Motion Tasks Representation: Extracting Knowledge from Human Experts

Xanthi S. Papageorgiou
Robotics & Cognitive Systems Unit
UBITECH
Athens, Greece
xpapageorgiou@ubitech.eu

Danai Vergeti
Robotics & Cognitive Systems Unit
UBITECH
Athens, Greece
vergetid@ubitech.eu

Dimitris Ntalaperas
Robotics & Cognitive Systems Unit
UBITECH
Athens, Greece
dntalaperas@ubitech.eu

***Abstract*—Service Robots (SR) are increasingly used for execution of complex tasks based on high-level goals. Also, the interaction with users in an easily-understandable way is very important. To this end, the imitation of human - like motion, is important not only because we "don't need to reinvent the wheel", but because humans, that the SR will share the world with, expect such solutions to the task. It is crucial to represent and organize the vast amount of knowledge, so that the robot can retrieve relevant knowledge faster and conveniently, in order to complete the tasks automatically. In this work, we propose a Holistic Knowledge Base scheme for task planning representation, towards the acceleration of robotic learning, based on robotic priors, scene structure, and demonstrations in a specific real-world context which dynamically changes over time and space.**

***Index Terms*—Motion Tasks Representation, Knowledge Base, Learning from Demonstration**

I. Introduction

Service robots are increasingly pervasive in our daily life and are anticipated to execute complex tasks based on high-level goals and interact with users in an easily-understandable way, [1]. This requires robots to effectively organize and represent situational knowledge — the in-situ information about humans, objects, places, and events in the robot's working environment, [2]. However, since situational knowledge is often highly related to users and objects on the spot, robots need to interact with humans to mediate the mismatch between perception and comprehension of situational knowledge, [3].

How to represent and organize the vast amount of knowledge and the complex relationships between these this knowledge in a more effective way, so that the robot can retrieve relevant knowledge faster and conveniently, and reasoning based on the original knowledge to help the robot complete the task automatically, has been a new research problem.

Proper motions and tasks should be learned from demonstration of human experts, in order to incorporate their skills into the robotic systems. Skill learning from demonstration has been thoroughly addressed in previous approaches. Several imitation learning systems and architectures based on the perception and analysis of human demonstrations have been proposed, [4], [5]. In most of the proposed approaches, the imitation process proceeds through three stages:

- perception and analysis of the human demonstration,
- representation of the demonstration,
- reproduction of the demonstrated task on the robot.

Two trends emerge from known approaches in the literature regarding the way demonstrations are represented and generated: the use of trajectory-level representations in the form of non-linear mappings between sensory and motor information, [6], [7], and the use of symbolic-level representations that decompose demonstrations into sequences of abstract skills or task segments, [8], [9]. A key issue in all these approaches is to find a generic representation which

- expresses actions as combinations of meaningful elements based on motor primitives,
- learns such motor primitives,
- uses them to recognize and synthesize actions,
- transfers them to different tasks and different embodiments, [10], [11].

In this context, recent works suggest that geometric-based representations of motion primitives may be well suited to learn larger scopes of motion primitives by allowing

- the representation and learning of non-Euclidean parameters, which often arise in a variety of robotics tasks, [12], [13],
- the learning of latent manifold spaces from demonstrations, where relevant motion pattern are efficiently captured.

Moreover, current efforts focus on generalizing the representation power of primitives in the 6D space and non-Euclidean domain, [14], [15], however the disentanglement of the subspaces remains an open problem.

In this paper, we propose a holistic Knowledge Base for task planning representation scheme, which addresses robotic reasoning in a specific real-world context that dynamically changes over time and space.

This approach is a part of a larger scale research project, that investigates the applicability of differential geometry methods and trajectory representation in imitation learning. Specifically, our research goal is to describe the motion using categories from homotopy theory and differential geometry theory. This approach is natural, not only because it allows an easy categorization of motion primitives and tasks, but, even more importantly, it allows a "transfer" of task to other

