

Knowledge Processing for Autonomous Robot Control

Moritz Tenorth and Michael Beetz

Intelligent Autonomous Systems Group
Department of Informatics
Technische Universität München
Karlstr. 45, D-80333 München
{tenorth,beetz}@cs.tum.edu

Abstract

Successfully accomplishing everyday manipulation tasks requires robots to have substantial knowledge about the objects they interact with, the environment they operate in as well as about the properties and effects of the actions they perform. Often, this knowledge is implicitly contained in manually written control programs, which makes it hard for the robot to adapt to newly acquired information or to re-use knowledge in a different context. By explicitly representing this knowledge, control decisions can be formulated as inference tasks which can be sent as queries to a knowledge base. This allows the robot to take all information it has at query time into account to generate answers, leading to better flexibility, adaptability to changing situations, robustness, and the ability to re-use knowledge once acquired. In this paper, we report on our work towards a practical and grounded knowledge representation and inference system. The system is specifically designed to meet the challenges created by using knowledge processing techniques on autonomous robots, including specialized inference methods, grounding of symbolic knowledge in the robot's control structures, and the acquisition of the different kinds of knowledge a robot needs.

Introduction

Autonomous robots are becoming more and more skilled in object manipulation and are extending their capabilities towards complex manipulation tasks, including everyday activities like setting a table, tidying up or preparing meals. Such tasks are extremely knowledge-intensive: Competently taking the decisions involved in these activities requires a robot to have access to various kinds of knowledge. Encyclopedic knowledge is required as a general vocabulary for describing the types and properties of objects and actions. It needs to be combined with spatial knowledge about the environment for planning actions and locating the objects required for a task. The abstract knowledge about object classes needs to be linked to physical objects the robot detects and localizes in the environment. Models of the effects of actions and about processes help a robot to reason about the consequences of its actions and to detect and eliminate execution flaws.

These kinds of knowledge and the required inferences differ in many respects from what is commonly investigated in

knowledge representation research in artificial intelligence (AI). Existing knowledge bases like Cyc (Lenat 1995) or SUMO (Niles & Pease 2001) are very extensive and cover a wide range of concepts, but they lack much of the practical knowledge a robot needs. Cyc, for example, describes grasping as a sub-class of holding an object, but does not tell the robot where to stand for grasping an object, which grasp to use, or which force to apply. The reason is that these knowledge bases were created for understanding texts rather than executing actions. Knowledge bases for robots therefore have some very specific demands compared to general-purpose knowledge representation systems as commonly regarded in AI research:

1. Robot-specific knowledge: Provide the different kinds of knowledge a robot needs, like a detailed general ontology, task descriptions, information about object properties, and environment models.
2. Grounding and integration with the robot: To have meaning to the robot, the abstract symbols in the knowledge base need to be related to actions, percepts, and to the robot's internal data structures.
3. Integration of knowledge sources: Some kinds of knowledge may be available from existing knowledge bases, others can be imported from sources on the Internet, others have to be acquired by the robot through sensing, abstracting and reasoning. These initially separate areas of knowledge need to be integrated and described in a common *interlingua* to be jointly used for inference tasks.
4. Special-purpose inference methods: Some inferences are required for a robot, but less important in other knowledge-based applications, for example spatio-temporal reasoning or the projection of action effects. In other cases, special-purpose inference can exploit special properties of the problem domain in order to solve theoretically hard problems in rather short time, as required by real-time robotic applications.

Creating practical, extensive knowledge bases that meet these demands and equip robots with sufficient knowledge to autonomously perform complex manipulation tasks thus remains a challenging problem. In this paper, we give an overview of our work towards such a knowledge processing system. Our focus is on explaining how the different components, the different kinds of knowledge and the inference

