Human Intelligence Assisted Robot Learning

Zi Wang (MIT CSAIL)

Robot learning is an emerging field where we apply machine learning techniques to robotics tasks in the hope that, one day, robots can be deployed in a variety of environments, capable of completing a range of tasks such as doing laundry, cooking a nice dinner or cleaning up a messy room. Under the umbrella of robot learning, our wildest dream is that we humans do not need to do the hard work of hand coding certain behaviors for the robot any more and the robot can just learn everything by itself through trial and error. However, it has gradually become clear to us that simply applying end-to-end approaches using, for example, deep reinforcement learning [7] based approaches with sparse rewards, to long-horizon problems is unfortunately very data inefficient and impractical in reality, and structures must be imposed to the learning problem to reduce the sample complexity [1]. Under the hood, though not emphasized, human intelligence is still playing an indispensable role to define those structures and certain behaviors of the robot.

Instead of "hiding" the usage of human intelligence in learning algorithms, can we make use of it more explicitly to assist machine learning and artificial intelligence? This is the central question of my research, and my goal is to find a sweet spot balancing between human intelligence and statistical machine learning. This means a shift in the mindset of how to enable robot learning: instead of collecting large-scale data with an army of robots operating in parallel [8], we use expert domain knowledge to make up for the weakness of our learning robots. In the following, I will give some examples of how we tackle the problem of robot learning with help from human intelligence from both practical and theoretical perspectives.

I. Robot learning: how to integrate human knowledge with learning algorithms

Kitchen2D [14] is a framework in a 2D simulated kitchen that demonstrates how active learning and human-defined abstract models of skills can be integrated. Skills are structured actions such as pouring and pushing. For a robot to be effective in a domain that combines novel skills with long-horizon, high-level task objectives, it is necessary to acquire models of these skills. In our approach, we use expert domain knowledge to define the model as a constraint on relevant parameters of the skill. The constraint describes the condition of the parameters under which the skill can achieve the desired effect. Then, the robot learns the constraint by actively gathering data through experimentation on using the skill with different parameter settings.

In **SPARE** [17], we move one more small step toward learning and let the learning algorithm figure out what the relevant parameters are for a model of the skill. Human intelligence, as the assistance to machine intelligence, provide the information on deictic references that defines a sparse set of relations among objects in the scene, so that the learning algorithm can decide which relations to use in order to find the set of appropriate parameters.

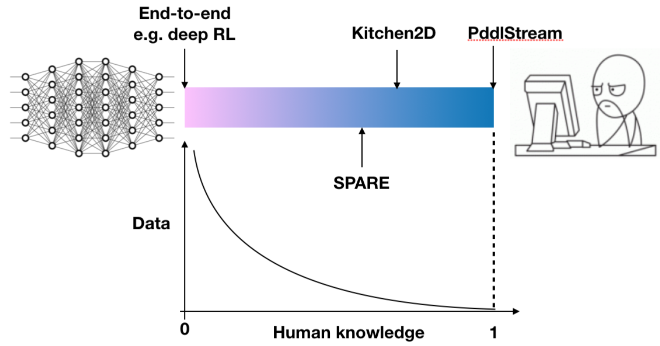

Fig. 1. My goal is to find the balance between human knowledge encoding and statistical learning in an autonomous agent. In the leftmost part of the spectrum, we do not impose much domain knowledge and end-to-end approaches [7] need large-scale data to achieve good performance. In the rightmost, no data need to be collected but an expert human engineer need to hand-code the entire framework, e.g. PddlStream [2]. My work, in contrast to both extremes, lies in the middle, e.g. Kitchen2D [14] and SPARE [17].

Once we obtain the models of the skills, we as engineers encode the models to a task and motion planner [2] and use the model to produce suitable parameter values to try out in the planner. In [5, 14, 16], we adaptively adjust the sequence of recommended parameter values so that the planning problem can be solved more efficiently.

II. Theoretical foundations on the sample complexity of active data collection in robot learning

The applications explained above boil down to the question of how to actively aquire data for learning and decision making. One can abstract it as a mathematical zero-th order optimization problem for non-convex, multi-peak functions. We further study the underlying principles of such a problem, entitled Bayesian optimization, and use the intuitions gained from the theoretical studies to help to design algorithms for robot learning through active data collection.

Bayesian optimization is one way to formulate the zero-th order optimization problem in both discrete and continuous action spaces, with a touch of the ideas in Bayesian modeling. We assume there is a scoring function that measures how good each action is. The goal is to optimize the function by actively querying the scores of a sequence of actions. I have worked on three fundamental challenges: (a) the design of action selection criteria based on Gaussian processes [11, 12]; (b) an empirical Bayes approach for theoretically-verified model selection [16]; and (c) scaling up to high-dimension actions [13] and large-scale observations [15]. My work has addressed the critical computational issues of Bayesian optimization in both decision making and model updating [16], and revealed previously unknown theoretical insights into the regret bound and connections among a number of existing action selection criteria. Notably, the decision criterion I designed [11] was featured in a popular Bayesian optimization package called GPflowOpt [6] and works from other groups [4, 10, 9, 3] have been developed building on my research findings.