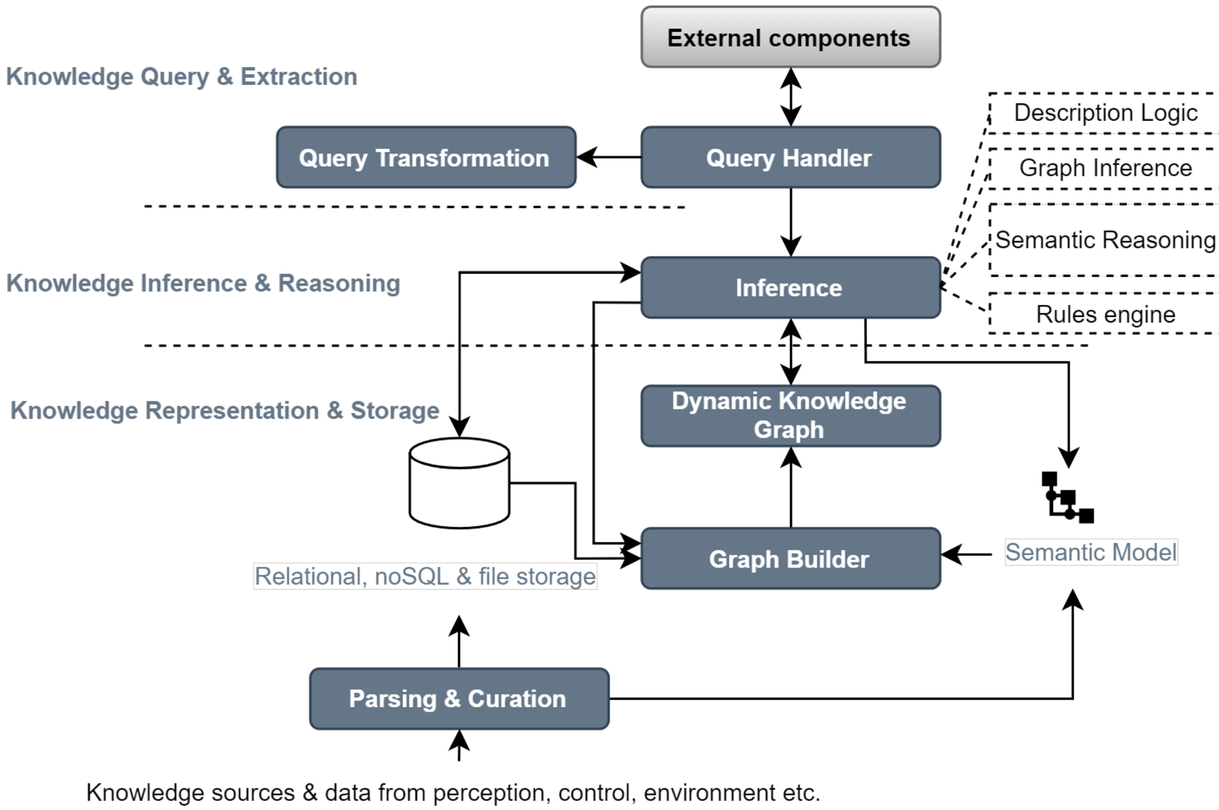

Fig. 1. Knowledge Base: High-level architectural layers.

robotic systems, using the machinery of differential geometry, thus facilitating a horizontal knowledge transfer. To this end, we plan to utilize the machinery developed and reported in this work.

II. Knowledge Representation Architecture

In this work, we propose a novel holistic knowledge representation architectural framework for developing solutions to the problem of incorporating robotic priors, scene structure, and human expert task demonstrations to accelerate robotic learning.

The proposed framework is to a great extent agnostic to the specifics of the incorporated knowledge. To accomplish this, a generic structure is used to store the knowledge, facilitating the use of different modalities, while a parsing subsystem transforms all the knowledge to this generic structure. A dynamic graph is used as the instrument of choice to represent the combined knowledge.

The scope of this work is not, obviously, to solve the core problem, but to describe a generic framework that can be used both as a test bed for different schemes, as well as to be deployed for actual use.

The framework is designed to be flexible, incorporating different prior architectures, different rule systems, and different modes of demonstration.

III. Holistic Knowledge Base for Task Planning Representation

The knowledge base needs to be integrated with sensing and acting, ground the symbolic knowledge in sensor data and support the ability of the robot to capture updated information from the environment [16], [17]. At the same time, meeting efficient data integration and interoperability, the knowledge base needs to act as an intermediate knowledge acquisition layer which will implement knowledge management and provide a set of knowledge extraction services. Efficiency and near-real-time (NRT), [18], response times of the knowledge extraction to other architectural components are two horizontal performance requirements that need to be also addressed. Under these high-level requirements, the knowledge base needs to provide a knowledge representation model and storage infrastructure, inference capabilities that go beyond semantic inference and efficient knowledge extraction under different perspectives (from cognition algorithms to perception and control-centric components).

The high-level architectural layers of the proposed knowledge base are depicted in Fig. 1.