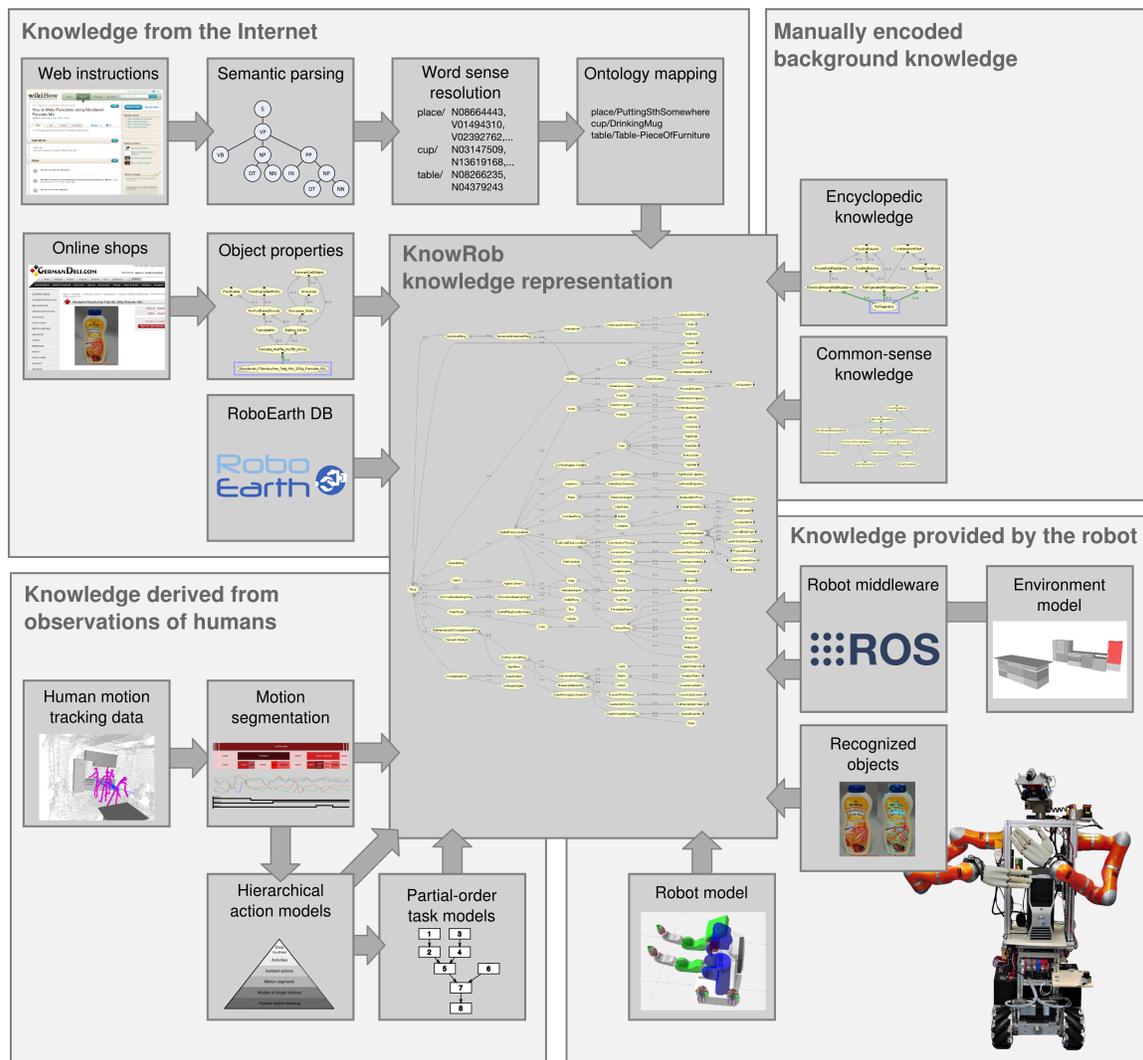


Figure 1: Overview of the different sources of knowledge integrated in the KNOWROB system.

techniques contribute to the overall system. Where possible, we refer to our prior work for more detailed descriptions of the individual components.

System architecture

KNOWROB, the knowledge processing system described in this paper, has been originally proposed in Tenorth & Beetz (2009) and has since been extended with several modules. The program code and ontology models are publicly available as open-source software as part of the ROS robot middle-ware.¹ Figure 1 describes the different kinds of knowledge and knowledge processing methods that are integrated in the system. Expressive representations for time, space, objects, environment models, actions, processes, the robot’s hardware and capabilities, as well as observations of human actions form the core of the system. Special inference methods for robotics-related applications operate on these

representations and provide for instance spatio-temporal inference, projection and planning capabilities. Several knowledge acquisition methods allow the robot to (semi-) autonomously acquire knowledge from sources on the Internet, using its own perceptual capabilities, or by observing human activities. Interfaces to the robot’s perception system serve for grounding abstract object information and for reasoning about physical objects in the environment. A tell/ask interface provides reasoning services to the robot control program in order to infer control decisions based on the content of the robot’s knowledge base.