III. Future work

There are two lines of future work that I am especially interested in tackling, following the philosophy that human and machines can work harmoniously together.

First, how do we understand human intelligence? Despite the practical experience of putting human knowledge into the right place that co-exist with learning approaches, it is still not clear what exactly the human expert has experienced to find what knowledge needs to be encoded and which part can be learned. We may be able to treat human as part of our dual-agent system once we have a good model for human instincts and decision-making. The understanding of human intelligence will also be inspiring to design better machine learning approaches.

Second, the dual problem of human intelligence assisted machine learning is machine intelligence supported human learning. It is predicted that automation is going to eliminate a lot of jobs, so we need to start thinking about how to create new jobs. One way to use our approach is to solve the inverse problem: how to enhance human intelligence in the dual system such that humans can be trained to tackle difficult and specialized problems which machines cannot solve very well, for example, care taking of the elderly and children, creative thinking and research, problem abstraction, etc. This problem is particularly interesting to the public because automation is real and losing the current mundane jobs will not necessarily result in a better job without preparation. We as scientists have a responsibility to find out a good way to address the relations between machines and humans.

References

[1] Jacob Andreas, Dan Klein, and Sergey Levine. Modular multitask reinforcement learning with policy sketches. In *International Conference on Machine Learning (ICML)*, 2017.

[2] Caelan Reed Garrett, Tomás Lozano-Pérez, and Leslie Pack Kaelbling. Sampling-based methods for factored task and motion planning. *IJRR*, 2018.

[3] Trong Nghia Hoang, Quang Minh Hoang, Ruofei Ouyang, and Kian Hsiang Low. Decentralized high-dimensional Bayesian optimization with factor graphs. In *AAAI*, 2018.

[4] Tarun Kathuria, Amit Deshpande, and Pushmeet Kohli. Batched Gaussian process bandit optimization via determinantal point processes. In *Advances in Neural Information Processing Systems (NIPS)*, 2016.

[5] Beomjoon Kim, Zi Wang, Leslie Pack Kaelbling, and Tomas Lozano-Perez. Learning to guide task and motion planning using score-space representation. *The International Journal of Robotics Research*, 2019.

[6] Nicolas Knudde, Joachim van der Herten, Tom Dhaene, and Ivo Couckuyt. GPflowOpt: A Bayesian optimization library using TensorFlow. *arXiv preprint arXiv:1711.03845*, 2017.

[7] Sergey Levine, Chelsea Finn, Trevor Darrell, and Pieter Abbeel. End-to-end training of deep visuomotor policies. *Journal of Machine Learning Research*, 2016.

[8] Sergey Levine, Peter Pastor, Alex Krizhevsky, Julian Ibarz, and Deirdre Quillen. Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. *The International Journal of Robotics Research*, 2018.

[9] Paul Rolland, Jonathan Scarlett, Ilija Bogunovic, and Volkan Cevher. High-dimensional Bayesian optimization via additive models with overlapping groups. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2018.

[10] Binxin Ru, Michael Osborne, and Mark McLeod. Fast information-theoretic Bayesian optimisation. In *International Conference on Machine Learning (ICML)*, 2018.

[11] Zi Wang and Stefanie Jegelka. Max-value entropy search for efficient Bayesian optimization. In *International Conference on Machine Learning (ICML)*, 2017.

[12] Zi Wang, Bolei Zhou, and Stefanie Jegelka. Optimization as estimation with Gaussian processes in bandit settings. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2016.

[13] Zi Wang, Chengtao Li, Stefanie Jegelka, and Pushmeet Kohli. Batched high-dimensional Bayesian optimization via structural kernel learning. In *International Conference on Machine Learning (ICML)*, 2017.

[14] Zi Wang, Caelan Reed Garrett, Leslie Pack Kaelbling, and Tomas Lozano-Perez. Active model learning and diverse action sampling for task and motion planning. In *International Conference on Intelligent Robots and Systems (IROS)*, 2018.

[15] Zi Wang, Clement Gehring, Pushmeet Kohli, and Stefanie Jegelka. Batched large-scale Bayesian optimization in high-dimensional spaces. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2018.

[16] Zi Wang, Beomjoon Kim, and Leslie Pack Kaelbling. Regret bounds for meta bayesian optimization with an unknown gaussian process prior. In *Neural Information Processing Systems (NeurIPS)*, 2018.

[17] Victoria Xia, Zi Wang, Kelsey Allen, Tom Silver, and Leslie Pack Kaelbling. Learning sparse relational transition models. In *International Conference on Learning Representations (ICLR)*, 2019.