Mental Simulation for Autonomous Learning and Planning Based on Triplet Ontological Semantic Model

Yuri Goncalves Rocha and Tae-Yong Kuc

College of Information and Communication Engineering, Sungkyunkwan University, South Korea

{yurirocha, tykuc}@skku.edu

Abstract

Cognitive science findings showed that humans are able to create simulated mental environments based on their episodic memory and use such environment for prospecting, planning, and learning. Such capabilities could enhance current robotic systems, allowing them predict the output of a plan before actually performing the action on the real world. It also allow robots to use this simulated world to learn new tasks and improve its current ones using Reinforcement Learning approaches. In this work, we propose a semantic modeling framework, which is able to express intrinsic semantic knowledge in order to better represent robots, places and objects, while also being a memory-efficient alternative to classic mapping solutions. We show that such data can be used to automatically generate a complete mental simulation allowing robots to simulate themselves and other modeled agents into known environments. This simulations allows robots to perform autonomous learning and planning without the need of human-tailored models.

1 Introduction

Mental simulation is one of the fundamental cognitive skills. It allow humans (and perhaps other animals too) to predict and anticipate outcomes by recalling past experiences. This ability is one of the main pillars of episodic memory [Boyer, 2008] paramount for task planning during navigation [Burgess, 2008]. Mental simulation theory presupposes three main components: behaviors can be simulated, perception systems can also be simulated, and, finally, outcomes can be anticipated by combining simulated behaviors and perception [Hesslow, 2012]. Early researches on the field also showed that the same mechanism is used to predict counterpart's thoughts and behaviors [Gordon, 1986]. Besides being one of the core mechanisms of the human brain, mental simulation is yet to be fully explored on robotic systems. Some works on the field suggested that such capability should not only be integrated into learning and planning algorithms, but also be their central architectural layer [Polceanu and Buche, 2017].

In order to simulate a given environment, it is paramount to understand it semantically. The semantic information adds another layer to the robot knowledge, allowing for a better understanding of the intrinsic concepts and relations that are inferred naturally by humans. Even though there are several applications of semantic data in the robotics field [Waibel et al., 2011, Kostavelis et al., 2016, Cosgun and Christensen, 2018], just a small fraction of them use it to perform mental simulations [Tenorth and Beetz, 2009, Beetz et al., 2015, Beetz et al., 2018].

Learning is one of the main applications of mental simulation. Humans, at first, learn by interacting with the environment and observing its output. After obtaining enough experience, the brain is able to simulate this environment, and use it to imagine new outcomes by applying a different behavior. In robotics, one of the main fields of Artificial Intelligence(AI) is Reinforcement Learning(RL). RL is inspired by the human way of learning, and it works by

exploring the environment and giving (or removing) rewards, depending on how well the robot executed a given task. More specifically, Deep Reinforcement Learning(deep-RL), has been used on several autonomous navigation applications [Tai et al., 2017, Shah et al., 2018, Kahn et al., 2018].

The contributions of this work are as follows:

- Expanding an Ontologic Semantic Framework in order to automatically generate a full simulation environment, including a simulated robot.

- An end-to-end deep-RL model for autonomous navigation trained using the mentally simulated environment.

2   Related Work

In the past decades, several works proposed ways to incorporate knowledge into computers. CYC [Lenat, 1995] and SUMO [Niles and Pease, 2001] gathered a large amount of encyclopedic knowledge into its database, however such knowledge lacked the information necessary for mobile robot tasks. The OMICS [Gupta et al., 2004] project created a similar database containing the necessary knowledge in order to a robot complete several indoor tasks. The RoboEarth [Waibel et al., 2011] project tried to create a World Wide Web for robots, where they would be able to share and obtain knowledge in an autonomous way. KnowRob [Tenorth and Beetz, 2009, Beetz et al., 2018] and OpenEASE [Beetz et al., 2015] created a complete knowledge processing system capable of semantic reasoning and planning, and also performing mental simulations (referred as *Mind's Eye*). Most of those works, however, focused on manipulation tasks only.

