“Knowing From” – An Outlook on Ontology Enabled Knowledge Transfer for Robotic Systems[1]

Mohammed DIAB [a] Mihai POMERLAN [b] Daniel BESSLER [b] Stefano BORGO [c] Jan ROSELL [a]

[a] *Universitat Politècnica de Catalunya*
[b] *Bremen University, Bremen, Germany*
[c] *Institute of Cognition Science and Technology, Trento, Italy*

Abstract. Encoding practical knowledge about everyday activities has proven difficult, and is a limiting factor in the progress of autonomous robotics. Learning approaches, e.g. imitation learning from human data, have been used as a way to circumvent this difficulty. While such approaches are on the right track, they require comprehensive knowledge modelling about the data present in records of activity episodes, and about the skills one attempts to have the robot learn. We provide a list of competency questions such knowledge modelling should answer, summarize some recent developments in this direction, and finish with a few open problems.

Keywords. Machine learning, Knowledge transfer, Manipulation capability, Autonomous robot

1. Motivation: the Challenge of “Knowing How”

Despite research interest into autonomous robot systems, progress in this field is slow. It takes considerable effort to teach a robot tasks which appear simple to a human, e.g. cracking an egg, and therefore autonomous robotic helpers for households or care facilities remain only a distant possibility. It is difficult to make explicit the practical knowledge of how to perform everyday tasks – it may be easy to walk, but is definitely hard to program a robot to do it. Recent development efforts attempt to elicit such knowledge indirectly: collect data from humans performing an activity, and use it to learn a model for inferring parameters or heuristics for organizing an activity’s structure. Alternatively, learning may use data collected from other robots’ attempts to perform the task.

We think such approaches are broadly correct, but we insist on the necessity to develop introspectable, formalized knowledge representations to guide them. The problem with machine learning, as a growing body of research demonstrates, is its tendency to learn shortcuts from idiosyncracies of the training data rather than the function intended by the human developer [1], a fact exacerbated by the opacity of some machine learned models. For a cyber-physical system acting in the real world however, it is crucial that the system is amenable to introspection and verification – it cannot be trusted otherwise.

Some of the authors of this paper are actively working on the development of an infrastructure to gather, store, and query information about activity episodes[2]. An episode is recorded in our infrastructure as a Narratively-Enabled Episodic Memory (NEEM). We highlight the "narrative" part: a rich level of semantics on top of the records of raw signals is necessary if such records are to be useful as training data, and we outline in this paper that the semantics should clarify what problems and tasks the data are about.

We begin with a list of competency questions that, ideally, knowledge modelling for autonomous robots should be able to answer. In particular, being able to answer such questions helps to acquire knowledge via learning and to transfer it between agents.

1. *What are the appropriate behaviors for a task in this situation?* Most tasks can be executed in different ways but not all are equally acceptable by humans: wasting of resources, very fast/slow movements, breaking of proxemic rules, causing danger are all behavioral factors that an ideal robot should take into account. Appropriateness is often cultural and contextual.
2. *What kind of knowledge is necessary to achieve a task?* Planning how to execute a task generally requires different types of knowledge: possible and allowed actions, likely consequences, factual and typical location of items, procedural "flair" (e.g. controller parametrizations), expectations about other agents, etc.
3. *What does a particular system believe to be true?* This kind of questions are important to explain why a system makes certain decisions, why it implements them in a certain way, and why the resulting behavior is justified or not.
4. *What knowledge can be transferred between agents?* It is important to understand what items of knowledge may be transferable, the level of abstraction at which they are most useful, and how to adapt them to the capabilities of the robot. Declarative knowledge can often be copied as-is, but practical knowledge is usually agent-dependent. Yet, it is expected that there is a level of commonality among different agents that have the "same skills".
5. *How is experience/training data gathered?* This question aims to make explicit what data is relevant and what events are considered (un)likely. Given the sensitivity of machine learning models to biases in their training data, it is not just the availability of the data that is important; we need data about the data.

2. Machine Learning for Robots: a Very Brief Overview

There is a very vast literature about applying machine learning techniques to a variety of applications, including robotics. We will only summarize a few results in this section, with a focus on how they relate to the competency questions we have identified. We believe the research results we discuss are typical of the larger situation in the field.

