left side

[28] www.robosherlock.org

RoboSherlock can complete complex perceptual tasks, use partial descriptions to find objects, and extract specific information such as pose, shape, or segmentation into functional parts (e.g., the handle) (Bálint-Benczédi *et al.*, 2016; Tenorth *et al.*, 2013a). Examples of perception results are shown in Figure 28. It boosts recognition performance and robustness by combining the strengths of multiple perception algorithms. It can perform object perception and simple recognition of human activities concurrently within one framework.

Combining expert modules

Moreover, it is compatible with the cognition-enabled control paradigm: perception tasks are represented similarly to robot plans, the perception pipeline can be reasoned about and can be adapted on the fly.

Reasoning about and automatically adapting the pipeline

RoboSherlock is designed for the "embodiment" of perception in autonomous robots and has been employed for this purpose for several major applications. It can interpret globally perceived coordinates relative to the robot (calibration) and convert perception data automatically to structures more suitable for robot manipulation. It also manages (partial) belief states about perceived objects and scenes. RoboSherlock can reason about the probabilities that two object hypotheses correspond to the same object in the real world and incrementally improve the information it has about objects. Using such a belief state, a robotic agent can form strong expectations about the world and formulate perception tasks as the validation of some part of the belief state rather than performing the perception task from scratch.

The current version of RoboSherlock employs a number of leading edge third-party robot perception systems. The perception libraries PCL (Point Cloud Library) and OpenCV are integrated into the RoboSherlock framework. In addition, the currently available methods include BLORT (Mörwald *et al.*, 2010), a toolbox for the detection, recognition, localization and tracking of objects, MOPED (Collet Romea *et al.*, 2011), a framework for multiple object pose estimation and detection, LINE-MOD (Hinterstoisser *et al.*, 2011), one of the best methods for generic rigid object recognition, and Google Goggles [29]: application for searching the web based on pictures.

All in all, RoboSherlock offers many of the services/structure required by EASE. The concepts behind RoboSherlock are also a good match for EASE. Besides its compatibility with the EASE scenario and cognition-enabled control, RoboSherlock stores incoming perception tasks along with the raw images used for accomplishing them. The architecture employs ways to store/retrieve past percepts and specialize perception based on experience, which will facilitate the collection and usage of NEEMs and PEAMs.

The specialized perception method to be investigated in Subproject R02 will be integrated into RoboSherlock and used for object detection, localization, and categorization in cluttered scenes, such as the fridge, cupboards, and drawers. EASE will also contribute to robot perception by providing new benchmark datasets. EASE can for example generate image benchmark databases from NEEMs collected during real robot tasks. A key advantage to being able to query images semantically, is that it will be straightforward to create a set of images that meet certain requirements. E.g., one could retrieve all images of table scenes with at least 5 objects captured from a distance of at least two meters, or of the opened refrigerator with at least two bottles standing close to each other.

Robot knowledge representation and processing EASE will use the KnowRob knowledge system as a key component for representing and reasoning about knowledge. KnowRob was designed with autonomous robots performing human-scale manipulation tasks in mind. The system contains **virtual knowledge bases** – collections of knowledge that are not explicitly represented but computed on demand from the robot's internal structures, its perception system, or external sources of information. KnowRob supports different types of knowledge, inference mechanisms, and interfaces for acquiring knowledge.

KnowRob for knowledge representation and reasoning

KnowRob contains encyclopedic knowledge with a conceptualization of the information

[29] https://support.google.com/websearch/answer/166331

needed for autonomous robot control. It extends ontologies commonly used in Artificial Intelligence to make the knowledge actionable. It contains a rich representation of actions, events, processes, situations, action effects and consequences, failures, knowledge preconditions of actions, etc. It also contains self-knowledge about the robot's sensors, actuators, and their respective capabilities. Given an action specification, the robot can use KNOWROB to decide whether it is capable of performing the action and whether missing capabilities can be obtained in any way.

Secondly, KNOWROB provides representation and reasoning capabilities for forward models to predict the outcome of an action as well as declarative specifications of an action's prerequisites and effects. These models support the robot with action planning, projecting future world states, and reasoning about the changes created by actions (Tenorth & Beetz, 2012).

Thirdly, KNOWROB also provides a rich representation and reasoning infrastructure for the perception, interpretation, analysis, and modeling of human activities. This infrastructure has been applied to the automated acquisition of activity models of table setting tasks (Beetz *et al.*, 2010c), to learn motion models of reaching tasks (Nyga & Beetz, 2012; Albrecht *et al.*, 2011), and for imitation learning from observation of manipulation tasks including gaze information (Ramirez-Amaro *et al.*, 2015c).

Integration of knowledge with perception and robot control

Furthermore, KNOWROB offers integration with ROBOSHERLOCK and the robot control system. It can integrate perceptual data with abstract knowledge from the knowledge base and generate answers during task execution. For example, to correctly flip a pancake using a spatula, KNOWROB retrieves the knowledge that a spatula should be grasped at the handle and together with ROBOSHERLOCK returns what that means given the currently perceived scene. It can also "listen in" on the control program and log its internal data structures as a dynamic, virtual knowledge base (Mösenlechner *et al.*, 2010). It can similarly to perceptual data, reason over the robot's internal data such as pose in combination with the knowledge to provide answers to aid task execution. Since the knowledge is generated from the data used for controlling the robot, the abstract representations are inherently grounded.

Finally, KNOWROB supports knowledge acquisition from different sources. These sources included observation of human activity (Beetz *et al.*, 2010c), logging of robot activity (Mösenlechner *et al.*, 2010), retrieving knowledge from the Web (Tenorth *et al.*, 2011), collections of common-sense knowledge like the Open Mind Indoor Common Sense database (Kunze *et al.*, 2010), and knowledge sharing techniques for robots such as the ROBOEARTH system (Waibel *et al.*, 2011). This will be useful for incorporating the knowledge generated in EASE and working with heterogeneous data.

KNOWROB is well-suited as a basis for the common representation and processing of knowledge in EASE, because it is has demonstrated to be able to support representation and inference of/on actionable robot knowledge, perceptual knowledge, and human activity. It will be used in Research Area R and serve as the representational basis for Research Area H. In particular, the NEEMs that are acquired from experiments with human everyday activity will be represented in KNOWROB and KNOWROB will through the help of OPENEASE provide powerful visualization, data mining, and learning tools to work with these experiment knowledge bases.

KNOWROB integrates different reasoning mechanisms used in EASE

Because KNOWROB is a hybrid knowledge representation and processing system where the use of knowledge bases and reasoning methods can be dynamically changed, it provides a suitable infrastructure for the integration of the different reasoning mechanisms investigated in EASE. The different inference methods researched in Areas P and R, in particular the semantics of action verbs (Research Area P02), ontological reasoning (Research Area P02), spatial reasoning (Research Area P03), probabilistic models (Research Area P01), *narrative-enabled episodic memories (NEEMs)* (R01), results of scene perception (Research Area R02), simulation-based reasoning (Research Area R03), as well as the specialized reasoning methods exploiting pragmatic manifolds (PEAMs) (Research Area R04) will be integrated into KNOWROB.

Motion control Artificial Intelligence research mostly abstracts away from how actions are to be executed, for example modeling actions through preconditions and effects only.

This successfully captures most of what we would consider the high-level, semantic information about an action. One has to presume actions can be successfully executed if their preconditions are met however. To make the models more realistic, action representations have been developed that model non-deterministic effects, probabilistic relations between precondition, context and effects, the inclusion of additional execution time information, etc. Nonetheless, the models remain relatively abstract. By restricting representation and reasoning in that manner, the symbolic reasoning systems consider the actions themselves as black boxes.

In contrast, in robot manipulation research we can easily see that small changes in execution of an action can have a large impact on the results. For example when cracking an egg for meal preparation, the effects will vary widely depending on the parametrizations of the motions, where the egg hits the surface, whether the contact point is sharp, etc.

Appropriate parametrization requires reasoning about the current scene and the prediction of the possible action effects depending on the situation and parametrizations. For this we require artificial intelligence methods at much more detailed and realistic models than is commonly the case.

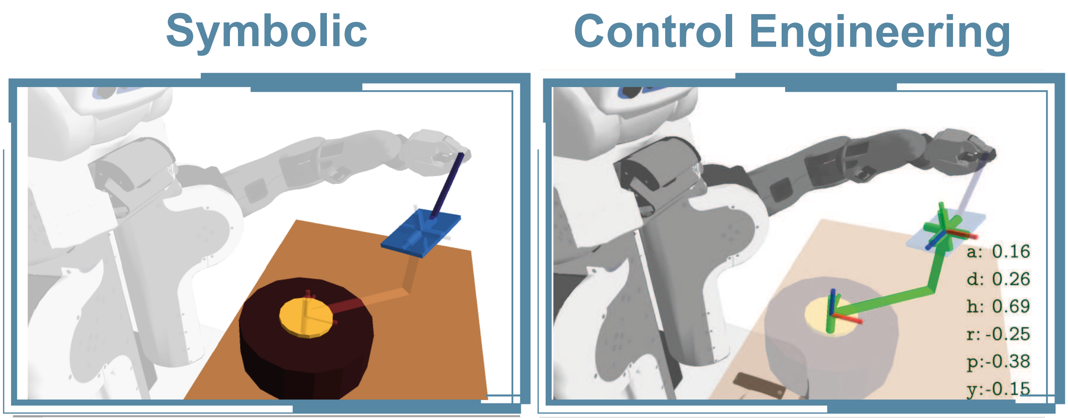


Figure 29: Two ways of representing “putting a spatula under a pancake”: the symbolic approach commonly specifies only the action in terms of objects/parts and higher-level action description, while the control engineering approach commonly specifies only coordinate frames, velocities, forces, etc.

In EASE we will use a motion control system developed in the EU project RoboHow, that bridges the gap between action representations in AI and robot control engineering. The two approaches are depicted in Figure 29. The AI-based approach on the left of the figure allows the programmer to describe the action in terms of objects, their parts, and effects, but provides no vocabulary for the detailed motions to be executed to perform the task successfully. The control engineering approach on the right provides a rich mathematical vocabulary to specify motions but only in terms of coordinate frames, without any interpretation. This control system completely lacks a semantic understanding of what it is doing.

Movement descriptions as first-class objects in the robotic agent’s knowledge representation

The action execution system to be used in EASE bridges the gap between the symbolic side and the control side. It introduces movement descriptions as first-class objects into the knowledge representation and the robot plan language. These movement descriptions are implemented as constraint-based movement specifications that are fine-grained, modular and transparent and serve as interlingua that is shared between both layers. On the symbolic layer, the constraints can be interpreted in a qualitative fashion like “on top of” or “pointing at”. On the subsymbolic layer, these constraints are expressed as relations between coordinate frames.

1.3 Positioning of EASE within its general research area

The CRC EASE will target a very important research area: the information processing technology needed to realize robot (co-)workers, assistants, and companions. These are expected to play key roles in dealing with various challenges of aging societies.

