

Curriculum-Based Deep Reinforcement Learning for Adaptive Robotics: A Mini-Review

Kashish Gupta and Homayoun Najjaran*

School of Engineering, University of British Columbia, 3333 University Way, Canada

Abstract

To facilitate the current and future automation needs, the research community constantly seeks to develop dynamic and efficient autonomous decision-making agents. These agents must not only be robust to modeling uncertainties, internal and external changes, but can adapt to a range of tasks also. Recent progress in deep reinforcement learning has corroborated to its potential to train such autonomous and robust agents. At the same time, the introduction of curriculum learning has made the reinforcement learning process significantly more efficient and allowed for training on much broader tasks. This combination, Curriculum-based Deep Reinforcement Learning (CDRL), presents a powerful solution to meet the increasing complexity of today's automation industry that demands highly intelligent machines. With this work we present a concise review of CDRL methods within the context of their application to the field of adaptive robotics.

Keywords

Reinforcement learning, Curriculum learning, Curriculum-based reinforcement learning, Adaptive robotics

Introduction

In this work, we briefly introduce the fields of reinforcement and curriculum learning and highlight curriculum-based deep reinforcement learning methods and their robotics applications. While the conceptual curriculum generation is a wide topic [1], we focus on recent works where a curriculum is utilized in a strict sense and the method is demonstrated for a robotics environment with a neural network-based reinforcement learning agent.

Reinforcement Learning

In the last couple of decades, the field of Deep Reinforcement Learning (DRL) has established itself

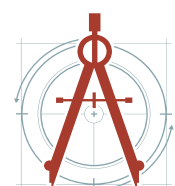
as a vital paradigm of Machine Learning (ML) and Artificial Intelligence (AI) [2]. Fueled by the recent advancements in processing hardware technology, the potential of the field has attracted research interest and scientific contributions from several domains like robotics, chemistry, neuroscience, gaming, business services and web management industries [3]. The field boasts an extremely powerful framework that allows for a software learner (dubbed agent) to observe and manipulate the world (dubbed environment). The outcome of the interaction between the agent and the environment, if successful, is a trained agent that maps simulated data into real-world actions. The nature

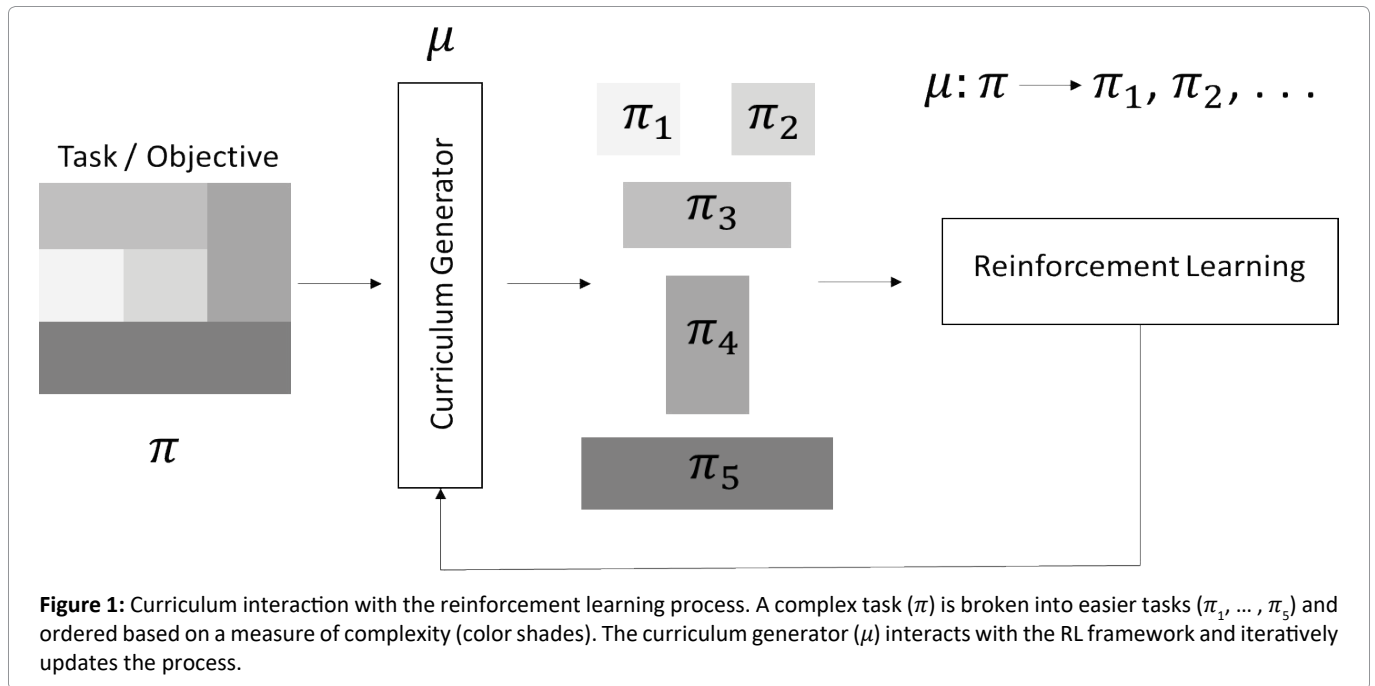
*Corresponding author: Homayoun Najjaran, School of Engineering, University of British Columbia, 3333 University Way, Kelowna, BC V1V1V7, Canada