2.1. What are the appropriate behaviors for a task in a given situation?

Ethics and AI is itself a hot topic for philosophical and vision papers, but we are not aware of any results tackling this with machine learning. To our knowledge, there are no

[2] Querying the activity episodes we have stored is possible at `open-ease.org`

datasets or benchmarks for what would be moral or culturally appropriate behavior – i.e., the usual way of pumping out results from machine learning is not applicable –, and this is clearly a situation that needs principled modelling of what norms on behavior are, how they can be acquired from humans, and how they might be verified in a learned system.

2.2. What kind of knowledge is necessary to achieve a task?

In an attempt to provide more generalizable action concepts, carefully engineered curricula were used to help an agent learn how to act in a simulated pixelworld [2] or a tabletop setting with objects of simple shapes [3]. "Simpler actions" are used to build complex ones. While a promising approach, the curricula are too specific to the demonstrations in the papers. E.g., what the authors call "containment" in the pixelworld will not give an agent a general intuition about containment. Also, capability reasoning is limited in this approach. The robot cannot identify subactions it would need to master first, the way a human can see they need a way to control objects if the task is to position them.

2.3. What does a particular system believe to be true?

A recent report surveys opportunities for knowledge graph techniques for explainable AI [4]. Briefly, current XAI techniques focus too much on feature correlations and offer little support for reasoning, e.g. causal reasoning. To this we would add that, in so far as an XAI technique relies on approximating a model making decisions with a simpler one, faithfulness issues appear. Therefore, it is important that an agent's knowledge about its own situation, at least at a level of abstraction that determines the actions it initiates, be represented in an interpretable way.

2.4. What knowledge can be transferred between agents?

The correspondence problem – e.g. different body shapes – is well known in imitation learning and approaches to adjust trajectories from a human teacher to a robot student exist [5]. However, the robot is dependent on the engineer to provide such trajectory adjustments, which may change when a new task needs to be learned. A deeper understanding of skills, their effects and physical requirements, is needed to autonomously reason about heuristics to tackle the correspondence problem.

2.5. How is experience/training data gathered?

Unfortunately many papers tend to not put much weight on this issue and its implications for what a robot learns. E.g. a claim is made that, via imitation learning, a robot is taught to kick a stationary ball forward [6]. The result is impressive given the difference between human and robot, but a skill hasn't been learned yet: the training data contains only humans who start facing the ball, and who successfully kick it. Additional effort is necessary to learn how to adapt what is imitated to new circumstances [5], but this new training process is liable to introduce error via unexamined biases in the training data [1].

3. Knowledge Transfer Toward Robots

Knowledge representation and reasoning in autonomous robot control is a fairly extensive field of research with developments in both service and industrial robotics. Olivares-Alarcos et al. provide a comprehensive comparison of different approaches [7]. We will further focus in this section on what has been already done to support learning and knowledge transfer towards robots.

An immediate problem is that knowledge is often implicitly encoded, e.g. in the control system of the robot or in physical models for simulation, raising the question of how to organize such heterogeneous information sources in a coherent knowledge base [8]. KnowRob [9,10] and Perception and Manipulation Knowledge (PMK) [11] are two examples of systems attempting to do so, by presenting themselves to their users via a logic-based query interface where logical expressions are grounded in subsymbolic procedures, e.g. robot localization algorithms giving probability distributions for location.

The need for a distributed knowledge service where experiential data for robots could be pooled and queried on demand has become clear in recent years. Fundamental work has been done in the RoboEarth project [12] where Rapyuta, a cloud-robotics platform, was developed [13], allowing the delegation of computationally heavy tasks into the cloud platform. It also provided the robots with access to a knowledge repository including task knowledge. Another example of a cloud-robotics platform is openEASE which attempts to provide a cloud storage for experiential knowledge [14]. openEASE stores activity episodes as heterogeneous data sets including symbolic descriptions of activities and quantitative data recorded during the activity, allowing robots to extract knowledge from stored situations that are, by some relevant metric, similar to the situations they face. The increased control over the training data allows the robot to formalize its own learning problem and curate the dataset so that it is appropriate to the task.

In order for a repository of episodic memories to be useful as a generator of training data or even a “cheat sheet” to suggest ready-made plans and parametrizations to the robot, it should be able to answer several questions such as, is there any episode from a situation similar to the one a robot wants to learn about? If so, how many episodes are matched? Can the performance in the recorded episode be adapted to the current situation? Some of these questions are addressed in the works presented in [15,16].