In this general research field EASE targets a knowledge-enabled approach to agent control. Knowledge intensive information systems had tremendous success in expert problem-solving domains and open question answering. In the robotics domain, however, knowledge intensive approaches to autonomous robot control have received little attention so far. An important reason for the absence of knowledge based approaches to robot control is that the problem of grounding symbolic representations in the perception-action loops of robots has turned out to be very difficult to solve. Also, to master manipulation actions robots need a lot of commonsense and naive physics knowledge. These are still not well understood in terms of computational models.

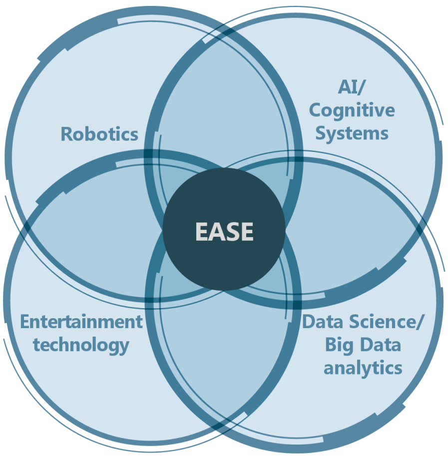


Figure 30: EASE in the intersection of disruptive technologies.

EASE proposes to modify robot control systems such that they generate NEEMs in order to collect subsymbolic robot experiences and annotate them with symbolic representations (narratives). NEEMs will provide the representational infrastructure to firmly ground the symbolic representations into the control systems. Collections of NEEMs will be suited for robotic agents to learn knowledge such as where objects are, when actions fail, etc. from experience. Therefore, we expect EASE to substantially advance our understanding of how we can build knowledge-enabled robotic agents.

EASE is uniquely positioned within the broader research field of investigating computational models of intelligent agency. EASE investigates the autonomous accomplishment of human-scale mobile manipulation tasks in realistic environments. It brings together expertise in the realization of autonomous fetch and place tasks, food preparation (pancake, popcorn, sandwich, and pizza making), part assembly, force controlled motion and manipulation capabilities, and the learning of sophisticated hand manipulation skills.

Building up competence in autonomous intelligent object manipulation is difficult and time consuming. It requires the tight integration of diverse robot capabilities, such as perception, reasoning, motion control, and manipulation, which are typically investigated in isolation. The necessity to have complete and tightly integrated robot systems capable of autonomous object manipulation is a barrier for research enterprises to quickly enter the same area as EASE. The expertise and operational systems contained in EASE give it an excellent starting point and a very promising longterm research perspective in performing leading-edge research.

Combining AI, Cognitive science, Robotics, Data science and Entertainment technology

EASE combines research and technological developments in AI and Cognition-enabled systems, Robotics, Data Science and big data analytics, and Entertainment technology. Artificial Intelligence, Cognitive Science, and Robotics form the research foundations of EASE (see Figure 30). Data Science and Entertainment technology provide tools that open up research opportunities that have not been in reach before. They will have substantial impact on making the EASE research goals achievable. Entertainment technology has recently made impressive progress in areas such as cheap and fast parallel computation devices (GPUs), realistic simulations, games for knowledge acquisition, sensing devices, virtual reality technology. Data science gives EASE the tools to interpret massive amounts of experience data and transform the data into valuable information and knowledge. Thus, EASE is a timely research enterprise that exploits today's disruptive technologies.

Artificial Intelligence With respect to Artificial Intelligence, EASE will focus its efforts on knowledge representation, processing, reasoning, learning, and planning. This includes commonsense and naive physics knowledge, as well as the embodiment of reasoning mechanisms for autonomous robot control. The use of AI representation techniques enables EASE to formally represent experience data such that it can be automatically reasoned about and can be connected to background knowledge. Thus, EASE builds on more than 60 years of research in knowledge representation, in particular on research in the areas of reasoning about action and change, naive physics, commonsense, temporal and spatial reasoning, and high volume knowledge bases. EASE also uses machine learning techniques to abstract and generalize the data contained in NEEMs into actionable knowledge. EASE will also substantially advance the state-of-the-art in AI by building one of the most comprehensive, embodied robot control systems realized using symbolic AI technology and applying it to the accomplishment of human-scale manipulation tasks.

We expect a key contribution of EASE to AI to be a novel way of equipping artificial (robotic) agents with commonsense and naive physics knowledge, by gathering and learning from very large collections of NEEMs from robots performing human-scale manipulation tasks. The hypothesis that a substantial part of commonsense and naive physics knowledge might be implicitly contained in distributions of episodic memories or NEEMs is a promising research direction for acquiring the commonsense knowledge closely connected to everyday manipulation activities.

Several subprojects, including Subproject H03, R01, and R05, will also investigate deep learning, for example in order to learn appropriate representations for collections of NEEMs. EASE is an interesting application domain for deep learning as NEEMs from related activity types might benefit greatly from learning representations that are tailored for the activities. Also, the comprehensive semantic annotations that NEEM narratives provide for the NEEM data might enable the learning of more modular representations that are easier to transfer and apply to other action categories.

Robotics EASE will also firmly build upon research in Robotics. We will focus in particular on autonomous mobile object manipulation and the integration of perception, reasoning, and goal-directed action with robot control. We will build on some of the most advanced approaches in impedance control (Borst *et al.*, 2009), which are capable of generating the sophisticated motions needed for many complex everyday manipulation tasks. A second basis will be learning-based action control that learns complex manipulation skills from experience (Elbrechter *et al.*, 2012).

Cognitive Sciences EASE research in Cognitive systems will be focused on understanding human agency and non-convential reasoning techniques, including simulation-based reasoning, narrative intelligence, episodic and semantic memories, Bayesian cognition, and Mind's eye functionality. EASE will use promising theories of cognition and agency from the Cognitive Sciences to build computational information processing models and test their embodiment in robotic agents. Research trends in the Cognitive Sciences that will be taken up in the EASE research agenda include the co-development of action and language, theories concerning episodic memories and the human memory system, the simulation theory of cognition, and Bayesian cognition. We expect that EASE can provide valuable feedback about these theories by investigating the impact of the theories on the mastery of everyday activity.

EASE distinguishes itself through key focus points

EASE aims at becoming an internationally leading and highly visible research institution in the field of AI-based and cognition-enabled robotic agents. EASE distinguishes itself from other research initiatives and large-scale projects in this field through its focus on:

- **Integrated, complete, experimental robotic systems** performing everyday manipulation activities as a driver and validator of research.

- **Embodied AI technology**, including knowledge representation, reasoning, and high-level semantics modeling objects and tasks, to increase the generality, robustness, and flexibility of robot control systems.

- Enabling robots to perform **complex human-scale manipulation activities autonomously** in an open world over extended periods of time.

EASE will realize its goals and objectives with an open research policy by publishing the code base open source, seeking intensive national and international cooperations, and organizing scientific events for community building. The importance of the field in which EASE is situated is illustrated by the attention it has received from large-scale programs of funding agencies in Europe, the United States, and Asia. At the European level, EASE would fall into the scope of the EU Challenge 2 Cognitive Systems of the Horizon 2020 framework. The U.S. has recently started the National Robotics Initiative (NRI) that also aims at investigating cognitive technology, though primarily in the context of human-robot interaction and cooperation. EASE PIs serve as NSF reviewers for this initiative and are included as international collaborators for multiple proposals submitted to it.

Robotics research, also in combination with cognitive methods, enjoys a high priority in Asia as well. Humanoid robotics and the application thereof in domestic and industrial settings are particularly far developed in countries where the challenges associated with population aging are most imminent, such as Japan and Korea. EASE intends to intensify the cooperations that already exist with three leading research laboratories at the University of Tokyo (the JSK Lab lead by Prof. Inaba, the Kuniyoshi Lab, and the Nakamura Lab). We will also cooperate with leading Korean research labs at the Seoul National University (Prof. Tak Zhang and Prof. Frank Park) and Hanyang University (Prof. Il-Hong Suh).

Position of EASE

EASE can take a leading position in the rapidly maturing research field of cognition-enabled robotics and take a distinct role in this community by focusing on human-scale activities in open worlds and by promoting embodied AI technologies. The existing cooperations that are to be intensified over the course of EASE also promise synergies at a global level.

1.3.1 Related research efforts

EASE will continue the strong cooperation with the former CoTeSys partners

In a special issue Cognition for Technical Systems of the members' magazine of the German AI Society, Beetz & Kirsch (2010) give an overview of Germany's Collaborative Research Centers and projects in the field. The editorial of the special issue also contains pointers to additional information resources. In the following, we will discuss the most related initiatives. The research enterprise most closely related to EASE is the excellence cluster CoTeSys[30] (2006-2014) in which the EASE speaker served as vice-coordinator (Buss & Beetz, 2010). Indeed, everyday manipulation and the application of AI-based control methods was a main research focus of CoTeSys. In 2011, CoTeSys refocused its research direction towards mathematical formalizations of cognitive control, team-based cognitive control, and cognitive architectures. The research area of everyday manipulation and its coordinator Prof. Beetz moved to Bremen. EASE intends to continue the strong cooperation with various members of the CoTeSys research team, including Prof. Gordon Cheng, Prof. Martin Buss, Prof. David Vernon, Dr. Michael Zehetleitner, and Prof. Erich Schneider.

[30] `http://www.cotesys.org`

CITEC (Cluster of Excellence Cognitive Interaction Technology, Bielefeld University) (Maycock *et al.*, 2010) is a Cluster of Excellence at the Bielefeld University that also investigates cognitive technology. CITEC pursues a three-fold mission: Creating cognitive abilities in technical systems from everyday devices to humanoid robots to make them more useful and pleasant to interact with for lay people; advancing our scientific understanding of the principles and mechanisms that enable seamless cognitive interaction; and ultimately, creating bridges between the cultures of engineering and humanities to better shape tomorrow's technology according to human needs. CITEC includes a very broad scope of applications with regard to interaction between humans and machines. In comparison EASE is focused deeper methodologically and in terms of application, considering a broader range of aspects within a well-defined area of applications and methods. I.e. EASE is focused on knowledge-enabled control and everyday manipulation activity, considering the full spectrum of how knowledge and skills are acquired, represented, reasoned about, and applied for autonomous execution without delay given the everyday activity domain. Thus, EASE and CITEC would be two research initiatives that complement each other well. Especially in the later stages where we expect to go more toward multi-agent scenario's, EASE plans to build upon the results from CITEC. Prof. Helge Ritter, affiliated with CITEC, is also a principal investigator in EASE.

EASE complements CITEC very well

Another related research endeavor was the **CRC Humanoid Robots**[31] at the University of Karlsruhe (2001-2012). The goal of this project was to generate concepts, methods and concrete mechatronical components for a humanoid robot that is to share his activity space with a human partner. With the aid of this partially anthropomorphic robot system, it will be possible to step out of the *robot cage* and realize direct contact with humans. The research conducted in the CRC Humanoid Robots differs from EASE both in terms of scope as well as the methodological approach. The CRC Humanoid Robots included the development of humanoid robot platforms and their low-level control systems. Methodologically, Karlsruhe's research methodologies were more directed towards developmental robotics techniques and imitation learning. The complementary research of Prof. Rüdiger Dillmann's and Prof. Tamim Asfour's research groups in the context of European research projects, most notably PACO+ and XPerience suggest interesting synergies for EASE.