Despite being thoroughly studied by cognitive science researches [Boyer, 2008, Burgess, 2008, Hesslow, 2012, Kahneman and Tversky, 1981], the mental simulation concept only started to be applied to computational systems few decades ago. Most of the early works focused on the "putting yourself on other's shoes" approach, where an agent would simulate itself on its counterpart perceived state in order to infer about its feelings and intentions. Leonardo [Gray and Breazeal, 2005] was developed to infer a human intention and aid the execution of this predicted task. In [Buchsbaumm et al., 2005], an animated mouse was able to imitate similar actors by inference using its own motor and action representations. [Laird, 2001] created a Quake bot able to predict its opponent next action by simulating itself on the opponent's current state, while [Kennedy et al., 2009] used its own behavior model to predict another agent's actions. Most of the recent works on robotics field, however, focused on the application of mental simulation to manipulation tasks planning and learning [Tenorth and Beetz, 2009, Beetz et al., 2015, Beetz et al., 2018, Kunze and Beetz, 2017], or comprehension and expression of emotions when socializing with humans [De Carolis et al., 2017, Horii et al., 2016]. J. Hamrick [Hamrick, 2019], however, showed that there are several similarities between mental simulation findings from cognitive science and model-based deep-RL approaches.

Deep Reinforcement Learning (deep-RL) has been applied to several different robot tasks, including but not limited to Human-Robot Interaction [Christen et al., 2019, Qureshi et al., 2018], dexterous manipulation [Gu et al., 2017, Rajeswaran et al., 2017] and autonomous map-less navigation [Kahn et al., 2018, Zhu et al., 2017]. RL methods can be divided into *model-based* and *model-free value-based* approaches. Model-based algorithms, such as [Zhu et al., 2017] use a predictive function that receive the current state and a sequence of actions and outputs the future states. The policy then select the sequence of actions that maximizes the expected rewards from the predicted states. Model-free approaches, such as [Christen et al., 2019], approximate a function that receives the current state and action and outputs the sum of the expected future rewards. The policy then picks the action that maximizes this output. Generally, model-based approaches are sample-efficient, while model-free methods are better at learning complex, high-dimensional tasks. Some approaches [Qureshi et al., 2018, Kahn et al., 2018] also tried to use hybrid methods which would explore the advantages of both model-based and model-free value-based approaches. Regarding value-based deep-RL methods, Deep Q Network (DQN) has been vastly used by the research community [Qureshi et al., 2018], due to its good generalization capabilities and relatively simple training method. DQN, however, can only approximate a discrete action space, requiring continuous applications to be discretized beforehand. Trying to solve this issue, some new approaches such as Deep Deterministic Policy Gradient (DPPG) have been used [Christen et al., 2019, Gu et al., 2017] due to its ability to approximate continuous action spaces.

3   Triplet Ontological Semantic Model

Researches on the cognitive science and neuroscience fields [Burgess, 2008] have shown that the human brain has its own "GPS" mapping system. Every time we revisit a known environment this GPS is responsible for navigating using past known information and update itself with novel data. By relying on relational information instead of precise metric position, the human brain remains unparalleled on its spatial scalability and data efficiency. Robots, on the other hand,
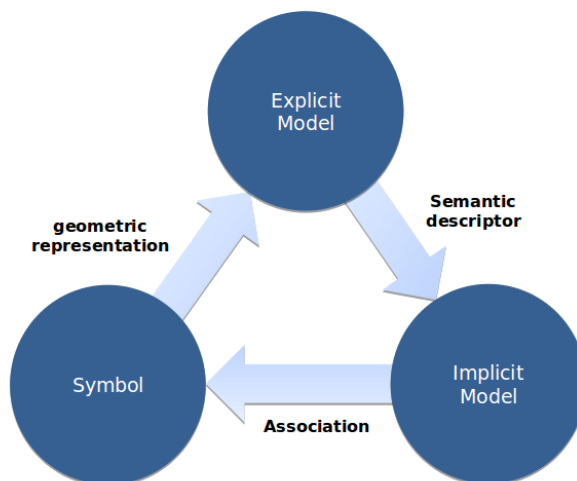
Figure 1: Triplet Ontologic Semantic Model (TOSM) representation.

still heavily rely on information-rich, yet memory inefficient, maps in order to localize themselves and navigate through known environments. Despite being precise, those maps require a large amount of data to be stored, which hinders the robot's long-term autonomy on large scale environments due to lack of storage space. Aiming to mimic the brain GPS model efficiency, the Triplet Ontological Semantic Model(TOSM) [Joo et al., 2019] was developed.