Accepted: May 04, 2021; Published: May 06, 2021

Copyright: © 2021 Gupta K, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Gupta and Najjaran. *Int J Robot Eng* 2021, 6:031





of the framework and its ability to train an agent without an explicit supervisor poses a significant advantage, which however is surmounted by its need for colossal amounts of data. While reinforcement learning continues to dominate the simulation-based applications, there is a principal need to study and apply sample-efficient extensions to the field, viz. curriculum learning, to enable applicability in practical and real-world scenarios [4].

Curriculum Learning

Formally proposed in 2009 [5], curriculum learning derives its inspiration from human learning style and the intuition that it is easier to learn by accumulating knowhow through simple tasks rather than trying to tackle a complex task directly. Figure 1 demonstrates the process as a flow diagram, where a given task (π) is broken down in multiple subtasks ($\pi_1, \pi_2, \dots, \pi_5$) and presented to the learning RL agent in a particular order. Sometimes studied as a specific case of continuation methods and dynamic programming [5], curriculum learning strategies have been widely employed to improve supervised and reinforcement learning algorithms. Curriculum-based Reinforcement Learning (CRL) agents have shown significant improvement in convergence rates, training time and generalization capabilities when compared to standalone RL agents [6]. Moreover, CRL agents have also been able to solve tasks where standalone RL agents prove intractable [7]. While the nature of the curriculum itself has a myriad of possibilities, the curriculum

design process can be either domain-expert specified [8] or automatic [9]. Domain-expert specified curricula rely heavily on human knowledge and thus lack new knowledge discovery, scalability, and robustness to unseen scenarios and may exhibit bias towards certain tasks. Automatic curricula, on the other hand, can offer what the latter lacks but are extremely difficult to generate and implement. The amalgamation of curriculum learning with reinforcement learning allows an agent to train in complex scenarios and learn a range of tasks, for which the domain of robotics stands to be the perfect recipient.

Adaptive Robotics

Evolved from traditional industrial robots and static automation, adaptive robots boast the ability to adapt to dynamic environmental changes, handle numerous tasks and safely collaborate with human agents. The interconnectivity and interoperability of adaptive robots further enable them to be an indispensable discipline for modern and future industrial, manufacturing, medical and assistive technological needs. Nonetheless, advancing robots beyond conventional automation pose unparalleled challenges like autonomous decision-making and semantic perception, specially for real-world implementations. Through the introduction of artificial intelligence and machine learning, the current state-of-the-art adaptive robots show tremendous potential and invites interest from numerous research domains.

The robotics research community investigating the notion of behavioral adaptivity is actively seeking solutions to alleviate challenges including efficient and automatic curriculum generation, multi-task learning and sim-to-real curriculum generation to improve applicability and implementation across a wide range of domains. Curriculum-based reinforcement learning methods allow the field of adaptive robotics to surpass its era of explicitly programmed robots with limited functionalities and thrust the progress in automation towards a true behaviorally adaptive machine intelligence to train and implement sophisticated and hard-to-engineer behaviors.