MIT has established the **Center for Brains, Minds, and Machines (CBMM)** as an NSF Science and Technology Center dedicated to the study of intelligence - how the brain produces intelligent behavior and how we may be able to replicate intelligence in machines. This center is in spirit very close to EASE in that both efforts focus on the combination of science and engineering of Intelligence and consider engineered computational methods as scientific means to investigate the *science* of intelligence. The CBMM is phenomenologically and methodologically broader than EASE whereas EASE focuses on everyday manipulation activities and knowledge-enabled agent control. A distinct strength of EASE is its strong commitment to the realization of competent robotic agents capable of mastering everyday activities.

On a national level, the DFG has established the **German priority research program (SPP) on Autonomous Learning** [32], in which members of the EASE consortium participate with the project Autonomous Learning for Bayesian Cognitive Robotics. Within the SPP, groups cooperate in the area of robot learning. In this context, extended visits of doctoral students have been performed and are envisaged with the research groups of Prof. Marc Toussaint (University Stuttgart) and Prof. Kristian Kersting (University of Bonn) to investigate aspects of first-order probabilistic learning and inference for robot control. In contrast to the project proposed within SPP, which investigates how robot control systems should be designed and implemented to provide the learning data for Bayesian Cognitive Robotics, EASE focuses on the issues concerning modeling, learning, and inference problems in Bayesian Cognitive Robotics.

[31] `http://www.sfb588.uni-karlsruhe.de`

[32] `http://autonomous-learning.org/`

The **SFB TR 62 A Companion-Technology for Cognitive Technical Systems** (Biundo & Wendemuth, 2010) investigates the vision of companion systems — cognitive technical systems that provide their functionality in a completely individualized way. The goal is to realize technical systems that adapt to an user's capabilities, preferences, requirements, and current needs and take into account both the situation and the emotional state of the individual user. Furthermore, they are continually available, co-operative, and reliable, and appear as competent and empathetic assistants to their users. EASE ś focus on embodied AI and robotic systems for human-scale manipulation make it quite different from the SFB TR 62. It complements the initiative well however, given that the competency in the performance of everyday activity would provide additional powerful mechanisms to adapt to a user's capabilities, preferences, and requirements.

A number of related collaborative projects can be found at the European level, in particular in the context of the **FP7 challenge for Cognitive Systems**. The EU projects that are closest to the EASE proposal with respect to their objectives are **First-MM** (Flexible Skill Acquisition and Intuitive Robot Tasking for Mobile Manipulation in the Real World, 2010–2013), **GeRT** (Generalizing Robot Manipulation Tasks, 2010–2013), and **IntellAct** (Intelligent observation and execution of Actions and manipulations, 2011–2014). First-MM aims at developing a novel robot programming environment that allows even non-expert users to specify complex manipulation tasks in real-world environments. The programming environment is to include a task specification language and probabilistic learning and inference mechanisms for learning manipulation skills from demonstration and from experience. GeRT starts with a small set of existing robot programs, for a certain robot manipulation task, and aims at giving the robot the ability to adapt them on the fly to novel objects and task variants. IntellAct investigates the programming of industrially relevant tasks, either in a factory or on a space station, using imitation learning where the system also extracts the semantic meaning of the observed actions through semantic event chains.

The EU project **RoboEarth** (Waibel *et al.*, 2011; Tenorth *et al.*, 2011, 2012) dealt with knowledge intensive robot applications. In RoboEarth, robots share a semantic knowledge base for action recipes and domain knowledge. The same holds for the **Rosetta** project (2009–2013), which included knowledge and skill representation and knowledge transformation and learning as work packages with the objective of reducing deployment effort in order to allow fast production changeover from product A to product B.

We further draw upon the results other EU projects, such as **PACO+** that investigated Object-Action Complexes (OACs) as a universal representation enabling efficient planning and execution. Another example is **RobotCub** which researched embedded cognition in the context of a humanoid robot. **iTALK** explores the co-development of language and action, which is relevant since natural language is particularly rich and can express the nuances of everyday manipulation. **POETICON++** investigated the use of natural language as a learning tool for the generalization of learned behaviors and generation of new behaviors and experiences.

More recently, several European projects have focused on improving the robustness and performance of robotic agents from experience. The project **RACE** has aimed at robots capable of storing experiences in their memory in terms of multi-level representations connecting actuator and sensory experiences with meaningful high-level structures, methods for learning and generalizing from experiences obtained from behavior in realistically scaled real-world environments, and robots demonstrating superior robustness and effectiveness in new situations and unknown environments using experience-based planning and behavior adaptation. While the objectives are similar to those of EASE the level of ambition of EASE is much higher. In EASE we intend to learn entire naive physics and common-sense knowledge bases from the robot experiences, cover much more complex manipulation tasks, transfer knowledge and insights from models of human activity, and study their relationship to everyday activity. The EU FET H2020 project **DREAM**[33] investigates sleep and dream-like processes within cognitive architectures. In con-

[33] `http://www.robotsthatdream.eu`

trast to EASE, DREAM is located in the developmental robotics research line and does not have the strong influence of symbolic representation and reasoning and is consequently not able to aim for knowledge intensive activities. Finally, the EU H2020 **RobDream**[34] project investigates how robots can optimize their performance in task execution by optimizing the parameters of their control programs in idle times and based on collected execution data. Their focus lies in the manufacturing domain. Moreover, the intended optimization does not include the acquisition, improvement, and use of comprehensive knowledge bases.

Another related research effort is the **EU flagship Human Brain Project**[35], a 10-year 1 billion € project that is to be carried out by the broad European research community. The Human Brain Project aims at developing a large-scale ICT infrastructure for understanding the brain and its diseases, and of translating this knowledge into new computing technology. One of the aims is to build a virtual environment in which modeled robots can act, based on existing open-source gaming platforms. The purpose of this tool will be to test the behavior of brain models however, whereas a component of EASE is to create virtual reality and simulation tools from which artificial agents learn how humans perform activities and build appropriate models based on this data.

EASE distinctionEASE differs from the above projects in that it investigates *human-scale, everyday* manipulation tasks in humans and autonomous robots. The tasks involved in these everyday activities are highly demanding with respect to *how* the activities should be executed in the particular situation and object context. As a consequence, the robot control systems to be investigated in EASE must be very knowledge-intensive – an approach that is not pursued in the other projects.

1.3.2 Progress beyond the state-of-the-art

We project that the main scientific impact of EASE will be along three main dimensions. First, EASE will build a **comprehensive knowledge base about people performing everyday activity**, which will represent aspects of problem-solving behavior that are difficult to capture because people are typically not aware of them. These aspects are mined using computer games and virtual environments designed to capture this implicit knowledge. We plan to make this knowledge base publicly available. Using these data, EASE researchers will investigate how everyday activity can be better structured and will identify PEAMs in the underlying problem space. Second, EASE will substantially advance **symbolic reasoning techniques for everyday activity problem solving** and provide efficient reasoning mechanisms that are capable of exploiting the structure of everyday activity. In addition, the reasoning mechanisms developed and investigated will apply to complex real-world manipulation tasks. Third, EASE will design, develop and investigate a **new generation of control systems for robotic agents**, such as robot (co-)workers, assistants, and companions, that can achieve much higher flexibility, generality, robustness, and performance than existing robot control systems (see Figure 30).

[34] `http://robdream.eu/`
[35] `http://www.humanbrainproject.eu/`

Why a Collaborative Research Center on "Everyday Activity Science and Engineering"?

Information processing models for mastering everyday activities

1. **EASE aims at scientific challenges with very large potential impact.** EASE is to investigate and develop information processing models for mastering everyday activity, based on comprehensive bodies of commonsense and naive physics knowledge, that can perform competent everyday decision making in a computationally efficient manner. These models will constitute an important step towards the realization of robotic (co-)workers, assistants, and companions: essential components of technology roadmaps for dealing with the challenges of aging societies. Having such information processing models will also have a substantial impact on better understanding, including measuring and assessing, the cognitive abilities of humans in performing their everyday activities.

EASE partners are in an excellent position to achieve the proposed research goals

2. **EASE is very challenging and yet technically feasible.** The research goal is very challenging and has so far received less attention than warranted. Yet, the proposing team of principal investigators and the university are in an excellent position to make the research enterprise feasible. The information processing methods developed in the context of the CRC Transregio Spatial Cognition, such as belief-based architectures for scene understanding by Schill and corpus-based ontology development by Bateman, will provide a good basis for researching processing models in EASE. Expertise in making operational robotic agents capable of performing complex manipulation tasks is provided by Beetz, among others, from the research performed in the cluster of excellence CoTeSys and several EU FP7 projects including RoboHow, RoboEarth, ACat, and SaPHaRI.

Threefold approach for achieving the research goals

3. **EASE has a well-defined goal, research agenda, and measures of success.** The goal of proposing information processing models with more competence in mastering everyday activity is tackled by (1) investigating models of human activity, (2) studying the formal foundations of the information processing principles, and (3) embodying the models into robotic agents. The measure of success will be how well a robotic agent can perform a morning's housework every day over an extended period of time.

EASE requires and supports a critical mass of interdisciplinary researchers in a long-term collaboration

4. **The EASE research program requires the framework of a CRC.** Realizing the research program of EASE requires us to bring together a critical mass of researchers into a common long-term research enterprise to investigate the mastering of everyday activities from different research perspectives including Artificial Intelligence, Robotics, Cognitive Neuroscience / Neuroinformatics and Linguistics. This is necessary to create the synergies and momentum required to succeed in such ambitious goals.

EASE benefits from its hosting institutions' expertise and vice versa

5. **EASE perfectly fits into the research profile of the University of Bremen (UB).** EASE is planned to constitute the core of UB's high-profile research area *Minds, Media, and Machines*, which receives additional support through the UB Excellence Program. Our Computer Science faculty can comprehensively cover the area with its professors and senior research scientists. The focus on cognitive science, computer science, and robotics is also mirrored by the large number of undergraduate and graduate students educated in this field at the university. The team of principal investigators is complemented with internationally leading experts in cognition-enabled autonomous robot control (Albu-Schäffer, Cheng, Ritter), which substantially increases the expertise available to EASE as well as introduces synergies with other leading research institutions such as the Cognitive Interaction Technology Excellence Cluster (CITEC) and the German Aerospace Center (DLR) for Mechatronics and Control. Pulling together the expertise and recent advances, we are in the right time and place to make the goals of EASE a reality.

1.4 National and international cooperation and networking

To further accelerate progress, we will establish and maintain a very strong international cooperation network. Autonomous robotics has recently changed from a research field where research groups develop comprehensive and complex robot control systems in isolation into one where more and more international cooperation and open-source code exchange is a strong motor for fast progress towards the tremendous research challenges (Quigley *et al.*, 2009; Metta *et al.*, 2006; Smits *et al.*, 2008). A prominent example of such an international cooperation network is the PR2 beta program[36].

