

The Assistive Kitchen — A Demonstration Scenario for Cognitive Technical Systems⁰

Michael Beetz, Jan Bandouch, Alexandra Kirsch, Alexis Maldonado, Armin Müller, Radu Bogdan Rusu
Intelligent Autonomous Systems, Department of Informatics
Technische Universität München, Germany

Abstract. This paper introduces the *Assistive Kitchen* as a comprehensive demonstration and challenge scenario for technical cognitive systems. We describe its hardware and software infrastructure. Within the *Assistive Kitchen* application, we select particular domain activities as research subjects and identify the cognitive capabilities needed for perceiving, interpreting, analyzing, and executing these activities as research foci. We conclude by outlining open research issues that need to be solved to realize the scenarios successfully.

I. Introduction

Cognitive technical systems are systems that are equipped with artificial sensors and actuators, integrated into physical systems, and act in a physical world. They differ from other technical systems in that they have *cognitive capabilities* including perception, reasoning, learning, and planning that turn them into systems that “*know what they are doing*” [4]. The cognitive capabilities will result in systems of higher reliability, flexibility, adaptivity, and better performance and systems that are easier to interact and cooperate with.

The cluster of excellence CoTESYS (Cognition for Technical Systems)¹ considers the assistance of elder people to be a key application where technical cognitive systems could profoundly impact the well-being of our society. Therefore, CoTESYS investigates the realization of an *assistive kitchen* (Figure 1), a ubiquitous computing, sensing, and actuation environment with a robotic assistant as one of its primary demonstration scenarios. The Assistive Kitchen aims at

- supporting and assisting people in their household chores through physical action;
- enhancing the cognitive capabilities of people doing household work by reminding them; and
- monitoring health and safety of the people.

To achieve these objectives, the Assistive Kitchen is to

- perceive, interpret, learn, and analyze models of household chore and activities of daily life (ADLs); and
- represent the acquired models such that the Assistive Kitchen can use them for activity and safety monitoring, health assessment, and for adapting itself to the needs and preferences of the people.

The Assistive Kitchen includes an autonomous robotic agent that is to learn and perform complex household chores. The robot must perform housework together with people or at least assist them in their activities. This requires safe operation in the presence of humans and behaving according to the preferences of the people they serve.

Clearly, assistive kitchens of this sort are important for sev-

eral reasons. First, they are of societal importance because they can enable persons with minor disabilities including sensory, cognitive, and motor ones to live independently and to perform their household work. This will increase the quality of life as well as reduce the cost of home care.

Assistive kitchens and living environments also raise challenging research problems. One of these problems is that performing household chores is a form of everyday activity that requires extensive commonsense knowledge and reasoning [1]. Another challenge is the low frequency of daily activities, which requires embedded systems and robotic agents to learn from very scarce experience. Besides, household chores include a large variety of manipulation actions and composed activities that pose hard research questions for current robot manipulation research. The management of daily activities also requires activity management that is very different from that commonly assumed by AI planning systems.

II. Assistive Kitchen Infrastructure

We start with the hardware and software infrastructure of the kitchen — the implementational basis that defines the possibilities and restrictions of the demonstration scenarios.

A. The Hardware Infrastructure

The hardware infrastructure consists of a mobile robot and networked sensing and actuation devices that are physically embedded into the environment.

A. 1. The Autonomous Mobile Robot

Currently, an autonomous mobile robot with two arms with grippers acts as a robotic assistant in the Assistive Kitchen (see Figure 1). The robot is a RWI B21 robot equipped with a stereo CCD system and laser rangefinders as its primary sensors. One laser range sensor is integrated into the robot base to allow for estimating the robot’s position within the environment. Small laser range sensors are mounted onto the robot’s grippers to provide sensory feedback for reaching and grasping actions. The grippers are also equipped with RFID tag readers that support object detection and identification. Cameras are used for longer range object recognition and to allow for vision-based interaction with people.

The robot can manipulate objects and its environment using its two industrial strength Amtec Powercube arms with simple grippers. While the arms and grippers are not very dexterous they permit the execution of simple manipulation actions such

⁰(1) The research reported in this paper is partly funded by the German cluster of excellence CoTESYS (Cognition for Technical Systems). (2) More information including videos and publications about the *Assistive Kitchen* can be found at ias.cs.tum.edu/assistivekitchen. (3) Due to space limitations this paper does not contain an outline of our research agenda towards realizing the scenarios and nor does it give a comprehensive discussion of related work.