In order to combine the different kinds of knowledge from different sources that will be described in the following sections, the robot needs to describe them in some common language that encodes their meaning and allows automated inference. KNOWROB uses the Web Ontology Language (OWL) for storing knowledge, and a Prolog-based representation and reasoning system for integrating the different inference modules.

¹<http://ros.org/wiki/knowrob>

Types of robot knowledge

In this section, we will describe in more detail which kinds of knowledge are represented in the KNOWROB system and where they can be acquired from.

Encyclopedic and common-sense knowledge

Encyclopedic knowledge is the kind of knowledge commonly found in a dictionary: definitions of types of actions and objects, such that cups are a kind of container, have a handle, and are used for drinking liquids. Such knowledge, also called an “ontology”, can, to a certain degree, be obtained from existing knowledge bases like OpenCyc. However, it usually has to be complemented with more detailed descriptions, such as the properties of specific types of objects. Tenorth *et al.* (2011) showed that for manufactured products, this kind of information can be mined from sources on the Web like shopping websites, which provide both a categorization of the products and detailed descriptions of their properties.

The encyclopedic knowledge is complemented by common-sense knowledge, additional information about the concepts that most humans immediately associate with them, like the fact that cups can break or that coffee may be spilled if a full cup is moved too quickly. Since these facts appear to be so obvious, they are normally not written down explicitly and are thus hard to find. There are, however, some initiatives such as the Open Mind Indoor Common Sense (OMICS) project that try to collect this knowledge from voluntary Internet users. Kunze, Tenorth, & Beetz (2010) describe techniques to convert this information from natural language into the OWL representation used in KNOWROB.

Objects and environment models

Spatial knowledge about the environment and the types and properties of the objects found therein is essential for a robot performing mobile manipulation tasks. Robots need this knowledge to navigate, to locate objects, and to read information about their properties. Our robots are equipped with a semantic environment map that is represented in the knowledge base in terms of instances of the respective objects, e.g. pieces of furniture, at the locations where they have been recognized (Figure 2). Tenorth *et al.* (2010) describe the representation and different use cases for such a semantic environment model in more detail.

Actions and processes

The knowledge base further contains an extensive ontology of actions and processes. While actions are actively and intentionally performed by a robot or a human, processes take place more or less automatically whenever their pre-conditions are fulfilled. Often, they are caused as side-effects of actions, like a baking process that transforms cookie dough into a cookie after the robot has put the dough into the oven. KNOWROB contains a unified representation for both actions and processes that supports representation, projection, and planning.

Actions can be described using a set of properties to specify their inputs and pre-conditions, their effects, as well as

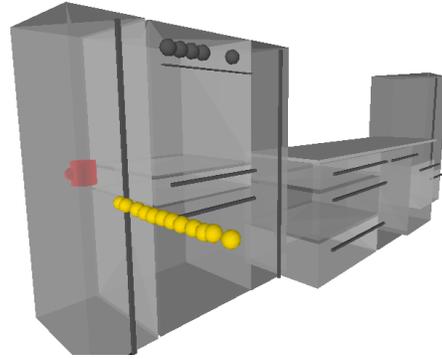


Figure 2: Semantic environment model describing the poses of pieces of furniture, the objects they contain, and their articulation properties.

their hierarchical composition. The representation is similar to Hierarchical Task Networks (HTN) used in robot planning. The representation in KNOWROB combines declarative specifications of the inputs and results of actions, used for planning purposes, with projection rules to qualitatively compute the effects of actions on the current world state. The current representation can be expanded to a certain level of granularity, below which actions are regarded as “black boxes”. We are currently investigating methods to overcome this limitation, for example using constraint-based motion specifications, which provide a more explicit and more transparent interface to the action execution level.