Open research – Laboratories without Walls

Open research is an important part of EASE. We take *Laboratory without Walls* as an approach to support this key aspect. It is realized through intensive networking with international research laboratories through (1) visiting doctoral students and joint supervision of doctoral students, (2) the provision of a unique integrated system setup and leading-edge laboratory with operational cognition-enabled, open-source autonomous mobile manipulation robots, (3) international cooperative research through open-source and standardized software, and (4) community building through international seminars, summer schools and conference workshops and tutorials.

EASE's Laboratories without Walls will be centered around the EASE Central Research Laboratory in Bremen. Here, EASE will bring together the main scenario environment, starting with two leading-edge mobile manipulation platforms and means for observing humans performing everyday activities. This will allow all subprojects to advance their objectives from the start rather than waiting for components that may be outside their control.

The laboratory will be an interdisciplinary environment where collaborators can integrate their research with EASE and work on EASE-related topics with additional tools. The laboratory will be able to provide space for up to 20 researchers. Current international cooperations include KTH Stockholm, LAAS-CNRS, the Center for Robotics and Intelligent Machines at Georgia Tech University, JSK Lab Tokyo, CMU Robotics Institute, and others. The cooperations are supported by corresponding letters of support and intent.

1.4.1 National cooperations

The EASE principal investigators collaborate with many colleagues in Germany. The affiliation of EASE PIs with other distinguished institutions and programs opens up the road for forming new cooperations with other PIs at those institutions as well.

An important partner for EASE is the DLR Robotics and Mechatronics institute in Oberpfaffenhofen. The Institute for Artificial Intelligence at UB, which is led by the EASE speaker, has started a strategic cooperation with the DLR Robotics and Mechatronics institute in the form of a virtual research group under the label Perceptive Autonomous Agents Laboratory (PAAL)[37]. The PAAL laboratory includes Zoltan-Csaba Marton (DLR, perception) and Michael Suppa (Roboception, perception) as senior researchers and several doctoral students who will all have their academic home at UB. The joint research areas to be covered by the PAAL laboratory are robot perception as well as the bridging between symbolic action control and control engineering. On the Bremen side the cooperation also includes Schill, Frese, and Zachmann as senior researchers.

[36] `http://spectrum.ieee.org/automaton/robotics/robotics-software/050410-willow-garage-giving-away-11-pr2-robots-worth-over-4-million`

[37] `http://paal.ai.uni-bremen.de`

1.4.2 International cooperations

EASE will promote itself and support world-wide collaboration through open-source release of software and a transfer project

EASE considers it important that key scientific results of EASE will be made available as professionally implemented and maintained open-source software. The reasons behind EASE's investment in open-source dissemination of research software are manifold. We expect that the simple and off-the-shelf use of EASE's research software infrastructure and results will be an effective promoter of EASE research. The availability of open-source software will support global cooperation on the EASE research topics. The transfer of successful research software into the open-source community will also ensure that the valuable contributions will be preserved and built upon beyond the tenure of the doctoral students in the research groups. After the successful start of EASE, we plan to install a transfer project OPENEASE in close cooperation with the open-source foundations Open Perception and Open Source Robotics Foundation.

Figure 31: EASE PI Beetz was among the teams to win a PR2 robot.

The Intelligent Autonomous Systems (IAS) group at the Institute for Artificial Intelligence at UB (formerly IAS at Technische Universität München) is a member of the PR2 beta program in which internationally leading robotics research laboratories work together as partners in an open-source research network. The cooperation partners are Georgia Institute of Technology, Massachusetts Institute of Technology, University of Pennsylvania, University of California at Berkeley, Stanford University, University of Southern California, University of Tokyo, University of Freiburg and Katholieke Universiteit Leuven (together in (Figure 31).

Bilateral research cooperations include exchanges of researchers with LAAS/CNRS in Toulouse (Dr. Rachid Alami, France)[38], Seoul National University (Prof. Tak Zhang, South Korea)[39], Carnegie Mellon University (Robotics Institute & Quality of Life Technology Center, Prof. Matt Mason, Prof. Sidd Srinivasa)[40], University of Tokyo (Prof. Masayuki Inaba, Prof. Yasuo Kuniyoshi, Prof. Yoshi Nakamura), La Sapienza University of Rome (Prof. Daniele Nardi, Prof. Fiora Pirri, Assoc. Prof. Barbara Caputo) , and Edinburgh Center for Robotics (Prof. David Lane).

Additional cooperations exist with the Robots and Intelligent Machines Center at Georgia Tech (Prof. Henrik Christensen), Center for Autonomous Systems (Prof. Danica Kragic), and the German Cluster of Excellence CoTeSys (Prof. Martin Buss).

A selected overview of collaborators, along with the nature of the collaboration is given in Table 3.

[38] A start-up cooperation project "Human-enabling Robot Assistance" together with Rachid Alami.

[39] Project "Machine Learning for the Generation of Flexible Motion Trajectories for Robot Manipulators" in the German-Korean Partnership Program (GEnKO) together with Prof. Tak Zhang and in cooperation with Prof. Frank Park.

[40] Start-up project is supported by a joint DFG-NSF grant entitled "Building Intelligent Mobile Manipulators for Assistive Care" with the PIs Sidd Srinivasa, Matt Mason (Director of the Robotics Institute), and Michael Beetz.

Selected International Cooperations of EASE	
Partner	**Forms of Cooperation**
KTH Stockholm Kragic, Jensfelt Center of Autonomous Systems	◦ cooperation in EU cognitive systems projects ◦ exchange of researchers (post docs & doctoral students) ◦ perception-guided manipulation
LAAS-CNRS Alami	◦ cooperation in EU cognitive systems projects ◦ exchange of researchers (post docs & doctoral students) ◦ joint publications ◦ planned: joint doctoral degrees
Georgia Institute of Technology Robotics and Intelligent Machines Christensen, Kemp	◦ cooperation in healthcare robotics ◦ exchange of researchers ◦ perception-guided manipulation
University of Tokyo Inaba, Okada, Kuniyoshi, Nakamura	◦ cooperation with three laboratories ◦ exchange of researchers (post docs & doctoral students) ◦ cooperation in Lisp-based robot plan languages ◦ joint publications
CMU Robotics Institute Srinivasa, Mason	◦ exchange of researchers (post docs & doctoral students) ◦ proposals for cooperation projects ◦ open-source cooperation ◦ planned: Dagstuhl seminar
Seoul National University Zhang	◦ cooperation with Global Frontier Research Program: Human-Level Machine Learning ◦ exchange of researchers (post docs & doctoral students) ◦ planned: joint conference workshops
La Sapienza University of Rome Nardi, Pirri, Caputo	◦ knowledge representation and reasoning ◦ vision ◦ planned: Erasmus exchange program
Italian Institute of Technology Sandini	◦ episodic memories ◦ language and actions ◦ developmental cognitive robotics
Edinburgh Center for Robotics Lane	◦ knowledge representation and reasoning ◦ machine learning ◦ planning and decision making ◦ planned: research visits and stays

Table 3: Selected International Cooperations of EASE

1.4.3 Exchange program for visiting researchers

EASE will establish an exchange program for visiting researchers as well as for sending PIs as visiting fellows to cooperating international institutions. Visiting researchers will be integrated into EASE research groups and can apply for a stay at the Hanse Wissenschaftskolleg Institute of Advanced Studies in Delmenhorst[41]. Candidates for Mercator fellows include Sidd Srinivasa (Robotics Institute, Carnegie Mellon University) and Charles Kemp (Robotics and Intelligent Machines, Georgia Institute of Technology). EASE aims at sending its principal investigators as visiting researchers to cooperating research groups at e.g., Carnegie Mellon University, Georgia Institute of Technology, Massachusetts Institute of Technology, Seoul National University, and University of Tokyo.

1.4.4 Organization of scientific events

EASE considers the establishment and promotion of an international research community in its field of investigation key to achieving research impact. The organization of scientific events is an important means for building such communities. EASE's efforts will include community building, such as the organization of scientific workshops at conferences, summer schools, Dagstuhl seminars, etc.

EASE partners have demonstrable experience and success in organizing scientific events

EASE PI's have demonstrable skill and experience in organizing successful scientific events in their fields of expertise. These efforts will be continued and applied to EASE-related scientific events. The Cognitive Systems group, for example, has established the conference on Spatial Cognition[42] as a means for strengthening the research topic of spatial cognition at an international level, where several other PIs of EASE including Kerstin Schill, Udo Frese or John Bateman, have key organizational roles in the conference series. In the area of human computation, Malaka has organized several national and international conferences including Smart Graphics 2011, the International Conference on Entertainment Computing 2012, and Mensch und Computer 2013. Beetz organized Dagstuhl seminars on Plan-based Control of Robotic Agents[43] and Cognition-enabled Manipulation [44] as a means of strengthening the field of AI-based robotics.

Another important building block are summer schools that help to promote research topics at the level of early stage researchers. EASE principal investigators have organized international research schools, such as the Player Summer School on Cognitive Robotics 2007 [45] and the CoTeSys-ROS Fall School on Cognition-enabled Mobile Manipulation[46]. Malaka is part of the executive committee of the Interdisciplinary College, an annual, intense one-week spring school which offers a dense state-of-the-art course program in Neurobiology, Neural Computation, Cognitive Science/Psychology, Artificial Intelligence, Robotics and Philosophy. Barkovsky, member of the Cognitive Systems group, has co-organized the International Spatial Cognition Summer Institute 2013 at the University of California, Santa Barbara (UCSB). EASE plans to continue organizing scientific meetings in the research related to Everyday Activity Science and Engineering.

[41] http://www.h-w-k.de/
[42] sc2012.informatik.uni-freiburg.de
[43] http://www.dagstuhl.de/01431, http://www.dagstuhl.de/03261)
[44] (http://www.dagstuhl.de/en/program/calendar/semhp/?semnr=09341)
[45] psscr07.cs.tum.edu
[46] http://ias.cs.tum.edu/events/cotesys-ros-school

1.4.5 Cooperation with external projects

The principal investigators of EASE will propose research projects whose aim also contributes to EASE's goals, to other funding agencies and funding lines. These projects are intended to strengthen EASE by investigating complementary and synergistic research questions, and thereby allow for a better coverage of the overall field. In many of these projects, EASE PIs and other project partners work together on a shared open-source basis, which facilitates the easy integration of project results. Let us consider the field of robot perception as an example. EASE does not include subprojects for the acquisition of semantic environment models, perceptual tracking, observing manipulated objects, or for complex scene understanding. Because these perceptual capabilities play important roles in everyday activity, EASE plans to cover these topics through cooperative projects.

EU project ROBOHOW covered complementary research goals to those of EASE and can be built upon further

Clearly, for EASE, cooperative projects in the Challenge 2 Cognitive Systems in the Horizon 2020 funding framework of the European Commission are particularly important. The EU Integrated Project ROBOHOW, is an excellent example for a cooperative project. ROBOHOW, in collaboration with internationally leading research groups, investigates how robot activity plans for novel human-scale manipulation tasks can be automatically constructed by robots from recipes and demonstration videos. ROBOHOW also investigates the translation of symbolic action descriptions into leading-edge control engineering frameworks.