¹The research agenda of the CoTESYS cluster of excellence is described in a companion paper [5] contained in this collection.

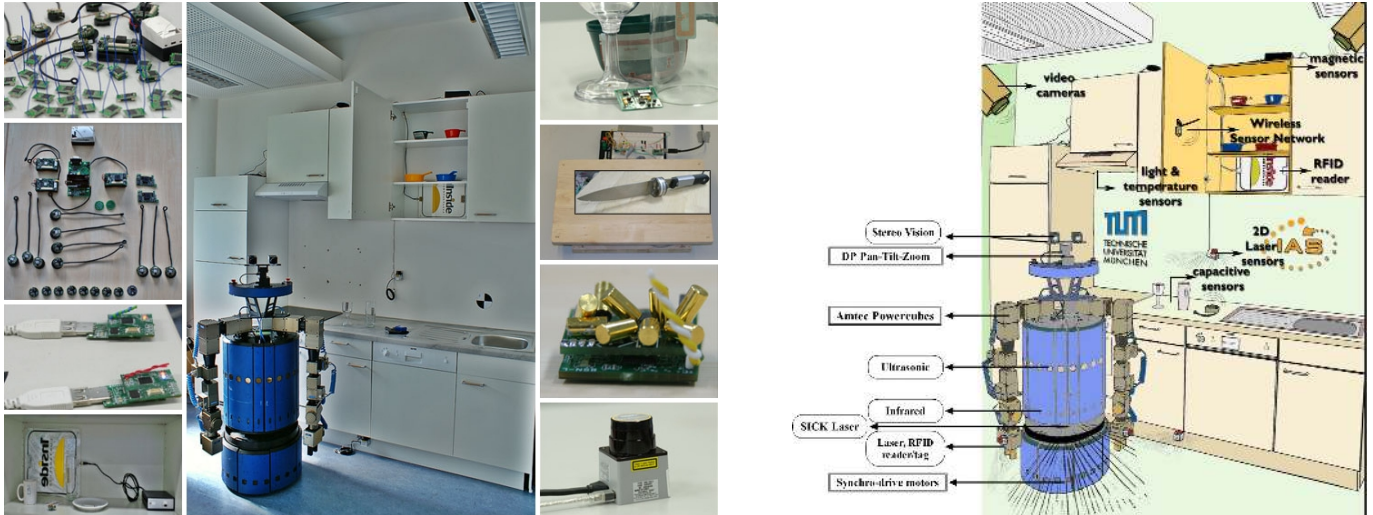


Fig.1: The Assistive Kitchen containing a robot and a variety of sensors.

as getting plates and glasses out of the cupboard and putting them onto the table.²

A. 2. Room Infrastructure

The sensor-equipped kitchen environment (see Figure 1) disposes of global environment sensors, sensor-equipped furniture, web-enabled appliances, “smart” objects, and instrumentations for people acting in the environment [6]. In this section we will look at these components in more detail.

Global Sensors of the Kitchen. We have mounted a set of static off-the-shelf cameras positioned to cover the kitchen area with high resolution in critical working areas. With these cameras, actions of the people and robots can be tracked from different locations to allow for more accurate positioning and pose estimation. In addition, laser range sensors are mounted at the walls for covering large parts of the kitchen. They provide accurate and valuable position data for the people present in the environment and their movements within the kitchen.

Sensor-equipped Furniture. The pieces of furniture in the Assistive Kitchen are also equipped with various kinds of sensors. For example, we have cupboards with long-range RFID tag readers that enable the cupboards to “know” the identities of the RFID tagged objects that are currently in the cupboard. Additionally, the cupboards are equipped with magnetic contact sensors that sense whether the cupboard doors are open or closed. Another example is a table which contains several integrated capacitive sensors as well as short-range RFID readers. The capacitive sensors report the capacitance of different areas on the table, when an object is placed there, while the RFID readers provide exact information on what object was placed there.

Web-enabled Kitchen Appliances such as the refrigerator, the oven, the microwave, and the faucet, allow for remote and wireless monitoring and control.