The TOSM representation can be described as three interconnected models as shown on Fig. 1. The explicit model subsumes everything that can be measured or obtained through sensorial means. It can be data such as size, three-dimensional pose, shape, color, texture, etc, which are already vastly used on current robot applications. The implicit model, on the other hand, contains intrinsic knowledge which cannot be obtained by sensors alone, thus needing to be inferred from the available semantic information. The implicit model comprise a large variety of data that range from physical properties such as mass and friction coefficient, relational data (e.g. object A is inside object B), until more complex semantic information such as "An automatic door opens if one waits in front of it". Finally, the symbolic model describes an element using a language-oriented way, such as name, description, identification number and symbols that can represent such element.

By creating an environment database using TOSM encoded data, a hierarchical mapping system was created, based on the findings of cognitive science. As shown on Fig. 2, different maps can be generated on-demand according to the specifications of the robot and the given task. This eliminates the demand to store several different maps by being able to build them only when needed, reducing the data redundancy and improving its storage efficiency. The TOSM can be also used to model places and robots, which combined with the object models can be used to generate high-level semantic maps.

In this work, we also used the TOSM encoded on-demand database to automatically generate a complete simulation environment without the need of domain expert tailored models. This allows the robot to update its mental simulated world automatically just by updating the on-demand database. In order to encode the TOSM data into a machine-readable format, the Ontology Web Language (OWL) was used. OWL is widely used and has an active community which created several tools and applications openly available. We used one of those tools, the Protégé framework, to manipulate and visualize the OWL triplets.

3.1 Robot Description

In order to describe a robot, it was divided into structural parts, sensors, wheels and joints, each of them described by its own explicit, implicit and symbolic information. All categories contain similar explicit data, such as pose, shape, color, size and material. The symbol data contains the part name and an identification number. On the other hand, the implicit data is unique for each category. For structural parts, it contains the mass and the material, while wheels also store whether or not it is a active wheel. Joints store which two parts it is connected to. Moreover, the implicit information can be different for each type of sensor. For example, cameras were described by image resolution, field of view, frames per second and, for RGB-D cameras, range. A laser range finder can have data such as range, view angle and number of samples.
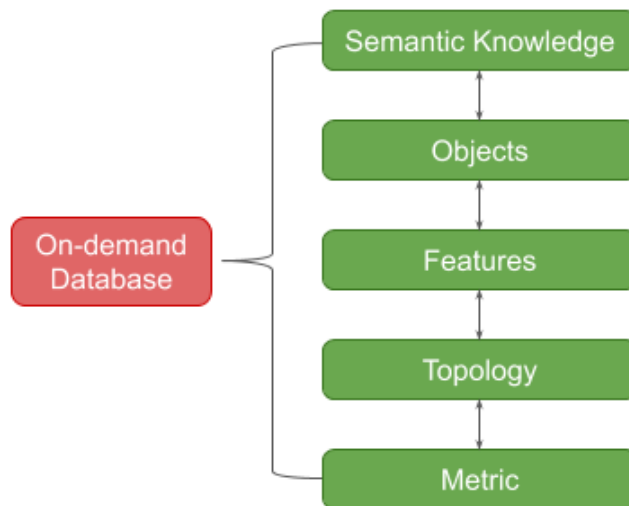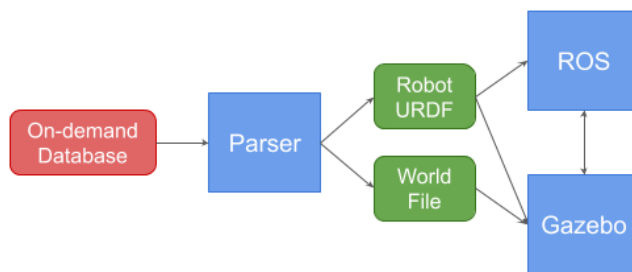
Figure 2: On-demand map generation.