To judge similarity of episodes, information present in a recorded episode should also allow answering questions such as what are the entities in the environment, what are their initial states, what are the sensors the robot/environment has, what type of sensors can be used to perceive the environment in the current situation [11].

As we have previously mentioned in our discussion of learning approaches for robots, it should be possible to store episodes of failure as well, both to understand what kind of failures can happen, when they do, and how to recover from them. This requires an ontological characterization of failure, its causal mechanisms, its impact on other activities the robot might perform, and what the criteria for successful failure handling are. These questions are addressed in the work presented in [17]

Stored episodes must also contain information relevant for the task and motion planning (TAMP) modules of a robot, such as: 1) the appropriate interaction parameters (friction, slip, maximum force) required by physics-based motion planners to correctly interact with rigid bodies; 2) the spatial relations (on, inside, right, etc.) between the objects in a cluttered scenario; 3) the feasibility of actions due to geometric issues like arm reachability or collisions; 4) the feasibility of actions due to object features (e.g., overweight objects); 5) the geometric constrains that limit the motion planner space; 6) action constraints regarding the interaction with objects; and 7) the initial scene for the planner regarding for instance the potential grasp poses. Knowledge can play a significant role to guide the TAMP by answering all aforementioned points [18,11].

4. Further Open Problems for a Formal Modelling of Knowledge Transfer

In this paper we have discussed knowledge transfer as a means to acquire and adapt suitable behaviors to execute tasks, and summarized some existing work into achieving this with knowledge-enabled approaches. If we look at this issue in full generality, the discussion in Sec. 3 leaves out important aspects which are hard to model, and that we anticipated in part in the list of competency questions of Sec. 1. In complex socio-technical systems, e.g. hospitals or retirement homes, it is not enough to learn how to accomplish a task in a standard way. When to act, how to adapt to surrounding events, how to interact with nearby humans are important aspects to the success of an activity, but these aspects are culturally dependent and rely on a shared understanding of the situation.

Some formal proposals have pushed forward in terms of a trait-based theory of culture [19]. The idea, supported by studies in anthropology, is to model culture as a combination of four types of traits: knowledge traits about factual knowledge, behavioral traits about recognized ways to behave, rule traits about general principles and guidelines and interpretation traits, i.e. functions that take as input perception and previous situations (among other things) and output the most common way to make sense of this information in the given cultural group; e.g. being in a restaurant if people are distributed over tables eating food, or in an emergency state when a certain sound is heard or there is a person lying unresponsive on the street. These approaches are far from being implemented, and the actual development of suitable cultural knowledge remains an open problem.

Another issue illustrated by our competency questions is modelling capabilities. Intuitively speaking, a capability is something which, if an agent possesses it, allows that agent to perform some kind of tasks. Complexity comes from finding a principled, formal answer about what kind of entity a capability is, what other entities it might depend on and in what way, and how it might compare to analogous capabilities of different agents. A simple capability model might be obtained by mapping sets of the robot's hardware parts to tasks; an example of this in the multi-robot, outdoor setting can be seen in the SHERPA project [20]. However, different robots may use different hardware for performing similar tasks, and defective hardware may be worked around by a process of adaptation. When adding new tasks to the repertoire of a robot it is not clear, without a deeper modelling of capabilities and how they relate to each other, how such a simple, hardware-to-tasks mapping should be updated – in particular, not when the new tasks don't use the old ones, with the old parametrizations, as subroutines.

Acknowledgments

This work was partially funded by Deutsche Forschungsgemeinschaft (DFG) through the Collaborative Research Center 1320, EASE and by the Spanish Government through the project DPI2016-80077-R. M. Diab is supported by the Spanish Government through the grants FPI 2017.

References

[1] Geirhos R, Jacobsen JH, Michaelis C, Zemel R, Brendel W, Bethge M, et al.. Shortcut Learning in Deep Neural Networks; 2020.

[2] Hay N, Stark M, Schlegel A, Wendelken C, Park D, Purdy E, et al. Behavior Is Everything: Towards Representing Concepts with Sensorimotor Contingencies. In: McIlraith SA, Weinberger KQ, editors. Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018. AAAI Press; 2018. p. 1861–1870. Available from: `https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/16413`.