Self-models of the robot’s capabilities

In order to reason about what it can and cannot do, or which capabilities are missing for performing an action, a robot needs a detailed model of its own components. Kunze, Roehm, & Beetz (2011) present the Semantic Robot Description Language, a formal description of the robot’s kinematics, of the semantics of the different body parts, its hardware and software components, and of higher-level capabilities such as navigation or object recognition. Actions can define dependencies on capabilities and components that a robot needs to have for this action to be executable. These dependencies, as well as inter-dependencies among capabilities and components, can be described and verified. If something is found to be missing, the robot can check whether this capability can be acquired in some way.

Specialized inference methods

Practical knowledge bases that are used by robots during task execution need to provide answers fast enough to not slow down the robot’s operation. The problem is that, in the general case, most of the required inferences are prohibitively complex. However, they often become much easier if constraints of the domain the robot operates in are taken into account. In these cases, specialized inference techniques can help to compute results fast enough for the robot to use.

Computable predicates

The *computable* predicates in KNOWROB are realized by attaching computational procedures to semantic relations in the knowledge base that describe how these relations can be evaluated based on the robot’s belief state. This concept can be used for different purposes: First, it helps to speed up the evaluation of certain relations by using fast computation instead of complex inference. Second, computable predicates can be used for calculating relations beyond pure logical inference, for example to derive qualitative spatial relations like ”on top of“ or ”inside of“ from metric information about object poses. And third, they allow to generate abstract symbolic views on continuous data structures on demand during the inference process. This custom abstraction helps to ensure consistence and currentness of data and keeps the original, continuous data as reference, thereby avoiding the loss of information due to premature abstraction into symbolic representations. An example of a computable definition can be found in the middle block in the right part of Figure 3. The definition defines that the relation *on-Physical* can be computed using *holds(on_physical(T,B), Time)*.

Spatio-temporal reasoning about change

Robots act in dynamic environments and need to be able to describe the world state at different points in time. KNOWROB’s object representation (Figure 3) supports spatio-temporal reasoning about the changing locations of objects. Object poses are described based on the event that created the pose information (see Figure 3 left). This approach allows to store multiple perceptions over time, enabling the robot to memorize past world states. It can also be used to store non-perceptual information about object poses, e.g. predictions where objects can be found or plans where objects shall be put, without rendering the knowledge base inconsistent. Based on this representation, qualitative relations between objects, e.g. a relation $rel(A,B)$, can be computed for an arbitrary point in time T using the *holds(rel(A,B), T)* predicate. It first reads the pose of the latest perception of the objects before time T and then computes the relation based on this data (Figure 3 right). Computable predicates can be used to provide a simplified query mechanism using the current time as default.

Integration with the robot

The integration of the knowledge base with other components of the robot control system is a very important topic that is much more than just a matter of system integration. It rather involves several hard research challenges, for instance how to integrate the abstract, symbolic knowledge with the robot’s perception and action execution system, how to ground symbols in perception, and how to infer which control decisions.

Interface to the perception system

The integration of the robot’s knowledge base with its perception system allows the robot to reason about objects it has perceived. Whenever the robot detects and recognizes objects, they are added to the world representation in the

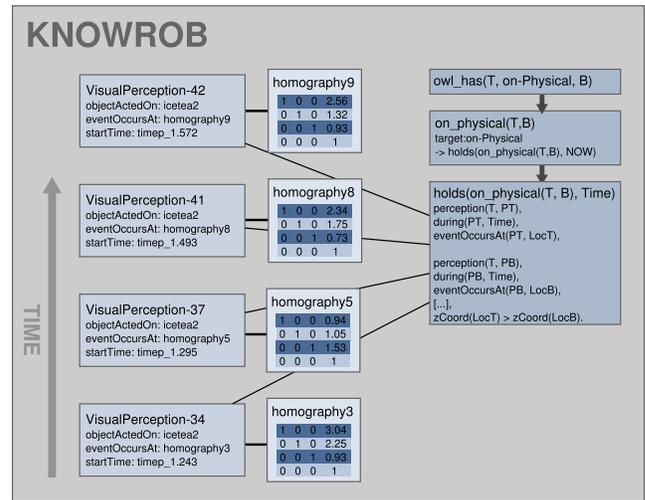


Figure 3: Spatio-temporal representation of object poses and computation of qualitative spatial relations based on this information.

knowledge base as described by Pangercic *et al.* (2010). The perception interface thereby builds up the representation depicted in the left part of Figure 3.