Figure 3: Data flow for the mental simulation.

3.2 Environment Description

The environment can be modeled in a similar fashion as the robot. It is divided mainly into objects and places. Regarding objects, the explicit model contains the same data as described on Subsection 3.1. The implicit model contains data such as mass, material and relational spatial information, such as "in front of", "on the left of", etc. With respect to places, on the other hand, the explicit model contains its boundary points, while the implicit model stores which objects/places are inside of it and which other places is it connected to. The symbol information is the same for both, storing the name of the place/object and an identification number.

4 Mental Simulation

By encoding the TOSM information using the OWL format, it is possible to do semantic reasoning and querying. Before doing any task, the robot can reason about its feasibility by knowing about its surrounding environment's characteristics and its own structure, limitations and properties. For example, a robot only equipped with a laser scanner can reason about its inability to navigate through a corridor made out of glass walls. We extended those reasoning capabilities by automatically generating a complete mental simulation environment using only the on-demand database data.

The data flow for the mental simulation can be seen on Fig 3. Whenever its needed, the robot requests the TOSM data to the on-demand database, and generate two different outputs. The first one is an Universal Robot Description Format(URDF) which is then fed into Robot Operating System(ROS) and Gazebo Simulator in order to control and simulate the virtual robot. The second one is a Gazebo World file which represents the whole environment simulation. Those files are generated on-demand and can be constantly updated whenever the real robot update its database.
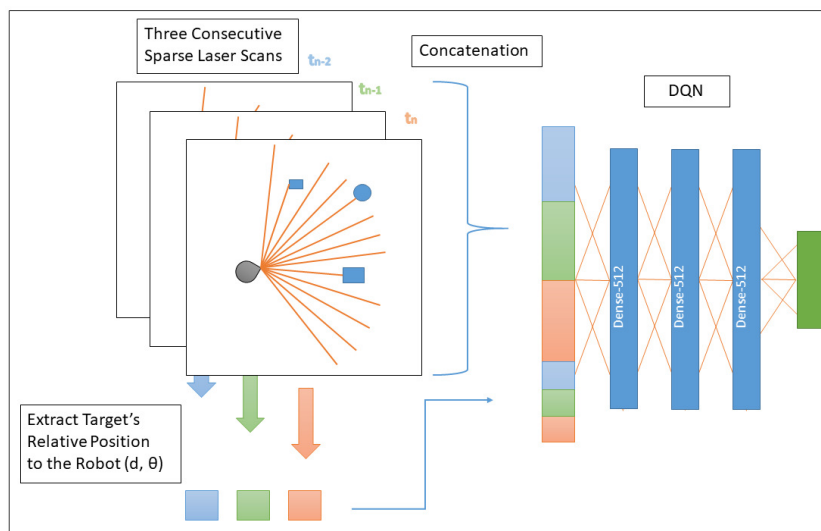
Figure 4: DQN structure.

5 Reinforcement Learning for Autonomous Navigation

In order to show one os the uses for the mental simulator, an autonomous navigation policy was trained using a DQN. The training was performed using a Core i7 CPU and a Nvidia GTX 1060. The OpenAI ROS framework [ezq,] was used in order to to abstract the layer between the reinforcement learning algorithm and the Gazebo/ROS structure. The task learning architecture is shown on Fig. 4. The observation space is composed by the latest three sparse laser scans concatenated with the last three relative distances between the robot and the target way-point. The action space are 15 different angular velocities equally distributed from $-0.5 rad/s$ to $0.5 rad/s$. The rewards were defined as

$$\begin{cases} r_{completion}, & \text{if at goal,} \\ r_{closer}, & \text{if getting closer to the goal,} \\ r_{collision}, & \text{if too close to an obstacle,} \end{cases}$$

where $r_{completion}$, r_{closer} and $r_{collision}$ were defined trivially.

The training was done using an ε-greedy exploration approach, where ε started at 1.0 and decayed until 0.1. The DQN was trained using batches of 64, with learning rate $\alpha = 0.001$ and discount factor $\gamma = 0.996$. The robot was trained for a total of 2000 episodes, where each episode would end in case of completion, collision or after 1000 steps.