“Smart” Objects In addition, kitchen utensils, tools and small appliances are equipped with integrated sensors. For example, we use a knife (see Figure 2) instrumented with a

6DOF force/torque sensor that allows us to record the force trajectories over extended periods of time. Because the shapes of the force trajectories are characteristic for the physical properties of the objects, we can learn object specific force profiles and use them to classify the objects being cut.

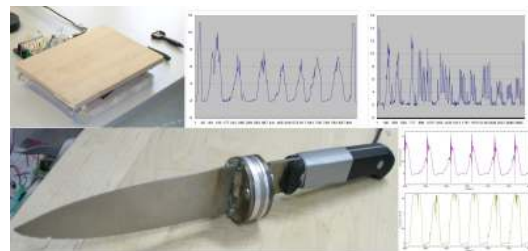


Fig.2: Knife with embedded force sensors networked within the wireless sensor network.

Another smart object is a sensor-instrumented coffee machine, its capabilities being extended from that of a normal kitchen appliance. The integrated sensors provide information whether the filter unit is open or closed and whether a coffee filter is installed, whether the machine is switched on or off, the amount of water in the water container, etc.

Instrumentation of People Small ubiquitous devices offer the possibility to instrument people acting in the environment with additional sensors. In our case, we have built a glove equipped with an RFID tag reader (see Figure 3) that enables us to identify the objects that are manipulated by the person wearing it. In addition, the person is equipped with tiny inertial measurement units that provide us with detailed information about the person’s limb motions.

Another body worn sensory device to be used in the Assistive Kitchen demonstration scenario is the gaze-aligned head mounted camera which allows the estimation of the attentional state of people while performing their kitchen work.³

²Robot platforms contributed by other members of the CoTESYS cluster include *Justin* (dexterous manipulation, www.robotic.dlr.de) and *Johnnie/Lola* (full body motion, www.amm.mw.tu-muenchen.de).

³The gaze-aligned head mounted camera is currently developed by Neurologische Klinik und Poliklinik, Ludwig-Maximilians-Universität München, which participates in the CoTESYS excellence cluster (www.forbias.de/).

B. Cognitive Sensor Networks

The sensors in the Assistive Kitchen are connected into distributed sensor networks (see Figure 3), which are enhanced with cognitive capabilities. To this end the sensors are wirelessly connected to small ubiquitous computing devices (like Gumstix) and to personal computers that perform state estimation and data-mining tasks. This way, activity data can be collected and abstracted into models in a distributed manner. Cognitive capabilities of the network include the estimation of meaningful states from sensor data, the continual acquisition, update, and use of activity models, and the estimation of context conditions that allow for the simplification of recognition tasks. Because of these capabilities cognitive sensor networks always have up-to-date models of objects, situations, and activities, which enable the networks to provide a continual service for answering queries about the environment and the activities that take place in it.

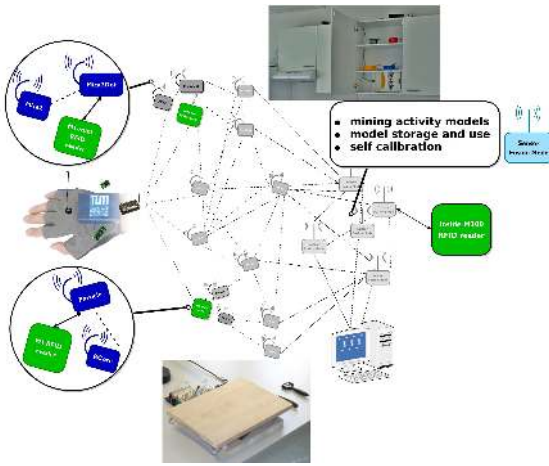


Fig.3: Sensor networks in the Assistive Kitchen.

Cognitive sensor networks can estimate states and recognize events that are meaningful in the application domain but must be obtained by combining, interpreting, and abstracting the sensor data of different sensors over time. For example, mounting RFID tags and acceleration sensors to kitchen utensils allows the environment to recognize *force-dynamic states* such as an object being picked up, carried, or put down. The recognition of force-dynamic states is essential for segmenting activities into meaningful subactions. The networks can also learn about places that play particular roles in the activities that are monitored. The system can learn where the people prepare food or where food is stored, etc.

C. Middleware Software Infrastructure

A critical factor for the successful implementation of the assistive kitchen is the middleware software infrastructure that has to provide a simple, reliable, uniform, and flexible interface for communicating with and controlling different physically distributed sensors and actuators.