To maintain a consistent belief state and to correctly resolve the identities of the perceived objects over time, the raw object detections from the perception system can be filtered, for instance using the approach presented by Blodow *et al.* (2010).

Inferring control decisions

In order to use the robot’s knowledge during task execution, the control decisions that need to be taken must be formulated in terms of inference tasks that can be solved based on the robot’s knowledge and its belief about the world. Beetz, Mösenlechner, & Tenorth (2010) introduced the Cognitive Robot Abstract Machine (CRAM), a toolkit for programming cognition-enabled robot control programs. The KNOWROB knowledge base is a central component of this framework and closely interacts with the CRAM Plan Language (CPL). Task specifications in CPL plans contain abstract descriptions of object or locations, called *designators*, which can be resolved during run-time by sending an inference task to the knowledge base. To execute a task like “open the container where you think cups are stored”, the robot has to reason about likely storage locations of cups in the environment, needs to locate the respective container, and also has to find out how it can be opened. This inference task can be formulated as follows; its result is visualized in Figure 2.

```
?-rdf_triple(knowrob:'in-ContGeneric',knowrob:'Cup67',B),
rdf_has(B,knowrob:openingTrajectory,Traj),
findall(P,rdf_has(Traj,knowrob:pointOnTrajectory,P),
Points).
```

Knowledge acquisition and exchange

When leaving the world of controlled, limited lab experiments, a robot needs much broader knowledge about all the

different kinds of objects it encounters. The efficient acquisition of this knowledge then becomes a challenging problem which we try to approach by exploiting existing sources of knowledge as much as possible. We work on using information from the Internet, originally created for humans, on analyzing observations of human manipulation activities that could serve as an example how to perform a task, and on methods for sharing information among robots. These methods are largely complementary: While web sites mainly provide abstract, symbolic knowledge, observations of humans give information about motions and locations.

Knowledge acquisition from the Web

The World Wide Web is a valuable source of knowledge that can be exploited to bootstrap robot knowledge bases: Several web sites like ehow.com and wikihow.com provide thousands of step-by-step instructions how to perform everyday tasks, other websites provides recipes for cooking meals. We have developed methods to translate such natural-language instructions into a logical representation in the robot's knowledge base (Tenorth, Nyga, & Beetz 2010) and finally into executable robot plans.

Information about the properties and appearance of products can be mined from shopping websites, where the product pages list object properties, while the website's category can be transformed into an ontology of products to be added to the knowledge base (Tenorth *et al.* 2011). Since most products are listed together with their pictures, they can not only be abstractly described, but can also be recognized in the environment (Pangercic, Haltakov, & Beetz 2011).

Learning from observations of human activities

Observations of humans can provide the robot with information that is hard to obtain from other sources, like the motions to perform a task. The challenge is how make this source of information accessible to the robot for interpretation and reasoning, that is, how to assign semantic meaning to the observed continuous motions. In a first step, we segment the observations and represent these segmented observations in the knowledge base in terms of action instances (Tenorth, Bandouch, & Beetz 2009). For the robot to use the observations, the segments are classified and described using the same language that is also used in the rest of the knowledge representation system for modeling actions, objects, and spatial information. Action parameters are determined based on co-temporal events like RFID tag detections that allow to determine properties like the *objectActedOn*. Starting from the fine-grained initial segmentation, we can apply knowledge about the hierarchical composition of actions to generate coarser-grained action descriptions, for example to go from the level of single reaching motions to the level of transport actions (Beetz *et al.* 2010).

Exchange of knowledge among robots

The acquisition of knowledge, i.e. the translation from any kind of input format into a formal representation in the robot's knowledge base, is often a complex and time-consuming procedure that is difficult to completely automatize. Ideally, it should only be needed once for each piece

of information: If robots could exchange information about tasks they have learned, object models they have created, or environments they have explored, it would save other robots from having to acquire this knowledge by themselves. Such an exchange system could thus significantly speed-up knowledge acquisition using a distributed approach.