IV. Knowledge Representation and Storage

The core knowledge of the proposed architecture should be able to provide semantic representations of the real-world and common sense enhanced with the domain knowledge of the application sector (as this is also captured and extended by the specific application cases) and to capture the dynamic changing context of the real environment where the robot acts and performs task planning. This includes some specific modules, as will be described in the following subsections.

A. Semantic Model

Knowledge representations for autonomous robots must be particularly rich in the way they represent actions, events, processes, situations, action preconditions and action effects, failures, knowledge of actions, as well as the robot self-knowledge. Based on these, the proposed semantic model will provide encyclopedic and domain knowledge under a common, formal, well-defined vocabulary for representing knowledge that can be used by the different components of the robot. This is complemented by common sense knowledge that provides additional information that is perceived as too obvious to be explicitly expressed and is associated with the concepts and is needed by the robots to perform tasks. Thus, the semantic model will include definitions of types of object parts, geometric representation to select actions, grasps, objects, as well as experience- and action-related knowledge to choose suitable parameters, robot parts etc. Our by investigating well established ontologies and models such as the Open Mind Indoor Common Sense (OMICS), the KnowRob ontology, Affordances Ontology, ORA Core Ontology, OpenCyc Ontology, RoboEarth system etc., while at the same time will integrate domain knowledge from the specific domain cases and the primitives.

B. Dynamic Knowledge Graph

As also indicated in [19], [20], the spatio-temporal aspects of the dynamic changing context of the real scene are also captured and represented as a Dynamic Knowledge Graph to be taken also under consideration during inference. The Dynamic Knowledge Graph is semantically enhanced and extended with entities and relations of the semantic model. The Dynamic Knowledge Graph will be continuously updated with information coming from the environment either the robotic perception, control system or the cognition algorithms/task planning, as well as experiences of the robot, and it will provide the primary point of decision-making of the robot to proceed with task execution and how.

C. Data Processing and Storage

The infrastructure, as well as data management services (collection, cleaning, encryption/decryption, anonymization, harmonization, storage) of the generated multidisciplinary data (relational data, noSQL data, images and files, videos, feeds, user feedback, machine learning parameters etc.) of the services (visual perception, cognition algorithms, task planning, control, training etc.). This data will either support the system service's proper functionality, the training of the modes, the knowledge base completion. Many of these data are expected to be derived from the appropriate data collection.

V. Knowledge Inference and Reasoning

The knowledge reasoning addresses five main reasoning strategies: spatial, temporal, context, retrieval modules, event-related reasoning module, and human-related reasoning. At the same time, it needs to serve different inference problems from the different (robotic) layers with different levels of expressivity, scalability, complexity, and efficiency. Thus, at the level of encyclopedic knowledge, our proposition system foresees semantic reasoning of the appropriate Ontology (using DL reasoners as well as semantic rules), while at the level of the Dynamic Knowledge Graph different ML and DL algorithms are used to perform graph inference and update and maintain the graph. Towards a more robust inference and reasoning approach, the proposed solution will also combine data coming from perception and control (not provided into the knowledge base) to perform forward-chaining and backward-chaining reasoning for task planning. Moreover, further approaches will be incorporated to resolve more higher-level problems such as rule engines (Drools). Through this three-level reasoning approach will perform optimized inference meeting the time-efficiency requirement of the cognition and motion generation models.

VI. Knowledge Query and Extraction

A set of services (API) for knowledge querying from the Knowledge Base, offering a higher level of abstraction. This layer acts as a proxy between the external components which either request a simple data query or complex inference hiding lower-level complexity. This layer routes and transforms the upper-level knowledge extraction requests to the relevant queries and inference mechanisms and orchestrates knowledge extraction depending on the type of query whether it is for perception, reasoning, or control.

VII. Conclusions

In this paper, we propose a Holistic Knowledge Base scheme for task planning representation, towards the acceleration of robotic learning, based on priors, scene structure, and demonstrations in a specific real-world context which dynamically changes over time and space. The proposed scheme contributes to the representation and organization of the vast amount of knowledge, so that the robot can retrieve relevant knowledge faster and conveniently, in order to complete the tasks automatically.

Acknowledgment

This work is a part of MARS: "Manufacturing Architecture for Resilience and Sustainability" project, that has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 101091783.