[3] Lázaro-Gredilla M, Lin D, Guntupalli JS, George D. Beyond imitation: Zero-shot task transfer on robots by learning concepts as cognitive programs. Science Robotics. 2019 01;4:eaav3150.

[4] Lecue F. On the role of knowledge graphs in explainable AI. Semantic Web. 2019 12;11:1–11.

[5] Hussein A, Gaber MM, Elyan E, Jayne C. Imitation Learning: A Survey of Learning Methods. ACM Comput Surv. 2017 Apr;50(2). Available from: `https://doi.org/10.1145/3054912`.

[6] Elbasiony R, Gomaa W. Humanoids Skill Learning Based on Real-Time Human Motion Imitation Using Kinect. Intell Serv Robot. 2018 Apr;11(2):149–169.

[7] Olivares-Alarcos A, Beßler D, Khamis A, Gonçalves P, Habib M, Bermejo J, et al. A Review and Comparison of Ontology-based Approaches to Robot Autonomy. The Knowledge Engineering Review. 2019 12;34.

[8] Bateman JA, Beetz M, Beßler D, Bozcuoglu AK, Pomarlan M. Heterogeneous Ontologies and Hybrid Reasoning for Service Robotics: The EASE Framework. In: Third Iberian Robotics Conference. ROBOT '17. Sevilla, Spain; 2017. .

[9] Beetz M, Beßler D, Haidu A, Pomarlan M, Bozcuoglu AK, Bartels G. KnowRob 2.0 – A 2nd Generation Knowledge Processing Framework for Cognition-enabled Robotic Agents. In: International Conference on Robotics and Automation (ICRA). Brisbane, Australia; 2018. .

[10] Tenorth M, Beetz M. KnowRob – A Knowledge Processing Infrastructure for Cognition-enabled Robots. Int Journal of Robotics Research. 2013 April;32(5):566 – 590.

[11] Diab M, Akbari A, Ud Din M, Rosell J. PMK—A Knowledge Processing Framework for Autonomous Robotics Perception and Manipulation. Sensors. 2019;19(5). Available from: `https://www.mdpi.com/1424-8220/19/5/1166`.

[12] Waibel M, Beetz M, Civera J, D'Andrea R, Elfring J, Gálvez-López D, et al. RoboEarth. IEEE Robotics and Automation Magazine. 2011;18(2):69–82.

[13] Mohanarajah G, Hunziker D, D'Andrea R, Waibel M. Rapyuta: A Cloud Robotics Platform. IEEE Transactions on Automation Science and Engineering. 2015;12(2):481–493.

[14] Beetz M, Tenorth M, Winkler J. Open-EASE – A Knowledge Processing Service for Robots and Robotics/AI Researchers. In: IEEE International Conference on Robotics and Automation (ICRA). Seattle, Washington, USA; 2015. Finalist for the Best Cognitive Robotics Paper Award.

[15] Bozcuoğlu AK, Kazhoyan G, Furuta Y, Stelter S, Beetz M, Okada K, et al. The Exchange of Knowledge Using Cloud Robotics. IEEE Robotics and Automation Letters. 2018;3(2):1072–1079.

[16] Diab M, Pomarlan M, Beßler D, Akbari A, Rosell J, Bateman J, et al. SKillMaN - A Skill-based Robotic Manipulation Framework based on Perception and Reasoning. Robotics and Autonomous Systems - Journal, accepted. 2020.

[17] Diab M, Pomarlan M, Beßler D, Akbari A, Rosell J, Bateman J, et al. An ontology for failure interpretation in automated planning and execution. In: Iberian Robotics conference. Springer; 2019. p. 381–390.

[18] Diab M, Akbari A, Rosell J, et al. An ontology framework for physics-based manipulation planning. In: Iberian Robotics conference. Springer; 2017. p. 452–464.

[19] Borgo S, Blanzieri E. Trait-based Module for Culturally-Competent Robots. International Journal of Humanoid Robotics. 2019;16(6):1950028.

[20] Yazdani F, Kazhoyan G, Bozcuoglu AK, Haidu A, Balint-Benczedi F, Beßler D, et al. Cognition-enabled Framework for Mixed Human-Robot Rescue Team. In: International Conference on Intelligent Robots and Systems (IROS). Madrid, Spain: IEEE; 2018. .