6 Results and Discussion

In order to show the usability of such framework, the 7th floor of the Corporate Collaboration Center, Sungkyunkwan University, was added to the on-demand database. Additionally, the differential robot shown on Fig. 5a was modeled. The comparison between the simulated world and the database data can be seen on Fig. 6 while the comparison between the real and simulated robot is shown on Fig. 5.

By automatically generating this simulation environment, we allow the robot to perform mental simulation without the aid of domain experts by reusing the same data it already uses for planning and navigation. Such approach further improves the robot autonomous behavior by letting it simulate itself(or even other robots) on *its own mind* and use this simulated environment to prospect about new actions. Currently, this can be done in two different ways:

- **Learning**: The robot can use the mental simulation to run reinforcement learning algorithms in order to train and learn the execution of new tasks. This is mainly done when the physical robot is idle(e.g. charging at night).

- **Planning**: The robot can simulate its current state and use it for testing a plan, generated by traditional planners, and check whether it succeeds or not. In the case of failure, the robot can re-plan without having to fail on the real environment, allowing for a more robust task execution.
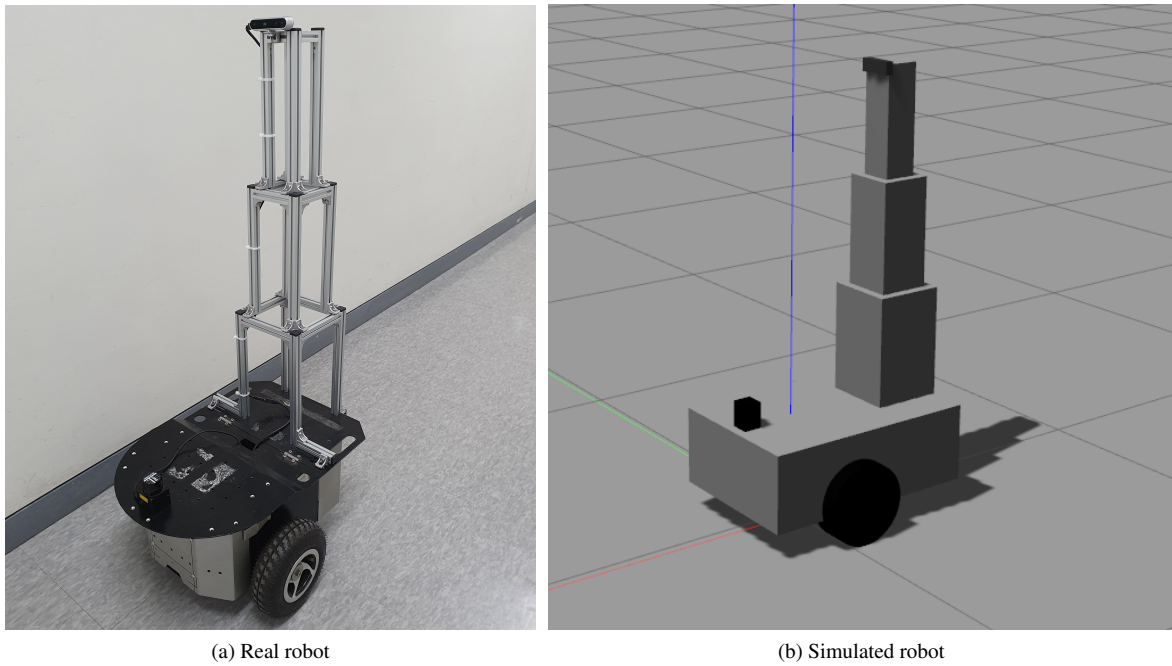
(a) Real robot

(b) Simulated robot

Figure 5: Real and simulated robots comparison.