It is easy to see that ROBOHOW covers complementary research goals to those of EASE and provides it with ample possibilities for cooperation with internationally leading research groups.

Other Challenge 2 projects that EASE PIs coordinate and participate in include: SAPHARI, which investigates safe physical manipulation in the presence of, and together with, humans. ACAT, which studies the learning of knowledge bases for the execution of action verbs. VICON, which provides support through the development of an advanced Virtual User Model which enables virtual testing and feedback throughout the development lifecycle. ROBLOG, which investigates the development of a cognitive robot for unloading of containers in logistics. Also nationally funded projects, in particular those in national priority programs, such as Autonomous Learning for Bayesian Cognitive Robotics in the DFG Autonomous Learning priority program, will be key components of EASE. EASE managers will continually monitor opportunities of co-funding together with UB's offices for national and EU funding, and the office for central research development.

References

Agre, P. E. 1988. *The dynamic structure of everyday life*. Tech. rept. Massachusetts Institute of Technology, Artificial Intelligence Lab.

Agre, P. E. 1995. Computational research on interaction and agency. *Artificial Intelligence*, **72**(1-2), 1–52.

Agre, P. E., & Horswill, I. 1997. Lifeworld Analysis. *Journal of Artificial Intelligence Research*, **6**, 111–145.

Ahn, L. v. 2006. Games with a Purpose. *IEEE Computer*, **39**(6), 92–94.

Albrecht, S., Ramirez-Amaro, K., Ruiz-Ugalde, F., Weikersdorfer, D., Leibold, M., Ulbrich, M., & Beetz, M. 2011 (October, 26–28). Imitating human reaching motions using physically inspired optimization principles. *In: 11th IEEE-RAS International Conference on Humanoid Robots*.

Albu-Schäffer, A., Ott, C., & Hirzinger, G. 2007. A Unified Passivity-based Control Framework for Position, Torque and Impedance Control of Flexible Joint Robots. *The International Journal of Robotics Research*, **27**(1), 23–39.

Allen, J. 1983. Maintaining knowledge about temporal intervals. *Communications of the ACM*, **26**(11), 832–843.

Anand, A., Koppula, H. S., Joachims, T., & Saxena, A. 2013. Contextually guided semantic labeling and search for three-dimensional point clouds. *I. J. Robotic Res.*, **32**(1), 19–34.

Anderson, J. E. 1995. *Constraint-Directed Improvisation for Everyday Activities*. Ph.D. thesis, University of Manitoba.

Anderson, T. S. 2015a. From Episodic Memory to Narrative in a Cognitive Architecture. *Pages 2–11 of:* Finlayson, M. A., Miller, B., Lieto, A., & Ronfard, R. (eds), *6th Workshop on Computational Models of Narrative (CMN 2015)*. OpenAccess Series in Informatics (OASIcs), vol. 45. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

Arbib, M. A. 2006. *Action to Language via the Mirror Neuron System*. Cambridge University Press.

Arechavaleta, G., Laumond, J.-P., Hicheur, H., & Berthoz, A. 2006a. The nonholonomic nature of human locomotion: a modeling study. *In: IEEE International Conference on Biomedical Robotics and Biomechatronics*.

Arechavaleta, G., Laumond, J.-P., Hicheur, H., & Berthoz., A. 2006b. Optimizing principles underlying the shape of trajectories in goal oriented locomotion for humans. *In: Proceedings of the International Conference on Humanoid Robots*.

Asfour, T., Regenstein, K., Azad, P., Schroder, J., Bierbaum, A., Vahrenkamp, N., & Dillmann, R. 2006a. ARMAR-III: An integrated humanoid platform for sensory-motor control. *Pages 169–175 of: Proceedings of the IEEERAS/RSJ International Conference on Humanoid Robots (Humanoids06)*.

Asimov, I. 1991. *I, Robot*. A Bantam book. Random House Publishing Group.

Baader, F., Brand, S., & Lutz, C. 2005. Pushing the EL envelope. *Pages 364–369 of: In Proc. of IJCAI 2005*. Morgan-Kaufmann Publishers.

Bailey, D. R. 1997. *When Push comes to Shove: A Computational Model of the Role of Motor Control in the Acquisition of Action Verbs*. Ph.D. thesis, University of California, Berkeley.

Balakirsky, S., Carpin, S., Kleiner, A., Lewis, M., Visser, A., Wang, J., & Ziparo, V. A. 2007. Towards heterogeneous robot teams for disaster mitigation: Results and performance metrics from RoboCup rescue. *Journal of Field Robotics*, **24**(11-12), 943–967.

Bálint-Benczédi, F., Mania, P., & Beetz, M. 2016. Scaling Perception Towards Autonomous Object Manipulation — In Knowledge Lies the Power. *In: International Conference on Robotics and Automation (ICRA)*.

Bartels, G., Kresse, I., & Beetz, M. 2013 (October 15–17). Constraint-based Movement Representation grounded in Geometric Features. *In: Proceedings of the IEEE-RAS International Conference on Humanoid Robots*.

Beetz, M. 1996. *Anticipating and Forestalling Execution Failures in Structured Reactive Plans*. Technical Report, YALE/DCS/RR1097, Yale University.

Beetz, M., Beßler, D., Winkler, J., Worch, J.-H., Bálint-Benczédi, F., Bartels, G., Billard, A., Bozcuoglu, A. K., Fang, Z., Figueroa, N., Haidu, A., Langer, H., Maldonado, A., Pais, A., Tenorth, M., & Wiedemeyer, T. 2016. Open Robotics Research Using Web-based Knowledge Services. *In: International Conference on Robotics and Automation (ICRA)*. Accepted for publication.

Beetz, M. 2000. *Concurrent Reactive Plans: Anticipating and Forestalling Execution Failures*. Lecture Notes in Artificial Intelligence, vol. LNAI 1772. Springer Publishers.

Beetz, M. 2001. Structured Reactive Controllers. *Journal of Autonomous Agents and Multi-Agent Systems. Special Issue: Best Papers of the International Conference on Autonomous Agents '99*, **4**(March/June), 25–55.

Beetz, M. 2002a. *Plan-based Control of Robotic Agents*. Lecture Notes in Artificial Intelligence, vol. LNAI 2554. Springer Publishers.

Beetz, M. 2002b. Plan Representation for Robotic Agents. *Pages 223–232 of: Proceedings of the Sixth International Conference on AI Planning and Scheduling*. Menlo Park, CA: AAAI Press.

Beetz, M., & Kirsch, A. 2010. Special Issue on Cognition for Technical Systems. *Künstliche Intelligenz*, **24**.

Beetz, M., Buss, M., & Wollherr, D. 2007. Cognitive Technical Systems — What Is the Role of Artificial Intelligence? *Pages 19–42 of:* Hertzberg, J., Beetz, M., & Englert, R. (eds), *Proceedings of the 30th German Conference on Artificial Intelligence (KI-2007)*. Invited paper.

Beetz, M., Mösenlechner, L., & Tenorth, M. 2010a (October 18-22). CRAM – A Cognitive Robot Abstract Machine for Everyday Manipulation in Human Environments. *Pages 1012–1017 of: IEEE/RSJ International Conference on Intelligent Robots and Systems*.

Beetz, M., Tenorth, M., Jain, D., & Bandouch, J. 2010c. Towards Automated Models of Activities of Daily Life. *Technology and Disability*, **22**(1-2), 27–40.

Beetz, M., Klank, U., Kresse, I., Maldonado, A., Mösenlechner, L., Pangercic, D., Rühr, T., & Tenorth, M. 2011 (October, 26–28). Robotic Roommates Making Pancakes. *In: 11th IEEE-RAS International Conference on Humanoid Robots*.

Beetz, M., Jain, D., Mösenlechner, L., Tenorth, M., Kunze, L., Blodow, N., & Pangercic, D. 2012. Cognition-Enabled Autonomous Robot Control for the Realization of Home Chore Task Intelligence. *Proceedings of the IEEE*, **100**(8), 2454–2471.

Beetz, M., Tenorth, M., & Winkler, J. 2015a. Open-EASE – A Knowledge Processing Service for Robots and Robotics/AI Researchers. *In: IEEE International Conference on Robotics and Automation (ICRA)*. Finalist for the Best Conference Paper Award and Best Cognitive Robotics Paper Award.

Beetz, M., Bálint-Benczédi, F., Blodow, N., Nyga, D., Wiedemeyer, T., & Marton, Z.-C. 2015b. RoboSherlock: Unstructured Information Processing for Robot Perception. *In: IEEE International Conference on Robotics and Automation (ICRA)*. Best Service Robotics Paper Award, Finalist for the Best Conference Paper Award.

Beetz, M., Bartels, G., Albu-Schäffer, A., Bálint-Benczédi, F., Belder, R., Beßler, D., Haddadin, S., Maldonado, A., Mansfeld, N., Wiedemeyer, T., Weitschat, R., & Worch, J.-H. 2015c. Robotic Agents Capable of Natural and Safe Physical Interaction with Human Coworkers. *In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

Bengio, Y., Courville, A., & Vincent, P. 2013. Representation Learning: A Review and New Perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.*, **35**(8), 1798–1828.

Bertero, M., Poggio, T. A., & Torre, V. 1987. *Ill-Posed Problems in Early Vision*. Tech. rept. AIM-924. Massachusetts Institute of Technology.

Bicchi, A., & Kumar, V. 2000. Robotic grasping and contact: a review. *In: International Conference on Robotics and Automation (ICRA)*.

Birbach, O., & Frese, U. 2013. A Precise Tracking Algorithm Based on Raw Detector Responses and a Physical Motion Model. *Pages 4746–4751 of: Proceedings of the IEEE International Conference on Robotics and Automation. IEEE International Conference on Robo-*

tics and Automation (ICRA-13), May 6-10, Karlsruhe, Germany. IEEE.

Birbach, O., Frese, U., & Bäuml, B. 2015. Rapid calibration of a multi-sensorial humanoid's upper body: An automatic and self-contained approach. *The International Journal of Robotics Research*, **34**(4-5), 420–436.

Biundo, S., & Wendemuth, A. 2010. Von kognitiven technischen Systemen zu Companion-Systemen. *KI - Künstliche Intelligenz*, 335–339.

Blair, D., & Meyer, T. 1997. *Tools for an Interactive Virtual Cinema*.

Blum, M., & Kannan, S. 1989. Designing Programs That Check Their Work. *Pages 86–97 of: STOC*.

Bollini, M., Tellex, S., Thompson, T., Roy, N., & Rus, D. 2012. Interpreting and Executing Recipes with a Cooking Robot. *In: International Symposium on Experimental Robotics (ISER), Quebec City, Canada*.