Work in [10] extends the work of [11] to design a curriculum based on Hindsight Experience Replay (HER) transitions. The design of curriculum is based on the measure of similarity between the intended and achieved goals while tuning the measure for diversity among goals to aid exploration. On the other hand, [12] argues that a natural curriculum can be learnt in competitive multi-agent environments and that such a setting can produce complex behaviours that surpass the complexity of the environment itself. Similarly, cooperative multi-agents [13], diverse environments [14] and self-plays [15] have demonstrated better generalization and faster training using intrinsic and natural curriculums. Work in [16] proposes a curriculum of start states for a constant goal state while the work in [16] estimates the next or intermediary goal of the appropriate complexity using a Generative Adversarial Network (GAN). [17] uses the backward reachability decomposition between different goal and start states to estimate reachability and obtain a measure of complexity of the task [18] quantifies the complexity of different training environments (teacher network) and generates a curriculum for the student networks to learn. [19] masks certain features of the goal vector to reduce its complexity and trains the agent on the reduced complexity goal states. It has also been shown that the development of a simple curriculum based on the accuracy requirements of a given task, on basis of a measure of a degree of competence like [20], leads to faster and more efficient training [21]. An extension to the work with Universal Value Function Approximators (UVFA) [22] is studied in [23]. While the current state of research mostly focusses on simulated objectives, notable efforts are being put into introducing simulation-trained intelligent and highly adaptive agents to the real-world [24-26].

In learning-based robotics, end-to-end autonomy often involves three principle components-perception, cognition, and control. The components are complementary in nature which allows for intelligent perception and cognition algorithms to enable autonomous control. Compared to traditional robotics where planning is an independent phase, the cognition phase is often embedded within the control framework of the learning-based robotics. In this way, the combination of cognition and control constitute the decision making process [27]. The curriculum-based reinforcement learning methods strictly focus on the decision-making framework of autonomous agents. Depending on the nature of the training data available, the decision-making agent framework can range from an end-to-end solution that also includes perception to one that only captures a low-dimensional control strategy. An example of the former can be the implementation of deep drive that replaces all the steps of a traditional Autonomous Vehicle (AV) process. A comparison of different implementations of learning-based AV can be found in [28]. In the same way, RL-based training is a quintessential choice for implementation that can collapse different required processes of industrial robots into the decision-making process of an adaptive robot. Task-specific trajectory planning from traditional control methods have proven to be more efficient for designated tasks [29,30] though may lack the flexibility to repurpose a robot to cope with the changes of dynamic environments. On the other hand, optimization techniques from various trajectory planning methods [31], have effectively been adopted for reinforcement learning agents [32,33] that allow for better convergence and learning of the autonomous decision-making process. The introduction of curriculum to RL training process (CRL) helps learning by improving sample efficiency and generalization capabilities of the agent. The ability of CRL to train incrementally complex agents based on scheduled experiences allow for maximal learning required for the promotion of adaptive robots in practical applications.

Acknowledgement

We would like to acknowledge the financial support of Natural Sciences and Engineering Council Canada (NSERC) under the Discovery Grant No. 341887 of Dr. Homayoun Najjaran.

Conflict of Interest

The authors report no conflict of interest.