We use and extend the open-source Player/Stage/Gazebo (P/S/G) software library to satisfy these requirements for the sensor-equipped environment as well as the robotic agent. Player provides a simple and flexible interface for robot control by making available powerful classes of interface abstractions for interacting with robot hardware, in particular sen-

sors and effectors. These abstractions enable the programmer to use devices with similar functionality with identical software interfaces, thus increasing the code transferability. An enhanced *client/server model* featuring auto-discovery mechanisms as well as permitting servers and clients to communicate between them in a heterogeneous network, enables programmers to code their clients in different programming languages. Programmers can also implement sophisticated algorithms and provide them as Player drivers. By incorporating well-understood algorithms into our infrastructure, we eliminate the need for users to individually re-implement them.

Using the P/S/G infrastructure a robot can enter a sensor-equipped environment, autodiscover the sensors and the services they provide and use these sensors in the same way as it uses its own sensors.

D. Simulation and Visualization

We have developed a simulation of the kitchen and robot acting as robotic assistant using the Gazebo toolbox for 3D, physic-based robot simulation. The simulator is realistic along many dimensions. In particular, the simulator will use sensing and actuation models that are learned from the real robots. The simulator is coupled to the cognitive sensor network. Thereby the simulator can be automatically initialized and updated according to the sensing data of the network.



Fig.4: The Assistive Kitchen and its simulation.

The use of these simulation tools promotes the research in assistive kitchen technology in various ways. First, the simulator supports generalization: we can model all kinds of robots in our simulation framework. We will also have different kitchen setups, which requires us to develop control programs that can specialize themselves for different robots and environments. The simulator allows us to run experiments fast and with little efforts and under controllable context settings. This supports the performance of extensive empirical studies.

III. Demonstration Scenarios

The demonstration scenarios are organized along two dimensions (see Figure 5). The first dimension is the domain tasks and activities under investigation. The second one are the cognitive aspects of perception, interpretation, learning, planning, and execution of these activities.

A. Activities/Scenario Tasks

We start by looking at three scenario tasks in the context of household chores: setting the table, cooking, and performing household chores for an extended period. These tasks challenge cognitive systems along different dimensions, which we will discuss in the remainder of this Section.

scenario/cognitive capability	domain task/activity		
	table setting	cooking	household chore
perception	*	*	*
activity model	*	*	*
env. model	*	*	
motion primitives	*	*	
macro actions	*	*	
everyday activity			*

Fig.5: Dimensions of demonstration scenarios.

Table setting refers to the arrangement of tableware, cutlery, and glasses on the table for eating. Table setting is a complex transportation task that can be performed as a single thread of activity.

There are various aspects of table setting that make it to a suitable challenge for cognitive systems: • the commonsense knowledge and reasoning needed are to perform the task competently; • the task itself is typically specified incompletely and in a fuzzy manner; and • the task can be carried out and optimized in many ways.

To account for incomplete task specifications, agents must use their experience or inquire in order to find out who sits where and whether people prefer particular plates, cups, etc. Other missing information can be found in the world wide web such as which kinds of plates and cutlery are needed and where they should be placed.

Aspects to be considered for the optimization of the activity include the the position where the robot should stand to place the objects on the table, the decision if objects should be stacked or carried individually, etc. Also, the task requires the execution of macro actions such as approaching the cupboard *in order to* open it and get out the plates. Agents also have to consider subtle issues in action selection, such as whether or not to interrupt the carrying action in order to close the door immediately or later. Finally, actions require substantial dexterity. Putting objects on table where people are sit requires the robot to perform socially acceptable reaching actions.

Cooking is the activity of preparing food to eat. Unlike setting the table, cooking requires the selection, measurement and combination of ingredients in an ordered procedure. The performance of a cooking activity is measured in terms of the quality and taste of the meal and its timely provision.

Cooking involves the transformation of food, the control of physical and chemical processes, the concurrent execution of different activities, the use of tools and ingredients, and monitoring actions. Concurrent activities have to be timed such that all parts of the meal are done at the desired time. Cooking comprises many different methods, tools and combinations of ingredients — making it a difficult skill to acquire.

Prolonged House Keeping and Household Chores are specific activities related to or used in the running of a household. The activities include cooking, setting the table, washing the dishes, cleaning, and many other tasks.