(a) Visualization of the data obtained from the on-demand DB. Objects
are represented as bounding boxes, while places are represented as col-
ored polygons on the floor

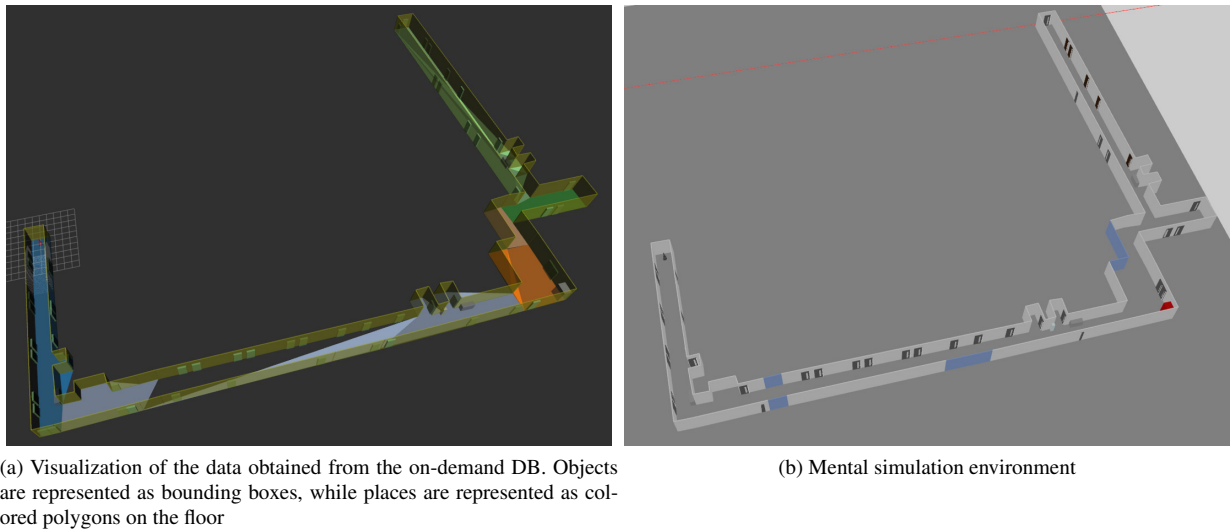(b) Mental simulation environment

Figure 6: Comparison between environment data queried from the DB and mental simulation.
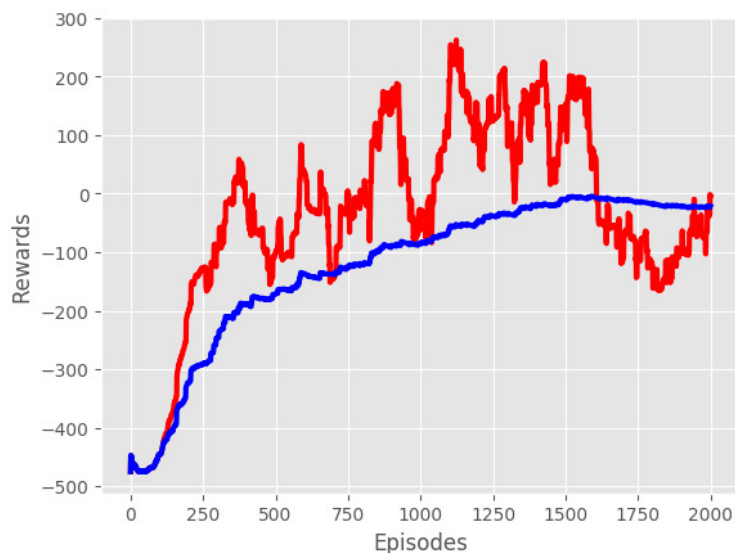
Figure 7: The blue line represents the cumulative average reward, while the red line shows the 100 episodes moving average.

The main advantage of such simulation is that it removes the necessity of a *tailor-made* simulation environment, allowing the robot to generate and update this environment automatically. It can be specially useful for reinforcement learning approaches, which, in theory, gets better the more experiences the robot collects. The robot should be able to run learning algorithms whenever it is idle, slowly improving itself. Naturally, a cluster running multiples CPUs and GPUs would learn orders of magnitude faster, allowing the robot to run the learning algorithms itself bring the robotics field one step closer to true robot autonomy. Finally, by uploading the on-demand database to a cloud infrastructure, robots should be able to share its own model and environment maps, allowing other robots to compare its performance on a given task with one another, and provide this information for its operators automatically.