References

- Narvekar S, Stone P (2019) Learning curriculum policies for reinforcement learning. In: AAMAS 19: Proceedings of the 18th International Conference on Autonomous Agents and Multi Agent Systems, 25-33.
- Gu S, Holly E, Lillicrap T, Levine S (2017) Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates. In Proceedings - IEEE International Conference on Robotics and Automation, 3389-3396.
- Sutton RS, Barto AG (2017) Reinforcement learning. An Introduction. (2nd edn).
- Henderson P, Islam R, Bachman P, Pineau J, Precup D, et al. (2018) Deep reinforcement learning that matters, 3.
- Bengio Y, Louradour J, Collobert R, Weston J (2009) Curriculum learning. In ACM International Conference Proceeding Series 382: 1-8.
- Narvekar S, Peng B, Leonetti M, Sinapov J, Taylor ME, et al. (2020) Curriculum learning for reinforcement learning domains: A Framework and Survey 21: 1-50.
- Florensa C, Held D, Wulfmeier M, Zhang M, Abbeel P (2017) Reverse curriculum generation for reinforcement learning.
- D mendoza barrenechea, E Johns (2017) Curriculum learning for robot manipulation using deep reinforcement learning.
- Karpathy A, Van De Panne M (2012) Curriculum learning for motor skills. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics 7310: 325-330.
- Fang M, Zhou T, Du Y, Han L, Zhang Z (2019) Curriculum-guided hindsight experience replay. In: Advances in Neural Information Processing Systems 32: 12623-12634.
- Andrychowicz M, Wolski F, Ray A, Schneider J, Fong R, et al. (2017) Hindsight experience replay.
- Bansal T, Pachocki J, Sidor S, Sutskever I, Mordatch I (2017) Emergent complexity via multi-agent competition.
- Yang J, Nakhaei A, Isele D, Fujimura K, Zha H (2018) CM3: Cooperative multi-goal multi-stage multi-agent reinforcement learning.
- Heess N, Dhruva TB, Sriram S, Lemmon J, Merel J, et al. (2017) Emergence of locomotion behaviours in rich environments.
- Sukhbaatar S, Lin Z, Kostrikov I, Synnaeve G, Szlam A, et al. (2017) Intrinsic motivation and automatic curricula via asymmetric self-play. 6th Int Conf Learn Represent ICLR 2018 Conf Track Proc.
- Florensa C, Held D, Geng X, Abbeel P (2018) Automatic goal generation for reinforcement learning agents.
- Ivanovic B, Harrison J, Sharma A, Chen M, Pavone M (2019) BaRC: Backward reachability curriculum for robotic reinforcement learning. IEEE International Conference on Robotics and Automation 2019: 15-21.
- Portelas R, Colas C, Hofmann K, Oudeyer PY (2019) Teacher algorithms for curriculum learning of Deep RL in continuously parameterized environments.
- Eppe M, Magg S, Wermter S (2019) Curriculum goal masking for continuous deep reinforcement learning. In 2019 Joint IEEE 9th International Conference on Development and Learning and Epigenetic Robotics, ICDL-EpiRob, 183-188.
- Baranes A, Oudeyer PY (2013) Active learning of inverse models with intrinsically motivated goal exploration in robots. Rob Auton Syst 61: 49-73.
- Fournier P, Sigaud O, Chetouani M, Oudeyer PY (2018) Accuracy-based curriculum learning in deep reinforcement learning.
- Schaul T, Horgan D, Gregor K, Silver D (2015) Universal value function approximators. PMLR 37: 1312-1320.
- Luo S, Kasaei H, Schomaker L (2020) Accelerating reinforcement learning for reaching using continuous curriculum learning.
- Breyer M, Furrer F, Novkovic T, Siegwart R, Nieto J (2018) Flexible robotic grasping with sim-to-real transfer based reinforcement learning.
- Bin Peng X, Andrychowicz M, Zaremba W, Abbeel P (2017) Sim-to-real transfer of robotic control with dynamics randomization.
- Ugnow H (2019) MEng individual project CuRL: Curriculum reinforcement learning for goal-oriented robot control.
- Dimitri Bertsekas P (2019) Reinforcement learning and optimal control book. Athena Scientific.
- Fayyad J, Jaradat MA, Gruyer D, Najjaran H (2020) Deep learning sensor fusion for autonomous vehicle perception and localization: A review. Sensors 20: 4220.

29. Bottin M, Rosati G (2019) Trajectory optimization of a redundant serial robot using cartesian via points and kinematic decoupling. *Robotics* 8: 101.
30. Bottin M, Rosati G, Cipriani G (2021) Iterative path planning of a serial manipulator in a cluttered known environment. In *Mechanisms and Machine Science* 91: 237-244.
31. Ratiu M, Adriana Prichici M (2017) Industrial robot trajectory optimization-A review. *MATEC Web Conf* 126: 02005.
32. Mahmood T, Ricci F (2009) Improving recommender systems with adaptive conversational strategies. In *Proceedings of the 20th ACM Conference on Hypertext and Hypermedia, HT'09*, 73-82.
33. Recht B (2019) A tour of reinforcement learning: The view from continuous control. *Annu Rev Control Robot Auton Syst* 2: 253-279.