Performing household chores illustrates well the flexibility and reliability of human everyday activity. An important reason is the flexible management of multiple, diverse jobs. Hu-

mans as well as intelligent agents usually have several tasks and activities to perform simultaneously. They clean the living room while the food sits on the oven. At any time, interrupts such as telephone calls, new errands and revisions of old ones might come up. In addition, they have regular daily and weekly errands, like cleaning the windows.

Since doing the household chores is a daily activity and done over and over again, agents are confronted with the same kinds of situations many times, they learn and use reliable and efficient routines for carrying out their daily jobs. Performing their daily activities usually does not require a lot of thought because routines are general. This generality is necessary since jobs come in many variations and require agents to cope with whatever situations they encounter. The regularity of everyday activity also requires agents to handle interruptions that occur while carrying out their routines. Routines, like cleaning the kitchen, are often interrupted if a more urgent task is to be performed, and continued later on. Therefore, agents have to remember where they put things in order to find them again later.

A main factor that contributes to the efficiency of everyday activity is the flexibility with which different routines can be interleaved. Agents often work on more than one job at a time. Or, they do jobs together — operating the washing machine in, and getting stored food from, the basement.

People can accomplish their household chores well — even in situations they have not encountered yet — for two reasons: they have general routines that work well in standard situations and they are, at the same time, able to recognize non-standard situations and adapt their routines to the specific situations they encounter. Making the appropriate adaptations often requires people to predict how their routines would work in non-standard situations and why they might fail.

B. Scenarios

For these activities the following scenarios are investigated.

B.1. Understanding Human Household Chore: Acquisition and Use of Activity Models

In the first demonstration scenario the task is to observe the activity in the kitchen in order to acquire abstract models of the activity. The models are then to be used to

- answer queries about the activity,
- monitor activity and assess its execution,
- generate visual summaries of a cooking activity including symbolic descriptions, and
- detect exceptional situations and the need for help.

In order to do so, the system is required to answer questions including the following ones: What are meaningful subactivities of “setting the table” and why? Which eating utensils and plates have to be used for spaghetti and how should they be arranged? Does John prefer a particular cup? Where do people prepare meals? How do adults set the table as opposed to kids? and why? Why does one put down the plates before bringing them to the table? (To close the door of the cupboard.) What happens after the table is set? Where are the forks kept? What is the fork used for?

The models needed for answering these questions are to be acquired (semi-)automatically from the world wide web, from observing people, and by asking for advice.



Fig.6: Sensing, reaching, and grasping activities using tiny inertial measurement units (left). Estimating the pose of people by accurately matching digital human models into image sequences (right).

B. 2. Self-modelling and -adapting Assistive Kitchens

Maps or environment models are resources that the Assistive Kitchen uses in order to accomplish its tasks more reliably and efficiently. Acquiring a model of an assistive kitchen is very different from environment mapping done by autonomous robots. For mapping the Assistive Kitchen the robot can make use of the sensors of the environment. There is also various semantic information associated with RFID tags in the environment and the sensors are services providing information about themselves and their use.

Here, we consider mapping to be the following task: *Given* (1) a sensor-equipped kitchen where appliances and other pieces of furniture might or might not be tagged with RFID tags that have information associated with them (such as their size) and (2) a stream of observations of activities in this kitchen acquired by the various sensors *acquire* a semantic, 3D object map of the kitchen.

In contrast to many other robot mapping tasks where the purpose of mapping is the support of navigation tasks, the kitchen map is to be a resource for understanding and carrying out household chores. To this end, the map needs to have a richer structure, explicitly reference task relevant objects such as appliances, know about the concept of containers and doors, such as cabinets and drawers.

The maps also have to contain sensors of the environment and their locations in order to semantically interpret the sensor data. For example, the mapper has to estimate the position and orientation of a laser range sensor using its own motions and their effects on the sensor data, in order to infer that the sensor can be used for determining its own position and how. Or, the robot has to locate an RFID tag reader, for example using the estimated position of its gripper and observing which of the sensors reports the RFID identifier of its gripper. Knowing that its gripper is inside a cabinet it can infer that the respective sensor can be used to sense what is inside this cabinet.

By observing the activity in the kitchen the system can also infer the function of objects. For example, that a particular cabinet is used to store the plates and glasses and that the cutlery is kept in a particular drawer. Finally, the geometry of the environment and recognized activities are used to finetune the sensors for particular activity recognition tasks.