Creating such a system, a kind of "Wikipedia for robots", is the goal of the RoboEarth project (Waibel *et al.* 2011). KNOWROB is a central component of this project, providing the representations for describing the knowledge to be exchanged as well as the inference procedures needed to autonomously exchange information: When exporting knowledge, it has to decide which information could at all be useful to others and how this information needs to be processed to be exchangeable (e.g. be transformed into a different coordinate system). When downloading information, it has to select which pieces of information could be useful in the current task context, if the robot has all required capabilities to make use of them, and if they have further dependencies that need to be resolved.

Related work

Recently, there have been several other attempts to re-integrate knowledge processing techniques into robotic systems. The focus of the ORO ontology (Lemaignan *et al.* 2010) is on human-robot interaction and on resolving ambiguities in dialog situations. This capability was for example described by Ros *et al.* (2010), where the robot inferred based on its knowledge about the objects in the environment and their properties which queries it should ask to disambiguate a command. ORO uses OWL as representation format and a standard DL reasoner for inference. An underlying 3D geometrical environment representation serves for computing spatial information and for updating the internal belief state about the positions of objects (Siméon, Laumond, & Lamiroux 2001).

The knowledge base presented by Daoutis, Coradeshi, & Loutfi (2009) is an important part of the PEIS ecology project (Physically Embedded Intelligent Systems). PEIS investigates distributed intelligent systems consisting of mobile robots, but also of sensors embedded into the environment which are all integrated into a common framework. The PEIS knowledge base is realized as an extension of the Cyc inference engine. On the one hand, this gives the system full access to the large Cyc ontology, but it comes at the cost of slower inference, of irrelevant knowledge in several branches of the ontology, and of a lack of knowledge in areas like robotics or mobile manipulation.

The OUR-K system by Lim, Suh, & Suh (2011) is the successor of the OMRKF framework (Suh *et al.* 2007). OUR-K is an extensive system that describes a variety of aspects centered around five main kinds of knowledge: contexts, objects, spaces, actions and features. Compared to the KNOWROB ontology, OUR-K is lacking the notion of processes, robot self-models, and having simpler action descriptions.

Conclusions

In this paper, we gave an overview of KNOWROB, a knowledge processing system for autonomous robots. KNOWROB integrates various kinds of knowledge, like encyclopedic knowledge, spatial information and common-sense knowledge, from multiple sources in a common representation and reasoning framework. It supports robot-specific reasoning tasks such as spatio-temporal reasoning about changing object configurations. We further pointed to methods we developed for acquiring knowledge from the Internet, from observations of human activities, and from the robot's own sensory system.

Though KNOWROB is a rather extensive and implemented system, there are still open challenges to be overcome: We need a better integration of non-symbolic information, e.g. to reason about motions, forces or geometric properties. Actions should be described in more detail, e.g. including the expected outcome or potential problems, and should be linked to learning techniques. In the European project ROBOHOW.COG², we are investigating how to extend the methods presented in this paper towards a complete system that can autonomously learn novel tasks by autonomously combining information from the Internet with visual and tactile information from human demonstrations.

We believe that equipping robots with sufficient knowledge and effective reasoning capabilities is key to realizing flexible and robust robot behavior and to scaling autonomous robots towards more advanced everyday manipulation tasks.

Acknowledgments

This work is supported in part within the DFG excellence initiative research cluster *CoTeSys* (www.cotesys.org) and the EU FP7 Projects *RoboEarth* (grant number 248942) and *ROBOHOW.COG* (grant number 288533).

