Projet de recherche doctoral

Titre: Hierarchical and Model-based Goal-oriented Behavior Generation for an Open-Ended Learning Agent

Context and objectives

Most existing software agents and robots suffer from insufficient versatility [Plappert et al., 2018], limiting the potential introduction of these agents and robots in our everyday life. In most cases, some expertise is required from a human engineer to design their behavior in anticipation of the situations they may encounter. The inflexibility of such hand-written controllers results in their incapability of addressing novel or unforeseen situations.

In sharp contrast with these limitations, open-ended learning [Doncieux et al., 2018] is an emerging machine learning paradigm in recent Artificial Intelligence research where autonomous agents have to discover their own tasks and learn how to solve them, driven by some intrinsic motivations.

Most of such “autotelic” agents set their own goal in a predefined goal space through intrinsic motivations such as curiosity or willingness to control their environment and call upon the reinforcement learning framework to improve their capability to reach their goals [Oudeyer et al., 2007 , Colas et al., 2022].

But a further defining property of open-ended learning that the current agents do not possess so far is the capability to extract on the fly from a rich sensorimotor flow an adequate representation of the problem they are solving. Such representations should compactly specify the states and actions that the agent considers so as to facilitate the immediate resolution of each problem the agent faces. Besides, we would like the agent to reuse knowledge learned from a task when solving other tasks.

The general objective of the PhD is to endow an agent with this key capability to immediately build an adequate representation of the problem it is solving. Such endeavor results in specific challenges that we describe below.

Specific challenges

In standard multitask or autotelic learning, all the tasks faced by an agent are specified using the same observation and action spaces, which generally correspond to the raw sensors and actuators of the agent. The policy of the agent is then represented as a unique task-conditioned multi-layer neural network, and the capability to transfer knowledge from one task to another builds on the generalization property of this network.

If the agent now designs specific state and action spaces for each tasks, several consequences arise. First, these spaces must be extracted from the raw sensorimotor flow, suggesting a hierarchical representation learning approach, where a unique neural network does not suffice. Second, if each task comes with its own state and action spaces, the input and output of the corresponding task-controllers change, hence relying on a unique network to leverage its generalization properties becomes much less straightforward.

To our knowledge, despite growing efforts in the design of more and more capable agents endowed with hierarchical learning, model-based learning, or intrinsically motivated learning capabilities (e.g. [Islam et al., 2022]), these specific challenges have not been addressed so far in the literature. If the thesis is successful, it will thus provide an original building block in the more general effort to design the next generation of more flexible and autonomous agents.

References :

[Colas et al., 2018] Cédric Colas, Tristan Karch, Olivier Sigaud, and Pierre-Yves Oudeyer. Autotelic agents with intrinsically motivated goal-conditioned reinforcement learning: a short survey. Journal of Artificial Intelligence Research, 74:1159–1199, 2022.

[Doncieux et al., 2018] Doncieux, S., Filliat, D., ... & Sigaud, O. (2018). Open-Ended Learning: A Conceptual Framework Based on Representational Redescription. *Frontiers in neurorobotics*, *12*.

[Islam et al., 2022] Riashat Islam et al. Discrete factorial representations as an abstraction for goal conditioned reinforcement learning. Advances in Neural Information Processing Systems, 36, 2022

[Oudeyer et al., 2007] Oudeyer, P. Y., Kaplan, F., & Hafner, V. V. (2007). Intrinsic motivation systems for autonomous mental development. *IEEE transactions on evolutionary computation*, *11*(2), 265-286.

[Péré et al., 2018] Péré, A., Forestier, S., Sigaud, O., & Oudeyer, P. Y. (2018). Unsupervised learning of goal spaces for intrinsically motivated goal exploration. *arXiv preprint arXiv:1803.00781*.

[Plappert et al., 2018] Matthias Plappert, Marcin Andrychowicz, et al. Multi-goal reinforcement learning: Challenging robotics environments and request for research. arXiv preprint arXiv:1802.09464, 2018