By using the mental simulation, a autonomous navigation task was learned. The average reward graph can be seen on Fig. 7. Despite being one of the simpler deep-RL approaches, DQN was shown to be good at generalizing a high-dimensional task. However, the whole training took around 30 hours on a mid-range computer. If the same training were performed on a mobile robot, the training times might be too prohibitive. Thus, sample-efficient learning algorithms should be more appropriate for this application.

7   Conclusion and Future Work

In this paper, we presented a method of generating an automatic mental simulation by using a TOSM on-demand database. By allowing robots to create and update mental simulations on a complete autonomous way, we removed the necessity of expert-tailored models, leading for more autonomous robotic systems. In order to show one of the possible applications of such method, we trained the robot to autonomously navigate on an known environment by using a Deep Q Network. We plan now to expand those applications, by including behaviors into the on-demand DB, allowing robots to share and configure RL policies by themselves. We also want to explore the usability of our framework when combined with classical planners.

Acknowledgment

References

[ezq,] Openai ros documentation. Date last accessed 04-Aug-2019.

[Beetz et al., 2018] Beetz, M., Beßler, D., Haidu, A., Pomarlan, M., Bozcuoğlu, A. K., and Bartels, G. (2018). Know rob 2.0—a 2nd generation knowledge processing framework for cognition-enabled robotic agents. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 512–519. IEEE.

[Beetz et al., 2015] Beetz, M., Tenorth, M., and Winkler, J. (2015). Open-ease. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1983–1990. IEEE.

[Boyer, 2008] Boyer, P. (2008). Evolutionary economics of mental time travel? *Trends in cognitive sciences*, 12(6):219–224.

[Buchsbaumm et al., 2005] Buchsbaumm, D., Blumberg, B., Breazeal, C., and Meltzoff, A. N. (2005). A simulation-theory inspired social learning system for interactive characters. In *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005.*, pages 85–90. IEEE.

[Burgess, 2008] Burgess, N. (2008). Spatial cognition and the brain. *Annals of the New York Academy of Sciences*, 1124(1):77–97.

[Christen et al., 2019] Christen, S., Stevsic, S., and Hilliges, O. (2019). Guided deep reinforcement learning of control policies for dexterous human-robot interaction. *arXiv preprint arXiv:1906.11695*.

[Cosgun and Christensen, 2018] Cosgun, A. and Christensen, H. I. (2018). Context-aware robot navigation using interactively built semantic maps. *Paladyn, Journal of Behavioral Robotics*, 9(1):254–276.

[De Carolis et al., 2017] De Carolis, B., Ferilli, S., and Palestra, G. (2017). Simulating empathic behavior in a social assistive robot. *Multimedia Tools and Applications*, 76(4):5073–5094.

[Gordon, 1986] Gordon, R. M. (1986). Folk psychology as simulation. *Mind & Language*, 1(2):158–171.

[Gray and Breazeal, 2005] Gray, J. and Breazeal, C. (2005). Toward helpful robot teammates: A simulation-theoretic approach for inferring mental states of others. In *Proceedings of the AAAI 2005 workshop on modular construction of human-like intelligence*.

[Gu et al., 2017] Gu, S., Holly, E., Lillicrap, T., and Levine, S. (2017). Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates. In *2017 IEEE international conference on robotics and automation (ICRA)*, pages 3389–3396. IEEE.

[Gupta et al., 2004] Gupta, R., Kochenderfer, M. J., Mcguinness, D., and Ferguson, G. (2004). Common sense data acquisition for indoor mobile robots. In *AAAI*, pages 605–610.

[Hamrick, 2019] Hamrick, J. B. (2019). Analogues of mental simulation and imagination in deep learning. *Current Opinion in Behavioral Sciences*, 29:8–16.

[Hesslow, 2012] Hesslow, G. (2012). The current status of the simulation theory of cognition. *Brain research*, 1428:71–79.

[Horii et al., 2016] Horii, T., Nagai, Y., and Asada, M. (2016). Imitation of human expressions based on emotion estimation by mental simulation. *Paladyn, Journal of Behavioral Robotics*, 7(1).