References

- Beetz, M.; Tenorth, M.; Jain, D.; and Bandouch, J. 2010. Towards Automated Models of Activities of Daily Life. *Technology and Disability* 22(1-2):27–40.
- Beetz, M.; Mösenlechner, L.; and Tenorth, M. 2010. CRAM – A Cognitive Robot Abstract Machine for Everyday Manipulation in Human Environments. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*.
- Blodow, N.; Jain, D.; Marton, Z.-C.; and Beetz, M. 2010. Perception and Probabilistic Anchoring for Dynamic World State Logging. In *Proceedings of 2010 IEEE-RAS International Conference on Humanoid Robots*.
- Daoutis, M.; Coradeschi, S.; and Loutfi, A. 2009. Grounding commonsense knowledge in intelligent systems. *J. Ambient Intell. Smart Environ.* 1(4):311–321.
- Kunze, L.; Roehm, T.; and Beetz, M. 2011. Towards semantic robot description languages. In *IEEE International Conference on Robotics and Automation (ICRA)*.
- Kunze, L.; Tenorth, M.; and Beetz, M. 2010. Putting People's Common Sense into Knowledge Bases of Household Robots. In *33rd Annual German Conference on Artificial Intelligence (KI 2010)*, 151–159. Karlsruhe, Germany: Springer.
- Lemaignan, S.; R., R.; L., M.; R., A.; and Beetz, M. 2010. Oro, a knowledge management module for cognitive architectures in robotics. In *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 3548–3553.
- Lenat, D. 1995. CYC: A large-scale investment in knowledge infrastructure. *Communications of the ACM* 38(11):33–38.
- Lim, G. H.; Suh, I. H.; and Suh, H. 2011. Ontology-based unified robot knowledge for service robots in indoor environments. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans* 41(3):492–509.
- Niles, I., and Pease, A. 2001. Towards a standard upper ontology. In *Proceedings of the international conference on Formal Ontology in Information Systems-Volume 2001*, 2–9. ACM.
- Pangercic, D.; Tenorth, M.; Jain, D.; and Beetz, M. 2010. Combining perception and knowledge processing for everyday manipulation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*.
- Pangercic, D.; Haltakov, V.; and Beetz, M. 2011. Fast and robust object detection in household environments using vocabulary trees with sift descriptors. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Workshop on Active Semantic Perception and Object Search in the Real World*.
- Ros, R.; Lemaignan, S.; Sisbot, E.; Alami, R.; Steinwender, J.; Hamann, K.; and Warneken, F. 2010. Which one? grounding the referent based on efficient human-robot interaction. In *19th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 570–575.
- Siméon, T.; Laumond, J.-P.; and Lamiroux, F. 2001. Move3D: a generic platform for path planning. In *4th International Symposium on Assembly and Task Planning*.
- Suh, I. H.; Lim, G. H.; Hwang, W.; Suh, H.; Choi, J.-H.; and Park, Y.-T. 2007. Ontology-based Multi-Layered Robot Knowledge Framework (OMRKF) for Robot Intelligence. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 429–436.
- Tenorth, M., and Beetz, M. 2009. KnowRob — Knowledge Processing for Autonomous Personal Robots. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 4261–4266.
- Tenorth, M.; Bandouch, J.; and Beetz, M. 2009. The TUM Kitchen Data Set of Everyday Manipulation Activities for Motion Tracking and Action Recognition. In *IEEE Int. Workshop on Tracking Humans for the Evaluation of their Motion in Image Sequences (THEMIS). In conjunction with ICCV2009*.
- Tenorth, M.; Kunze, L.; Jain, D.; and Beetz, M. 2010. KNOWROB-MAP – Knowledge-Linked Semantic Object Maps. In *Proceedings of 2010 IEEE-RAS International Conference on Humanoid Robots*.
- Tenorth, M.; Klank, U.; Pangercic, D.; and Beetz, M. 2011. Web-enabled Robots – Robots that use the Web as an Information Resource. *Robotics & Automation Magazine* 18(2):58–68.
- Tenorth, M.; Nyga, D.; and Beetz, M. 2010. Understanding and Executing Instructions for Everyday Manipulation Tasks from the World Wide Web. In *IEEE International Conference on Robotics and Automation (ICRA)*, 1486–1491.
- Waibel, M.; Beetz, M.; D'Andrea, R.; Janssen, R.; Tenorth, M.; Civera, J.; Elfiring, J.; Gálvez-López, D.; Häussermann, K.; Montiel, J.; Perzylo, A.; Schießle, B.; Zweigle, O.; and van de Molengraft, R. 2011. RoboEarth - A World Wide Web for Robots. *Robotics & Automation Magazine* 18(2):69–82.

²<http://www.robohow.eu>