PROMOTING REINFORCEMENT LEARNING IN ROBOTICS: AN EVOLUTIONARY PERSPECTIVE AND QUANTUM IMPLEMENTATIONS

Efe Utku
1284290
Department of Applied Physics
e.utku@student.tue.nl

1 Evolutionary Perspective on RL

In psychology and decision theory, “decision making” can be defined as “the cognitive process of choosing between two or more alternatives, ranging from the relatively clear cut to the complex(VandenBos, n.d.); and it is a very fundamental ability of intelligent agents. A key to possess this ability is to have motivations such that every made decision can be reasoned back to its motivationally driven process. The reason this sort of technique is uniformly used by humans and other primates to fishes and insects is not a coincidence; it's rather an evolutionary necessity to survive in an ever-changing environment. Every favourable decision made in the context of a driving motivation is rewarded by the dopaminergic reward system(Arias-Carrión, Stamelou, Murillo-Rodríguez, Menéndez-González, & Pöppel, 2010) in most of the living creatures and rewards are crucial objects that induce learning, approach behaviour and choices.(Schultz, 2015) In some sense, the relation between motivation and decision is analogous to the relation between reward and learning; and this is the case that needs to be applied to AI robotics to achieve more capable agents in terms of decision making, adaptability and; as an outcome, functionality. Currently, a promising area in ML called reinforcement learning (RL) is being broadly studied which forms a potential candidate for that role.

Reinforcement learning (RL) is similar to motivation-based learning in humans that evolved over millions-of-years; but it uses Markov Decision Processes (MDP) to optimize the reward and consequently performs operant conditioning on the AI agent. Despite the differences between human decision making and reinforcement learning in AI agents, the way they evaluate information as a mean to the outcome is comparable to one another. They both work based on principles of reward/ punishment, allowing the agent to adapt its behaviour and tend to maximize the reward that can be obtained by the agent without any declaration of inputs/ outputs.(Kaelbling, Littman, & Moore, 1996) This a fairly different strategy than the ones that are used by e.g. supervised/ unsupervised learning models, which work solely on layers and exact mathematical mapping of the output from the input. In addition, the usage of MDPs in RL to model the environment adds a level of non-determinism and makes the agent more adaptable, which is essential for robotic systems acting in interactive environments. Although it can be argued that supervised learning models are also adaptive with hidden layers; the increase in hidden layers eventually forms a “black box” where computations are done, and motivated reasoning is lost. When, one considers motivation/ reward as the driving force and a parameter describing the distance from the current state to the desired state; the loss becomes more significant.

2 Future Quantum Implementations of RL

Another characteristic that promotes RL in the field of robotics is due to the possible quantum enhancements it admits. RL can be viewed as an optimization problem; given the problem and state definitions, it searches the space of behaviours to find the best solution, favouring the motivation/ reward. That's why some aspects of RL are closely related to search and planning issues in AI (Kaelbling et al., 1996), which are proven to be boosted by the developments in the field of quantum computing. Quantum accelerated algorithms exploit quantum phenomena to perform complex calculations faster and more accurate. They can be used; from simple optimization tasks to problems in the complexity class NP-Complete.(Petschnigg, Brandstotter, Pichler, Hofbaur, & Dieber, 2019) Quantum reinforcement learning can be realized using the properties of state superposition principle and quantum parallelism (Dong, Chen, Li, & Tarn, 2008); and, one research points out that classical RL can be boosted by $O(\sqrt{N})$ (Dunjko, Taylor, & Briegel, 2016), depicting a quadratic acceleration. While, these potentials are not far-future, it must also be noted that there are a variety of problems with quantum computation, like the fragility of quantum states and hardware issues. For now, only costly lab environments can provide the required conditions and issues like these should be dealt with before being able to use them in everyday robotics.

As a summary, RL is a method making use of stochastic control processes to form a model of decision-making for AI agents. For several reasons, it can be seen similar to the way humans process and evaluate decisions. This is one of the main reasons this method should be implemented to robotics research. Another reason is related with the promises this method holds for the future. By using quantum computers and appropriate hardware, in a couple of decades, one will be able to build quantum robots; which are essentially made up of many quantum systems(Dong, Chen, Zhang, & Chen, 2006). These robots will be able to run on Quantum RL software as to perform significantly complex computations and have proper interactions with their environment. It can also encourage future researches such as projective simulation for AI, which allows the agent to project itself into future situations before it takes real action(Briegel & Cuevas, 2012); just like human mental time travel evolved over time.

References

Arias-Carrión, O., Stamelou, M., Murillo-Rodríguez, E., Menéndez-González, M., & Pöppel, E. (2010). Dopaminergic reward system: a short integrative review. *International Archives of Medicine*, *3*(1), 24. doi: 10.1186/1755-7682-3-24

Briegel, H. J., & Cuevas, G. D. L. (2012). Projective simulation for artificial intelligence. *Scientific Reports*, *2*(1). doi: 10.1038/srep00400

Dong, D., Chen, C., Li, H., & Tarn, T. J. (2008). Quantum reinforcement learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*. doi: 10.1109/TSMCB.2008.925743

Dong, D., Chen, C., Zhang, C., & Chen, Z. (2006). Quantum robot: Structure, algorithms and applications. *Robotica*. doi: 10.1017/S0263574705002596

Dunjko, V., Taylor, J. M., & Briegel, H. J. (2016). Quantum-Enhanced Machine Learning. *Physical Review Letters*. doi: 10.1103/PhysRevLett.117.130501

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *CoRR*, *cs.AI/9605103*. Retrieved from `http://arxiv.org/abs/cs.AI/9605103`

Petschnigg, C., Brandstotter, M., Pichler, H., Hofbaur, M., & Dieber, B. (2019). Quantum computation in robotic science and applications. *2019 International Conference on Robotics and Automation (ICRA)*. doi: 10.1109/icra.2019.8793768

Schultz, W. (2015). Neuronal reward and decision signals: From theories to data. *Physiological Reviews*, *95*(3), 853–951. doi: 10.1152/physrev.00023.2014

VandenBos, G. R. (n.d.). *Apa dictionary of psychology.* Retrieved from `https://dictionary.apa.org/decision-making`