Borst, C., Wimböck, T., Schmidt, F., Fuchs, M., Brunner, B., Zacharias, F., Giordano, P. R., Konietschke, R., Sepp, W., Fuchs, S., *et al.*. 2009. Rollin'Justin-Mobile platform with variable base. *Pages 1597–1598 of: Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*.

Brachman, R. 2002. Systems That Know What They're Doing. *IEEE Intelligent Systems*, November/December, 67 – 71.

Bradshaw, J., Feltovich, P., Jung, H., Kulkarni, S., Taysom, W., & Uszok, A. 2003. Dimensions of Adjustable Autonomy and Mixed-Initiative Interaction. *Pages 17–39 of: Agents and Computational Autonomy*.

Brooks, R. 1999. *Cambrian Intelligence: The Early History of the New AI*. Cambridge, MA: MIT Press.

Brooks, R. A. 1991a. Intelligence without Reason. *Pages 569–595 of: Proceedings of the Twelfth International Joint Conference on Artificial Intelligence*.

Brooks, R. A. 1991b. Intelligence without Representation. *Artificial Intelligence*, **47**(1-3), 139–159.

Buchanan, B. G., & Feigenbaum, E. A. 1978. Dendral and Meta-Dendral: Their Applications Dimension. *Artif. Intell.*, **11**(1-2), 5–24.

Buss, M., & Beetz, M. 2010. CoTeSys – Cognition for Technical Systems. *Künstliche Intelligenz*.

Bylander, T. 1991. Complexity Results for Planning. *Pages 274–279 of: Proceedings of the Twelfth International Joint Conference on Artificial Intelligence*.

Cangelosi, A., & Schlesinger, M. 2015. *Developmental robotics: From babies to robots*. Cambridge, MA: MIT Press/Bradford Books.

Canny, J. F. 1988. *The Complexity of Robot Motion Planning*. MIT Press.

CARETeam. 2009. *Robotic Visions to 2020 and beyond — The Strategic Research Agenda (SRA) for robotics in Europe*. http://www.robotics-platform.eu/cms/index.php?idcatart=119& client=1& lang=1?

Christensen, H. I. 2011. Cognitive Systems - From Internet to Robotics. *ERCIM News*, **2011**(84), 14–15.

Collet Romea, A., Martinez Torres, M., & Srinivasa, S. 2011. The MOPED framework: Object recognition and pose estimation for manipulation. *International Journal of Robotics Research*, **30**(10), 1284 – 1306.

Dario, P., Verschure, P. F. M. J., Prescott, T. J., Cheng, G., Sandini, G., Cingolani, R., Dillmann, R., Floreano, D., Leroux, C., MacNeil, S., Roelfsema, P., Verykios, X., Bicchi, A., Melhuish, C., & Albu-Schäffer, A. 2011b. Robot Companions for Citizens. *Procedia CS*, **7**, 47–51.

Decety, J. 1996. Do imagined and executed actions share the same neural substrate? *Cognitive Brain Research*, **3**(2), 87–93.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. 2009. ImageNet: A Large-Scale Hierarchical Image Database. *In: CVPR09*.

Eiter, T., Lutz, C., Ortiz, M., & Simkus, M. 2009. Query Answering in Description Logics with Transitive Roles. *Pages 759–764 of: IJCAI 2009, Proceedings of the 21st International Joint Conference on Artificial Intelligence, Pasadena, California, USA, July 11-17, 2009*.

Elbrechter, C., Haschke, R., & Ritter, H. 2012. Folding paper with anthropomorphic robot hands using real-time physics-based modeling. *Pages 210–215 of: Humanoid Robots (Humanoids), 2012 12th IEEE-RAS International Conference on*. IEEE.

Erol, K., Nau, D. S., & Subrahmanian, V. S. 1995. Complexity, Decidability and Undecidability Results for Domain-Independent Planning. *Artif. Intell.*, **76**(1-2), 75–88.

Fedrizzi, A. 2010 (November). *Action-Related Places for Mobile Manipulation*. Ph.D. thesis, Technische Universiät München.

Feigenbaum, E. A. 1993. Tiger in a Cage: The Applications of Knowledge-based Systems (1993) - Abstract. *Page 852 of:* Fikes, R., & Lehnert, W. G. (eds), *AAAI*. AAAI Press / The MIT Press.

Feldman, J., & Narayanan, S. 2011 (August). Simulation Semantics, Embodied Construction Grammar, and the Language of Events. *In: Language Action Tools for Cognitive Artificial Agents. Workshop Proceedings of the Conference of the American Association for Artificial Intelligence (AAAI)*.

Feldman, J., & Narayanan, S. 2004. Embodied meaning in a neural theory of language. *Brain and Language*, **89**(2), 385–392.

Fellbaum, C. 1998. *WordNet: An Electronic Lexical Database*. MIT Press.

Ferrucci, D., & Lally, A. 2004a. UIMA: an architectural approach to unstructured information processing in the corporate research environment. *Natural Language Engineering*, **10**(3-4), 327–348.

Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A. A., Lally, A., Murdock, J. W., Nyberg, E., Prager, J., Schlaefer, N., & Welty, C. 2010. Building Watson: An Overview of the DeepQA Project. *AI Magazine*, **31**(3), 59–79.

Fox, M., & Long, D. 2001. Hybrid STAN: Identifying and managing combinatorial optimisation sub-problems in planning. *Pages 445–452 of: Proceedings of International Joint Conference on Artificial Intelligence*.

Ghallab, M. 1996. On Chronicles: Representation, On-line Recognition and Learning. *Pages 597–607 of: PRINCIPLES OF KNOWLEDGE REPRESENTATION AND REASONING-INTERNATIONAL CONFERENCE-*. MORGAN KAUFMANN PUBLISHERS.

Gibbs, R. A., Belmont, J. W., Hardenbol, P., Willis, T. D., Yu, F., Yang, H., Ch'ang, L.-Y., Huang, W., Liu, B., Shen, Y., *et al.*. 2003. The international HapMap project. *Nature*, **426**(6968), 789–796.

Giovannetti, Tania, Schwartz, Myrna, Buxbaum, & Laurel. 2007. The Coffee Challenge: A new method for the study of everyday action errors. *Journal of Clinical and Experimental Neuropsychology*, **29**(7), 690–705.

Giovannetti, T., Libon, D. J., & Hart, T. 2002. Awareness of naturalistic action errors in dementia. *Journal of the International Neuropsychological Society JINS*, **8**(5), 633–644.

Gravot, F., Haneda, A., Okada, K., & Inaba, M. 2006. Cooking for humanoid robot, a task that needs symbolic and geometric reasonings. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 462–467.

Guerra-Filho, G., & Aloimonos, Y. 2007. A Language for Human Action. *IEEE Computer*, **40**(5), 42–51.

Gödel, K. 1931. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, **38**(1), 173–198.

Halevy, A., Norvig, P., & Pereira, F. 2009. The unreasonable effectiveness of data. *Intelligent Systems, IEEE*, **24**(2), 8–12.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. 2009. The WEKA Data Mining Software: An Update. *SIGKDD Explor. Newsl.*, **11**(1), 10–18.

Hammond, K. J., Converse, T. M., & Grass, J. W. 1995. The stabilization of environments. *Artificial Intelligence*, **72**(1–2), 305–327.

Hartigan, I. 2007. 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.

Hassabis, D., Kumaran, D., Vann, S. D., & Maguire, E. A. 2007. Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Sciences*, **104**(5), 1726–1731.

Hawrylycz, M. J., Lein, E. S., Guillozet-Bongaarts, A. L., Shen, E. H., Ng, L., Miller, J. A., van de Lagemaat, L. N., Smith, K. A., Ebbert, A., Riley, Z. L., *et al.*. 2012. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, **489**(7416), 391–399.

Hayes, P. J. 1985a. Naive Physics I: Ontology for Liquids. *Pages 71–107 of:* Hobbs, J. R., & Moore, R. C. (eds), *Formal Theories of the Commonsense World*. Norwood, NJ: Ablex.

Hayes, P. J. 1985b. The Second Naive Physics Manifesto. *Pages 1–36 of:* Hobbs, J. R., & Moore, R. C. (eds), *Formal Theories of the Commonsense World*. Norwood, NJ: Ablex.

Herdt, V., Le, H. M., Große, D., & Drechsler, R. 2016. ParCoSS: Efficient Parallelized Compiled Symbolic Simulation. *In: Computer Aided Verification*.

Herrmann, E., Call, J., Hernández-Lloreda, M. V., Hare, B., & Tomasello, M. 2007. Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *science*, **317**(5843), 1360–1366.

Hesslow, G. 2012. The current status of the simulation theory of cognition. *Brain Research*, **1428**, 71 – 79. The Cognitive Neuroscience of Thought.

Hinterstoisser, S., Holzer, S., Cagniart, C., Ilic, S., Konolige, K., Navab, N., & Lepetit, V. 2011. Multimodal templates for real-time detection of texture-less objects in heavily cluttered scenes. *Pages 858–865 of:* Metaxas, D. N., Quan, L., Sanfeliu, A., & Gool, L. J. V. (eds), *ICCV*. IEEE.

Horswill, I. 1995. Analysis of Adaptation and Environment. *Artificial Intelligence*, **73**, 1–30.

Horswill, I. 1996a. *Integrated Systems and Naturalistic Tasks*. In: Strategic Directions in Computing Research AI Working Group.

Horswill, I. D. 1994. *Specialization of Perceptual Processes*. Tech. rept. Massachusetts Institute of Technology.

Hothorn, T., & Everitt, B. S. 2014. *A Handbook of Statistical Analyses Using R*. 3rd edn. Boca Raton, Florida, USA: Chapman & Hall/CRC Press.

Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A., & Fitzgibbon, A. 2011. KinectFusion: realtime 3D reconstruction and interaction using a moving depth camera. *Pages 559–568 of: Proceedings of the 24th annual ACM symposium on User interface software and technology*. UIST '11. New York, NY, USA: ACM.

Jain, D., Mösenlechner, L., & Beetz, M. 2009. Equipping Robot Control Programs with First-Order Probabilistic Reasoning Capabilities. *Pages 3626–3631 of: IEEE International Conference on Robotics and Automation (ICRA)*.

Jasny, B., & Kennedy, D. 2001 (February). *The Human Genome*. Science, Special Issue.

Jeannerod, M. 2001. Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage*, **14**(1 Pt 2), S103–109.

Jenkins, O. C., & Matarić, M. J. 2004b (Jul). A Spatio-temporal Extension to Isomap Nonlinear Dimension Reduction. *Pages 441–448 of: The International Conference on Machine Learning (ICML 2004)*.

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., & Darrell, T. 2014. Caffe: Convolutional architecture for fast feature embedding. *Pages 675–678 of: Proceedings of the ACM International Conference on Multimedia*. ACM.

Kemp, C. C., Edsinger, A., & Torres-Jara, E. 2007. Challenges for Robot Manipulation in Human Environments. *IEEE Robotics and Automation Magazine*, **14**(1), 20–29.

Klein, J., & Zachmann, G. 2004a. Point Cloud Collision Detection. *Computer Graphics forum (Proc. EUROGRAPHICS)*, **23**(3), 567–576.