B. 3. Action and Motion Primitives

People doing their household activities perform very complex movements smoothly and effortlessly, and improve such movements with repetition and experience. Such activities

like putting the dishes in the cupboard, or setting the table, are simple for people but present challenges for robots.

Consider picking up a cup from the table and placing it in the cupboard. Any person will turn and walk towards the cup, while simultaneously stretching the arm and opening the hand. Then the hand will grasp the cup firmly, and lift it. It will apply just the right amount of force, because through experience, the person has learned how much a cup should weigh. If the cup were heavier as expected, this would be detected and corrected immediately. Then, taking the cup close the body for a better stability, the person will walk to the cupboard, open the door with the other hand, and place the cup inside.

Several noteworthy things take place. First, the motions used to get the cup are similar to the ones used for reaching and picking up other objects. Such movements can be formed by combining one or more basic movements, called motion primitives. These primitives can be learned by observation and experimentation. Each one of this primitives can be parametrized to generate different movements. Second, it is known that noise and lag are present in the nervous system, which affect both sensing and motor control. But people manage despite this limitations to have elegant and precise control of our limbs. A robot with a traditional controller would have great difficulties carrying out simple movements in such conditions. Humans deal with this problem by building forward and reverse models for motor control [7]. Third, during a movement, any abnormal situation is quickly detected and a corrective action is taken.

In this demonstration scenario, we would like a robot to exhibit a similar capabilities. The robot will observe the activities of people, obtain motion primitives from them, and improve them through own experimentation. Models of activity are also learned for reaching and grasping movements, and include information regarding the effects of the control commands. This allows the robot make predictions and detect abnormal situations when the sensory data differs from the expected values.

B. 4. Planning and Learning Macro Actions

The next higher level of activities in the kitchen are macro actions. We consider macro actions to be the synchronized execution of a set of action primitives that taken together perform a frequent macro activity in the task domain. These activities are so frequent that the agents learn high-performance skills from experience for their reliable and flexible execution. Examples for such macro activities are opening a tetrapak and filling a cup with milk, opening a cabinet to take a glass, or buttering a bagel.

Here the challenge is to compose macro actions from coordinated action primitives such that the resulting behavior is skillful, flexible, reliable, and fluent without noticeable transiting between the subactions [8]. Another challenge is that such macro actions must be learned from very little experience — compared with other robot learning tasks [3].

The learning of macro actions is an application of *action aware control* where the agents learn performance and predictive models of actions and use these models for planning, learning, and executing macro actions. The construction of such macro actions requires the application of transforma-

tional learning and planning methods and the combination of symbolic and motion planning with learned dynamic models.

B. 5. Self-adapting High-level Controller

Robotic agents can not be fully programmed for every application. Thus, in this demonstration scenario we realize robot control programs that specialize to their respective robot platform, work space, and tasks (see Figure 7).

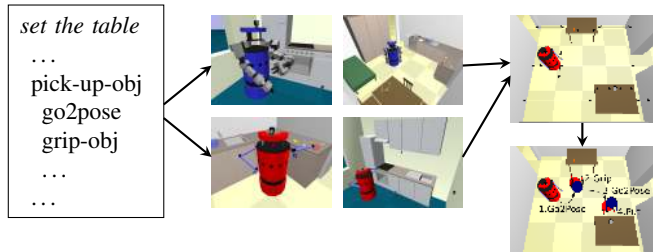


Fig.7: Self-adaptation of different robots in different kitchens.

Specifically we realize a high-level control program for setting the table. The program learns from experience where to stand when taking a glass out of the cupboard, how to best grasp particular kitchen utensils, where to look for particular cutlery, etc. This requires the control system to know the parameters of control routines and to have models of how the parameters change the behavior [2]. Also, the robots are required to perform their tasks over extended periods of time, which asks for very robust control.

B. 6. Learning to Carry out Abstract Instructions

The final scenario discussed in this paper is the acquisition of new high-level skills. Let us consider the setting of a table as an illustrative example (see Figure 8). Upon receiving “set the table” the robot retrieves instructions from webpages such as *ehow.com*. These instructions are typically sequences of steps to be executed in order to carry out the activities successfully. The challenges of this execution scenario are: (1) translate the abstract instructions into an executable robot control program, (2) supplement missing information through observations of kitchen activities, (3) transform the action sequence into an activity structure that can be carried out more reliably, efficiently, and flexibly. Instructions typically abstract away from these aspects of activity specification.