[Joo et al., 2019] Joo, S.-H., Manzoor, S., Rocha, Y. G., Lee, H.-U., and Kuc, T.-Y. (2019). A realtime autonomous robot navigation framework for human like high-level interaction and task planning in global dynamic environment. *arXiv preprint arXiv:1905.12942*.

[Kahn et al., 2018] Kahn, G., Villaflor, A., Ding, B., Abbeel, P., and Levine, S. (2018). Self-supervised deep reinforcement learning with generalized computation graphs for robot navigation. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1–8. IEEE.

[Kahneman and Tversky, 1981] Kahneman, D. and Tversky, A. (1981). The simulation heuristic. Technical report, STANFORD UNIV CA DEPT OF PSYCHOLOGY.

[Kennedy et al., 2009] Kennedy, W. G., Bugajska, M. D., Harrison, A. M., and Trafton, J. G. (2009). "like-me" simulation as an effective and cognitively plausible basis for social robotics. *International Journal of Social Robotics*, 1(2):181–194.

[Kostavelis et al., 2016] Kostavelis, I., Charalampous, K., Gasteratos, A., and Tsotsos, J. K. (2016). Robot navigation via spatial and temporal coherent semantic maps. *Engineering Applications of Artificial Intelligence*, 48:173–187.

[Kunze and Beetz, 2017] Kunze, L. and Beetz, M. (2017). Envisioning the qualitative effects of robot manipulation actions using simulation-based projections. *Artificial Intelligence*, 247:352–380.

[Laird, 2001] Laird, J. E. (2001). It knows what you're going to do: adding anticipation to a quakebot. In *Proceedings of the fifth international conference on Autonomous agents*, pages 385–392. ACM.

[Lenat, 1995] Lenat, D. B. (1995). Cyc: A large-scale investment in knowledge infrastructure. *Communications of the ACM*, 38(11):33–38.

[Niles and Pease, 2001] 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*, pages 2–9. ACM.

[Polceanu and Buche, 2017] Polceanu, M. and Buche, C. (2017). Computational mental simulation: A review. *Computer Animation and Virtual Worlds*, 28(5):e1732.

[Qureshi et al., 2018] Qureshi, A. H., Nakamura, Y., Yoshikawa, Y., and Ishiguro, H. (2018). Intrinsically motivated reinforcement learning for human–robot interaction in the real-world. *Neural Networks*, 107:23–33.

[Rajeswaran et al., 2017] Rajeswaran, A., Kumar, V., Gupta, A., Vezzani, G., Schulman, J., Todorov, E., and Levine, S. (2017). Learning complex dexterous manipulation with deep reinforcement learning and demonstrations. *arXiv preprint arXiv:1709.10087*.

[Shah et al., 2018] Shah, P., Fiser, M., Faust, A., Kew, J. C., and Hakkani-Tur, D. (2018). Follownet: Robot navigation by following natural language directions with deep reinforcement learning. *arXiv preprint arXiv:1805.06150*.

[Tai et al., 2017] Tai, L., Paolo, G., and Liu, M. (2017). Virtual-to-real deep reinforcement learning: Continuous control of mobile robots for mapless navigation. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 31–36. IEEE.

[Tenorth and Beetz, 2009] Tenorth, M. and Beetz, M. (2009). Knowrob—knowledge processing for autonomous personal robots. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4261–4266. IEEE.

[Waibel et al., 2011] Waibel, M., Beetz, M., Civera, J., d'Andrea, R., Elfring, J., Galvez-Lopez, D., Häussermann, K., Janssen, R., Montiel, J., Perzylo, A., et al. (2011). Roboearth-a world wide web for robots. *IEEE Robotics and Automation Magazine (RAM), Special Issue Towards a WWW for Robots*, 18(2):69–82.

[Zhu et al., 2017] Zhu, Y., Mottaghi, R., Kolve, E., Lim, J. J., Gupta, A., Fei-Fei, L., and Farhadi, A. (2017). Target-driven visual navigation in indoor scenes using deep reinforcement learning. In *2017 IEEE international conference on robotics and automation (ICRA)*, pages 3357–3364. IEEE.