Klein, J., & Zachmann, G. 2004b. Point Cloud Surfaces using Geometric Proximity Graphs. *Computers & Graphics*, **28**(6), 839–850.

Kontchakov, R., Lutz, C., Toman, D., Wolter, F., & Zakharyaschev, M. 2010a. The Combined Approach to Query Answering in DL-Lite. *In: Principles of Knowledge Representation and Reasoning: Proceedings of the Twelfth International Conference, KR 2010, Toronto, Ontario, Canada, May 9-13, 2010*.

Körding, K. P., & Wolpert, D. M. 2004a. Bayesian Integration in Sensorimotor Learning. *Nature*, **427**, 244–247.

Körding, K. P., & Wolpert, D. M. 2004b. Probabilistic Inference in Human Sensorimotor Processing. *In:* Thrun, S., Saul, L. K., & Schölkopf, B. (eds), *Advances in Neural Information Processing Systems 16: Proceedings of the 2003 Conference*. MIT Press.

Kosslyn, S. M., & Rabin, C. 1999. *Mental imagery*. Cambridge, MA: MIT Press.

Kosslyn, S. M., Thompson, W. L., & Alpert, N. M. 1997. Neural systems shared by visual imagery and visual perception: a positron emission tomography study. *Neuroimage*, **6**(4), 320–334.

Kraemer, D. J., Macrae, C. N., Green, A. E., & Kelley, W. M. 2005. Musical imagery: sound of silence activates auditory cortex. *Nature*, **434**(7030), 158.

Krisnadhi, A., & Lutz, C. 2007. Data Complexity in the EL family of DLs. *In: Proceedings of the 2007 International Workshop on Description Logics (DL2007), Brixen-Bressanone, near Bozen-Bolzano, Italy, 8-10 June, 2007*.

Kunze, L., Tenorth, M., & Beetz, M. 2010. Putting People's Common Sense into Knowledge Bases of Household Robots. *Pages 151–159 of: 33rd Annual German Conference on Artificial Intelligence (KI 2010)*. Karlsruhe, Germany: Springer.

Lakshmanan, K., Sachdev, A., Xie, Z., Berenson, D., Goldberg, K., & Abbeel, P. 2012. A Constraint-Aware Motion Planning Algorithm for Robotic Folding of Clothes. *In: Proceedings of the 13th International Symposium on Experimental Robotics (ISER)*.

Latombe, J.-C. 1991. *Robot Motion Planning*. Boston, MA: Kluwer Academic Publishers.

LaValle, S. M. 2006a. *Planning Algorithms*. Cambridge University Press.

Lawton, M. P., & Brody, E. M. 1969. Assessment of older people: self-maintaining and instrumental activities of daily living. *The Gerontologist*, **9**(3), 179–186.

Le, H. M., Große, D., Herdt, V., & Drechsler, R. 2013. Verifying SystemC using an intermediate verification language and symbolic simulation. *Pages 116:1–116:6 of: Design Automation Conference*.

Leidner, D., Dietrich, A., Beetz, M., & Albu-Schäffer, A. 2015. Knowledge-Enabled Parameterization of Whole-Body Control Strategies for Compliant Service Robots. *Autonomous Robots (AURO): Special Issue on Whole-Body Control of Contacts and Dynamics for Humanoid Robots*.

Lenat, D. B., & Feigenbaum, E. A. 1991. On the Thresholds of Knowledge. *Artif. Intell.*, **47**(1-3), 185–250.

Leskovec, J., & Sosič, R. 2014 (June). *Snap.py: SNAP for Python, a general purpose network analysis and graph mining tool in Python*. `http://snap.stanford.edu/snappy`.

Levesque, H. J., & Brachman, R. J. 1987. Expressiveness and tractability in knowledge representation and reasoning. *Computational Intelligence*, **3**, 78–93.

Li, Q., Haschke, R., & Ritter, H. 2015. A visuo-tactile control framework for manipulation and exploration of unknown objects. *Pages 610–615 of: Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on*. IEEE.

Lisca, G., Nyga, D., Bálint-Benczédi, F., Langer, H., & Beetz, M. 2015. Towards Robots Conducting Chemical Experiments. *In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

Liu, H., & Singh, P. 2004b. ConceptNet—a practical commonsense reasoning tool-kit. *BT technology journal*, **22**(4), 211–226.

Long, D., & Fox, M. 1998. Domain-independent planner compilation. *In: Working notes of the Workshop on Knowledge Engineering and Acquisition for Planning, held in conjunction with AIPS-98. unpublished material*.

Lutz, C. 2008. The Complexity of Conjunctive Query Answering in Expressive Description Logics. *Pages 179–193 of: Automated Reasoning, 4th International Joint Conference, IJCAR 2008, Sydney, Australia, August 12-15, 2008, Proceedings*.

Lutz, C., & Wolter, F. 2012. Non-Uniform Data Complexity of Query Answering in Description Logics. *In: Principles of Knowledge Representation and Reasoning: Proceedings of the Thirteenth International Conference, KR 2012, Rome, Italy, June 10-14, 2012*.

Matuszek, C., Cabral, J., Witbrock, M. J., & DeOliveira, J. 2006b. An Introduction to the Syntax and Content of Cyc. *Pages 44–49 of: AAAI Spring Symposium: Formalizing and Compiling Background Knowledge and Its Applications to Knowledge Representation and Question Answering*. Citeseer.

Maycock, J., Dornbusch, D., Elbrechter, C., Haschke, R., Schack, T., & Ritter, H. J. 2010. Approaching manual intelligence. *KI - Künstliche Intelligenz*, 1–8.

Human Brain Project Team. 2012. *The Human Brain Project — A Report to the European Commission*. Tech. rept. .

McCarthy, J. 1968. Programs with Common Sense. *Pages 403–418 of: Semantic Information Processing*. MIT Press.

McCarthy, J., & Hayes, P. J. 1969. Some Philosophical Problems from the Standpoint of Artificial Intelligence. *Pages 463–502 of:* Meltzer, B., & Michie, D. (eds), *Machine Intelligence 4*. Edinburgh University Press.

McDermott, D. 1982. A Temporal Logic for Reasoning About Processes and Plans. *Cognitive Science*, **6**(2), 101–155.

McDermott, D. 1985. Reasoning about Plans. *Pages 269–317 of:* Hobbs, J. R., & Moore, R. C. (eds), *Formal Theories of the Commonsense World*. Norwood, NJ: Ablex.

McDermott, D. 1992a. *Transformational planning of reactive behavior*. Tech. rept. Department of Computer Science, Yale University, New Haven, CT.

McDermott, D. 1987. A critique of pure reason. *Compu-*

tational Intelligence, **3**, 151–160.

McDermott, D. 1992b. *Transformational Planning of Reactive Behavior*. Research Report YALEU/DCS/RR-941. Yale University.

Metta, G., Vernon, D., & Sandini, G. 2005. *The RobotCub Approach to the Development of Cognition*.

Metta, G., Fitzpatrick, P., & Natale, L. 2006. YARP: Yet Another Robot Platform. *International Journal of Advanced Robotics Systems, special issue on Software Development and Integration in Robotics*, **3**(1), 43–48.

Miller, G. A. 1995. WordNet: a lexical database for English. *Communications of the ACM*, **38**(11), 39–41.

Minsky, M. 1986. *The Society of Mind*. New York, NY, USA: Simon & Schuster, Inc.

Mohr, D., & Zachmann, G. 2010a. FAST: Fast Adaptive Silhouette Area based Template Matching. *Pages 39.1–39.12 of: Proceedings of the British Machine Vision Conference (BMVC)*. BMVA Press.

Mohr, D., & Zachmann, G. 2010b. Silhouette Area Based Similarity Measure for Template Matching in Constant Time. *Pages 43–54 of: 6th International Conference of Articulated Motion and Deformable Objects (AMDO)*. Port d'Andratx, Mallorca, Spain: Springer Verlag.

Morgenstern, L. 2001. Mid-Sized Axiomatizations of Commonsense Problems: A Case Study in Egg Cracking. *Studia Logica*, **67**(3), 333–384.

Morris, R., Tarassenko, L., & Kenward, M. (eds). 2005. *Cognitive Systems: Information processing meets brain science*. Academic Press Inc (London) Ltd. Foresight Report.

Mörwald, T., Prankl, J., Richtsfeld, A., Zillich, M., & Vincze, M. 2010. BLORT - The Blocks World Robotic Vision Toolbox. *In: Best Practice in 3D Perception and Modeling for Mobile Manipulation (in conjunction with ICRA 2010)*.

Mösenlechner, L., Demmel, N., & Beetz, M. 2010 (October 18-22). Becoming Action-aware through Reasoning about Logged Plan Execution Traces. *Pages 2231–2236 of: IEEE/RSJ International Conference on Intelligent RObots and Systems.*

Motik, B. 2008. KAON2 - Scalable Reasoning over Ontologies with Large Data Sets. *ERCIM News*, **2008**(72).

Müller, A., Kirsch, A., & Beetz, M. 2007 (September). Transformational Planning for Everyday Activity. *Pages 248–255 of: Proceedings of the 17th International Conference on Automated Planning and Scheduling (ICAPS'07)*.

Murphy, R., Tadokoro, S., Nardi, D., Jacoff, A., Fiorini, P., & Erkmen, A. 2008. *Rescue Robotics*. Springer Publishers. "in Handbook of Robotics, B. Sciliano and O. Khatib, Eds.". Chap. " ", pages 1151–1174.

Nebel, B. 1994. Artificial Intelligence: A Computational Perspective. *Pages 237–266 of: Essentials in Knowledge Representation*. CSLI publications.

Newborn, M. 2000. Deep Blue's contribution to AI. *Ann. Math. Artif. Intell.*, **28**(1-4), 27–30.

Newcombe, R. A., Davison, A. J., Izadi, S., Kohli, P., Hilliges, O., Shotton, J., Molyneaux, D., Hodges, S., Kim, D., & Fitzgibbon, A. 2011. KinectFusion: Real-time dense surface mapping and tracking. *Pages 127–136 of: Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on*.

Newell, A., & Simon, H. A. 1976. Computer Science as Empirical Inquiry: Symbols and Search. *Communications of the ACM*, **19**(3), 113–126.

Nielsen, M. 2012. *Reinventing discovery: The New Era of Networked Science*. Princeton University Press.

Nilsson, N. J. 1991. Logic and Artificial Intelligence. *Artif. Intell.*, **47**(1-3), 31–56.

Nyga, D., & Beetz, M. 2012 (October, 7–12). Everything Robots Always Wanted to Know about Housework (But were afraid to ask). *In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

Nyga, D., & Beetz, M. 2015. Cloud-based Probabilistic Knowledge Services for Instruction Interpretation. *In: International Symposium of Robotics Research (ISRR)*.

Nyga, D., Tenorth, M., & Beetz, M. 2011 (May, 9–13). How-Models of Human Reaching Movements in the Context of Everyday Manipulation Activities. *In: IEEE International Conference on Robotics and Automation (ICRA)*.