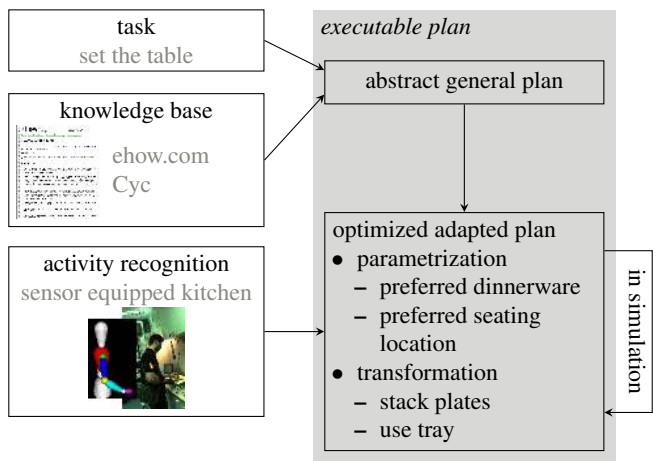


Fig.8: Learning to set the table.

Let us now look at some of the hard challenges in this scenario. First, translating abstract instructions into a working robot control program requires answers to the following research questions. (1) How can the plan libraries of autonomous household robots be specified so generally, reliably, transparently, and modularly that a robot can compose (almost) working plans from abstract instructions? In order for newly composed sequences of plan steps to work it helps if the individual plan steps are specified as “universal plans”, that is they achieve – if necessary – all preconditions needed for producing the desired effects. (2) Debugging newly created plans from instructions requires the robot to predict what will happen if it executes the new plan, to identify the flaws of the plan with respect to its desired behavior, and to revise the plan in order to avoid the predicted flaws. (3) Optimizing tasks like table setting also requires the technical cognitive system to observe people setting the table, to infer the structure of the activity and reason about why people do not follow the abstract instructions like a robot but perform the task the way they do. This way the robot would learn that people stack plates when carrying them in order to minimize the distance they have to walk. The robot would then transform its plan analogously and test whether this change of activity structure would result in improved performance.

IV. Concluding Remarks

This paper has presented assistive kitchens as demonstration platforms for cognitive technical systems that include various research challenges for cognitive systems. In particular, we expect the investigation of cognitive capabilities in the context of human everyday activity, which has received surprisingly little attention in previous research efforts, to substantially promote the state-of-the-art of cognition for technical systems. We refer to a longer version of this paper — currently in preparation — for more detailed discussion of the research issues and the related research in this area.

REFERENCES

- [1] M. Beetz. *Plan-based Control of Robotic Agents*, volume LNAI 2554 of *Lecture Notes in Artificial Intelligence*. Springer Publishers, 2002.
- [2] M. Beetz. Plan representation for robotic agents. In *Proceedings of the Sixth International Conference on AI Planning and Scheduling*, pages 223–232, Menlo Park, CA, 2002. AAAI Press.
- [3] M. Beetz, A. Kirsch, and A. Müller. RPL-LEARN: Extending an autonomous robot control language to perform experience-based learning. In *3rd International Joint Conference on Autonomous Agents & Multi Agent Systems (AAMAS)*, 2004.
- [4] R. Brachman. Systems that know what they’re doing. *IEEE Intelligent Systems*, pages 67 – 71, November/December 2002.
- [5] M. Buss, M. Beetz, and D. Wollherr. Cotesys — cognition for technical systems. submitted to: 4th COE-HAM Workshop, Tokyo Denki University, Japan, 2007.
- [6] R. B. Rusu, A. Maldonado, M. Beetz, and B. Gerkey. Extending player/stage/gazebo towards cognitive robots acting in ubiquitous sensor-equipped environments. In *Accepted for the ICRA Workshop for Networked Robot Systems, 2007, Rome, Italy*, 2007.
- [7] S. Schaal and N. Schweighofer. Computational motor control in humans and robots. *Current Opinion in Neurobiology*, 15:675–682, 2005.
- [8] F. Stulp and M. Beetz. Optimized execution of action chains using learned performance models of abstract actions. In *Proceedings of the 19th Int. Joint Conference on Artificial Intelligence (IJCAI)*, 2005.