References

[1] A. Kattepur and B. P, "Roboplanner: Autonomous robotic action planning via knowledge graph queries," in *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*, ser. SAC '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 953–956. [Online]. Available: https://doi.org/10.1145/3297280.3297568

[2] B. Javia and P. Cimiano, "A knowledge-based architecture supporting declarative action representation for manipulation of everyday objects," in *Proceedings of the 3rd Workshop on Model-Driven Robot Software Engineering*, ser. MORSE '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 40–46. [Online]. Available: https://doi.org/10.1145/3022099.3022105

[3] R. Fang, M. Doering, and J. Y. Chai, "Embodied collaborative referring expression generation in situated human-robot interaction," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 271–278. [Online]. Available: https://doi.org/10.1145/2696454.2696467

[4] C. Frith, D. Wolpert, S. Schaal, A. Ijspeert, and A. Billard, "Computational approaches to motor learning by imitation," *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, vol. 358, no. 1431, pp. 537–547, 2003.

[5] Y. Zhou, J. Gao, and T. Asfour, "Movement primitive learning and generalization: Using mixture density networks," *IEEE Robotics & Automation Magazine*, vol. 27, pp. 22–32, 2020.

[6] A. J. Ijspeert, J. Nakanishi, and S. Schaal, "Movement imitation with nonlinear dynamical systems in humanoid robots," *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*, vol. 2, pp. 1398–1403 vol.2, 2002.

[7] S. P. Chatzis and Y. Demiris, "Echo state gaussian process," *IEEE Transactions on Neural Networks*, vol. 22, no. 9, pp. 1435–1445, 2011.

[8] K. Lee and Y. Demiris, "Towards incremental learning of task-dependent action sequences using probabilistic parsing," *2011 IEEE International Conference on Development and Learning (ICDL)*, vol. 2, pp. 1–6, 2011.

[9] M. Wächter, S. Schulz, T. Asfour, E. Aksoy, F. Wörgötter, and R. Dillmann, "Action sequence reproduction based on automatic segmentation and object-action complexes," in *2013 13th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2013, pp. 189–195.

[10] A. Paraschos, C. Daniel, J. R. Peters, and G. Neumann, "Probabilistic movement primitives," in *Advances in Neural Information Processing Systems*, C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger, Eds., vol. 26. Curran Associates, Inc., 2013.

[11] S. Tosatto, G. Chalvatzaki, and J. Peters, "Contextual latent-movements off-policy optimization for robotic manipulation skills," 05 2021, pp. 10 815–10 821.

[12] N. Jaquier, L. Rozo, D. G. Caldwell, and S. Calinon, "Geometry-aware manipulability learning, tracking, and transfer," *The International Journal of Robotics Research*, vol. 40, no. 2-3, pp. 624–650, 2021, pMID: 33994629. [Online]. Available: https://doi.org/10.1177/0278364920946815

[13] L. Rozo and V. Dave, "Orientation probabilistic movement primitives on riemannian manifolds," 10 2021.

[14] N. D. Ratliff, J. Issac, and D. Kappler, "Riemannian motion policies," *CoRR*, vol. abs/1801.02854, 2018. [Online]. Available: http://arxiv.org/abs/1801.02854

[15] H. B. Mohammadi, S. Hauberg, G. Arvanitidis, G. Neumann, and L. D. Rozo, "Learning riemannian manifolds for geodesic motion skills," *CoRR*, vol. abs/2106.04315, 2021. [Online]. Available: https://arxiv.org/abs/2106.04315

[16] M. Tenorth and M. Beetz, "Knowrob: A knowledge processing infrastructure for cognition-enabled robots," *The International Journal of Robotics Research*, vol. 32, no. 5, pp. 566–590, 2013. [Online]. Available: https://doi.org/10.1177/0278364913481635

[17] M. Wächter, E. Ovchinnikova, V. Wittenbeck, P. Kaiser, S. Szedmak, W. Mustafa, D. Kraft, N. Krüger, J. Piater, and T. Asfour, "Integrating multi-purpose natural language understanding, robot's memory, and symbolic planning for task execution in humanoid robots," *Robotics and Autonomous Systems*, vol. 99, pp. 148–165, 2018.

[18] [Online]. Available: https://www.precisely.com/blog/big-data/difference-between-real-time-near-real-time-batch-processing-big-data

[19] A. Saxena, A. Jain, O. Sener, A. Jami, D. K. Misra, and H. S. Koppula, "Robobrain: Large-scale knowledge engine for robots," *CoRR*, vol. abs/1412.0691, 2014. [Online]. Available: http://arxiv.org/abs/1412.0691

[20] C. Jiang, S. W. Lu, and M. Jägersand, "Constructing dynamic knowledge graph for visual semantic understanding and applications in autonomous robotics," *CoRR*, vol. abs/1909.07459, 2019. [Online]. Available: http://arxiv.org/abs/1909.07459