Nyga, D., Bálint-Benczédi, F., & Beetz, M. 2014 (May 31-June 7). PR2 Looking at Things: Ensemble Learning for Unstructured Information Processing with Markov Logic Networks. *In: IEEE International Conference on Robotics and Automation (ICRA)*.

Okada, K., Kojima, M., Sagawa, Y., Ichino, T., Sato, K., & Inaba, M. 2006. Vision based behavior verification system of humanoid robot for daily environment tasks. *Pages 7–12 of: Proceedings of the 6th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*.

Pastra, K., & Aloimonos, Y. 2011. The minimalist grammar of action. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **367**(1585), 103–117.

Pfeifer, R., & Scheier, C. 2001. *Understanding Intelligence*. Bradford Books. University Press Group Limited.

Prats, M., Wieland, S., Asfour, T., Del Pobil, A. P., & Dillmann, R. 2008. Compliant interaction in household environments by the Armar-III humanoid robot. *Pages 475–480 of: Humanoids 2008-8th IEEE-RAS International Conference on Humanoid Robots*. IEEE.

Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., Berger, E., Wheeler, R., & Ng, A. 2009 (May 12–17). ROS: an open-source Robot Operating System. *In: In IEEE International Conference on Robotics and Automation (ICRA)*.

R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ramirez-Amaro, K., Beetz, M., & Cheng, G. 2015c. Understanding the intention of human activities through semantic perception: observation, understanding and execution on a humanoid robot. *Advanced Robotics*, **29**(5), 345–362.

Renz, J., & Nebel, B. 1999. On the Complexity of Qualitative Spatial Reasoning: A Maximal Tractable Fragment of the Region Connection Calculus. *Artif. Intell.*, **108**(1-2), 69–123.

Rizzolatti, G., Fogassi, L., & Gallese, V. 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, **2**, 661–669.

Rizzolatti, G. 1998. Language within our grasp. *Trends in Neurosciences*, **21**(5), 188–194.

Rizzolatti, G., & Luppino, G. 2001. The Cortical Motor System. *Neuron*, **31**(6), 889–901.

Rosenbaum, D. A., Breckenridge, M., & Murphy, M. 1969. *Human Motor Control*. Academic Press.

Rosenbaum, D. A., Chapman, K. M., Weigelt, M., Weiss, D. J., & van der Wel, R. 2012. Cognition, action, and object manipulation. *Psychological bulletin*, **138**(5), 924.

Russell, S. J., & Norvig, P. 2014. *Artificial Intelligence: a modern approach*. New international edn. Harlow, England: Pearson Education Limited.

Schaal, S. 2007. The New Robotics – Towards Human-centered Machines. *HFSP Journal Frontiers of Interdisciplinary Research in the Life Sciences*, **1**(2), 115–126.

Schank, R., & Abelson, R. 1977. *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schill, K., Zetzsche, C., & Hois, J. 2009. A Belief-based Architecture for Scene Analysis: From Sensorimotor Features to Knowledge and Ontology. *Fuzzy Sets and Systems*, **160**(10), 1507–1516.

Schnitzler, A., Salenius, S., Salmelin, R., Jousmaki, V., & Hari, R. 1997. Involvement of primary motor cortex in motor imagery: a neuromagnetic study. *Neuroimage*, **6**(3), 201–208.

Schwartz, M. F., & Buxbaum, L. J. 1997. *Naturalistic action*. Psychology Press. Pages 269–289.

Schwartz, M. F., Segal, M., Veramonti, T., Ferraro, M., & Buxbaum, L. J. 2002. The Naturalistic Action Test: A standardized assessment for everyday action impairment. *Neuropsycological Rehabilitation*, **12**(4), 311–339.

Sellen, J. 1996. Lower Bounds for Geometrical and Physical Problems. *SIAM J. Comput.*, **25**(6), 1231–1253.

Shadmehr, R., & Mussa-Ivaldi, F. 1994. Adaptive representation of dynamics during learning of a motor task. *Journal of Neuroscience*, **14**(5), 3208–3224.

Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., Blake, A., Cook, M., & Moore, R. 2013. Real-time human pose recognition in parts from single depth images. *Communications of the ACM*, **56**(1), 116–124.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. 2016. Mastering the game of Go with deep neural networks and tree search. *Nature*, **529**, 484–503.

Simon, H. 1956. Rational choice and the structure of the environment. *Psychological Review*, **63**, 129–138.

Singh, P., Lin, T., Mueller, E. T., Lim, G., Perkins, T., & Zhu, W. L. 2002. Open Mind Common Sense: Knowledge Acquisition from the General Public. *Pages 1223–1237 of: Proceedings of the First International Conference on Ontologies, Databases, and Applications of Semantics for Large Scale Information Systems*.

Slavicsek, B. 2000. *A Guide to the Star Wars Universe*. Lucas Books. Random House Publishing Group.

Smits, R., de Laet, T., Claes, K., Soetens, P., de Schutter, J., & Bruyninckx, H. 2008. Orocos: A Software Framework for Complex Sensor-Driven Robot Tasks. *IEEE Robotics and Automation Magazine*.

Srinivasa, S. S., Ferguson, D., Helfrich, C., Berenson, D., Collet, A., Diankov, R., Gallagher, G., Hollinger, G. A., Kuffner, J., & Weghe, M. V. 2010. HERB: a home exploring robotic butler. *Auton. Robots*, **28**(1), 5–20.

Stork, D. G. 1997. Happy Birthday, HAL! *IEEE Computer*, **30**(1), 24–25.

Structure, T., & of the Universe Roadmap Team, E. 2003. *Beyond Einstein: from the big bang to black holes*. Research Roadmap. Nasa.

Stuckler, J., Holz, D., & Behnke, S. 2012. Demonstrating everyday manipulation skills in Robocup@ home. *IEEE Robotics and Automation Magazine*, 34–42.

Suddendorf, T., & Corballis, M. C. 2007. The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences*, **30**(03), 299–313.

Tenorth, M., & Beetz, M. 2012 (October, 7–12). A Unified Representation for Reasoning about Robot Actions, Processes, and their Effects on Objects. *In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

Tenorth, M., & Beetz, M. 2013. KnowRob – A Knowledge Processing Infrastructure for Cognition-enabled Robots. Part 1: The KnowRob System. *International Journal of Robotics Research (IJRR)*, **32**(5), 566–590.

Tenorth, M., & Beetz, M. 2015. Representations for robot knowledge in the KnowRob framework. *Artificial Intelligence*.

Tenorth, M., Bandouch, J., & Beetz, M. 2009. The TUM Kitchen Data Set of Everyday Manipulation Activities for Motion Tracking and Action Recognition. *In: IEEE International Workshop on Tracking Humans for the Evaluation of their Motion in Image Sequences (THEMIS), in conjunction with ICCV2009*.

Tenorth, M., Klank, U., Pangercic, D., & Beetz, M. 2011. Web-enabled Robots – Robots that Use the Web as an Information Resource. *Robotics & Automation Magazine*, **18**(2), 58–68.

Tenorth, M., Perzylo, A. C., Lafrenz, R., & Beetz, M. 2012 (May 14–18). The RoboEarth language: Representing and Exchanging Knowledge about Actions, Objects, and Environments. *In: IEEE International Conference on Robotics and Automation (ICRA)*. Best Cognitive Robotics Paper Award.

Tenorth, M., Profanter, S., Bálint-Benczédi, F., & Beetz, M. 2013a (November 3–7). Decomposing CAD Models of Objects of Daily Use and Reasoning about their Functional Parts. *Pages 5943–5949 of: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

Thrun, S., Burgard, W., & Fox, D. 2005. *Probabilistic Robotics*. Cambridge: MIT Press.

Todorov, E. 2004. Optimality principles in sensorimotor control. *Nature Neuroscience*, **7**(9), 907–915.

Tulving, E. 1972. Episodic and semantic memory. *Chap. 10, pages 381–403 of:* Tulving, E., & Donaldson, W. (eds), *Organization of memory*. New York: Academic Press.

Tulving, E. 2002a. Episodic Memory: From Mind to Brain. *Annual review of psychology*, **53**(1), 1–25.

Tulving, E. 1995. Organization of memory: Quo vadis. *The cognitive neurosciences*, **839847**.

Vernon, D., Beetz, M., & Sandini, G. 2015. Prospection in Cognition: The Case for Joint Episodic-Procedural Memory in Cognitive Robotics. *Frontiers in Robotics and AI*, **2**(19).

Waibel, M., Beetz, M., D'Andrea, R., Janssen, R., Tenorth, M., Civera, J., Elfring, J., Gálvez-López, D., Häussermann, K., Montiel, J., Perzylo, A., Schießle, B., Zweigle, O., & van de Molengraft, R. 2011. RoboEarth - A World Wide Web for Robots. *Robotics & Automation Magazine*, **18**(2), 69–82.

Walther-Franks, B., Smeddinck, J., Szmidt, P., Haidu, A., Beetz, M., & Malaka, R. 2015. Robots, Pancakes, and Computer Games: Designing Serious Games for Robot Imitation Learning. *In: CHI Conference on Human Factors in Computing Systems*.

Weller, R., & Zachmann, G. 2010. ProtoSphere: A GPU-Assisted Prototype-Guided Sphere Packing Algorithm for Arbitrary Objects. *Pages 8:1–8:2 of: ACM SIGGRAPH ASIA 2010 Sketches*. New York, NY, USA: ACM.

Weller, R., & Zachmann, G. 2011. Inner Sphere Trees and Their Application to Collision Detection. *Chap. 10, pages 181–202 of:* Coquillart, S., Brunnett, G., & Welch, G. (eds), *Virtual Realities*. Springer (Dagstuhl).

Winkler, J., Tenorth, M., Bozcuoglu, A. K., & Beetz, M. 2014. CRAMm – Memories for Robots Performing Everyday Manipulation Activities. *Advances in Cognitive Systems*, **3**, 47–66.

Wu, Z., & Palmer, M. 1994. Verbs semantics and lexical selection. *Pages 133–138 of: Proceedings of the 32nd annual meeting on Association for Computational Linguistics*. Association for Computational Linguistics Morristown, NJ, USA.

Zachmann, G. 1998 (March). Rapid Collision Detection by Dynamically Aligned DOP-Trees. *Pages 90–97 of: Proc. of IEEE Virtual Reality Annual International Symposium; VRAIS '98*.

Zachmann, G. 2002 (Nov.11–13). Minimal Hierarchical Collision Detection. *Pages 121–128 of: Proc. ACM Symposium on Virtual Reality Software and Technology (VRST)*.

Zachmann, G., & Weller, R. 2006 (June 14–17). Kinetic Bounding Volume Hierarchies for Deformable Objects. *In: ACM Int'l Conf. on Virtual Reality Continuum and Its Applications (VRCIA)*.

Ziebart, B., Ratliff, N., Gallagher, G., Mertz, C., Peterson, K., Bagnell, J. A. D., Hebert, M., Dey, A., & Srinivasa, S. 2009 (October). Planning-based Prediction for Pedestrians. *In: Proc. IROS 2009*.