# Reinforcement Learning Essay

## Introduction

Reinforcement Learning (RL) is an area within Machine Learning, concerned with creating *intelligent agents* in the form of algorithms with the capacity to learn through trial and error with minimal human intervention (Ravichandiran, 2018). Practically RL seeks to mimic biology by training algorithms with reward and/or punishment to inform the decision making of the agent for continued learning. This is much like how internal neurological processes teaches humans and animals to avoid things that causes negative emotions and seek things that causes positive emotions (Géron, 2019). Thus, where for supervised machine learning seeks to teach through *examples*, RL seeks to teach through *experience.*

## Reinforcement Learning Fundamental Principles

As with any area within Machine Learning, many variations of RL algorithms has been developed since the research area’s beginning in the 1980’s (Sutton & Barton, 2018). The basics however is fairly straight forward. The agent receives information about its current state and environment (the input), then based on a *policy function* decides upon an action that will bring it to a new state. The agent decides on the best course of action by maximizing a reward function, an accumulated score calculated by evaluating predicted *states* or *state-action pairs* over time, rewarding or punishing certain states and/or actions based on a custom definition (Ravichandiran, 2018*).* When training reinforcement learning models a non-deterministic environment is often assumed, meaning the same action in the same state might not always result in the same outcome (Sutton & Barton, 2018).

An important aspect of reinforcement learning is its adaptability to various problem (Friedman, 2019). With a general policy/reward function, algorithms can solve many different problems with minimal human intervention. By running many iterations, the agent learns more of its environment and the rules governing it, resulting in incrementally better predictions about which actions will bring the most rewards over time. This general approach to problem solving gives the model a high utility in solving a variety of problems. Famously a model has been able to teach itself over 500 old Atari games (Hausknecht et. Al., 2014), (Kaiser et. Al. 2019). Another beat the human champion at the Asian game Go (Chen et. Al., 2018), a task previously thought a near impossible task for a computer.

## Use

Reinforcement learning is seeing more and more commercial applications spanning many different areas. These include resource management in computer systems, advertising, gaming, traffic light controls and more (Sutton & Barto, 2018). Lorica (2017), boils the range of applications down to three main themes: Optimizing (fx supply chain or yield management), controlling (fx autonomous vehicles, turbine control) or monitoring and maintaining (fx quality control and inventory monitoring). A shared quality of these applications is that they “supply an environment, an action, a response and a method for optimization” (Lorica, 2017) allowing reinforcement learning to provide valuable insights and improve over time.

## Limitations and critique

As with any AI, deep reinforcement learning comes with challenges and risks that need to be addressed before the technology will be safe to use (Friedman, 2019). Machine learning and neural net algorithms can be black boxes of incomprehensible logic after they’ve been trained. This makes it hard to conclude exactly how an algorithm arrives at its conclusions, which poses moral and ethical challenges (Géron, 2019). It can be impossible to predict if the algorithm will act appropriately to novel situations. An example of this is the problem of self driving cars prioritizing whose life to safe in life or death situations.

This challenge is even more pronounced in deep reinforcement learning, as one of the goals of the technology is to enable the intelligent agent to face as wide an array of challenges as possible. As the agent seeks to maximize its reward function, it can easily choose solutions non-sensical or even immoral to a human observer (Friedman, 2019). Thus extraordinary care has to be taken when implementing solutions involving DRL, complicated by the fact that tailoring reward or policy functions to specific tasks to limit risks, reduces the agents ability to function as a generalist problem solver or introduce even more undesirable avenues for the agent to exploit its reward/policy functions (Ravichandiran, 2018)

Another limitation of DRL is that even simple models can require an extraordinary amount of training before reaching desirable results. Where humans only need to try simple tasks a couple of times to pass them acceptable, our machine counterparts might need millions of test runs to produce an acceptable outcome, making training of models exceptionally time consuming (Sutton & Barto, 2018)

*Friedman, L., 2019 ‘MIT 6.S091 Introduction to Deep Reinforcement Learning (Deep RL)’ , Youtube*

*<* *https://www.youtube.com/watch?v=zR11FLZ-O9M&t=1812s >*

*Ravichandiran, S 2018, Hands-On Reinforcement Learning with Python : Master Reinforcement and Deep Reinforcement Learning Using OpenAI Gym and TensorFlow, Packt Publishing, Limited, Birmingham. Available from: ProQuest Ebook Central. [5 April 2022].*

*Sutton, R. & Barto, A. 2018, Reinforcement Learning: An Introduction, Second edn., ISBN: 9780262039246, The MIT Press, Cambridge, Massachusetts, <*[*https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf*](https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf)*>.*

*Géron, A. 2019, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, Second Edition, ISBN: 9781492032649, O’Reilly Media, Inc.*

*Lorica, B. 2017, Practical Applications of Reinforcement Learning in Industry: An Overview of Commercial and Industrial Applications of Reinforcement Learning, viewed 31/May/2020, <*[*https://www.oreilly.com/radar/practical-applications-of-reinforcement-learning-in-industry/*](https://www.oreilly.com/radar/practical-applications-of-reinforcement-learning-in-industry/)*>.*

*Chen, Y., Huang, A., Wang, Z., Antonoglou, I., Schrittwieser, J. & Silver, D. 2018, ‘Bayesian Optimization in Alphago’, arXiv.org, ISSN: 2331–8422, <*[*https://arxiv.org/pdf/1812.06855.pdf*](https://arxiv.org/pdf/1812.06855.pdf)*>*

*Garychl 2018, Applications of Reinforcement Learning in Real World, viewed 31/May/2020, <*[*https://towardsdatascience.com/applications-of-reinforcement-learning-in-real-world-1a94955bcd12*](https://towardsdatascience.com/applications-of-reinforcement-learning-in-real-world-1a94955bcd12)*>.*

*Kaiser, L., Babaeizadeh, M., Milos, P., Osinski, B., Campbell, R., Czechowski, K., Erhan, D., Finn, C., Kozakowski, P., Levine, S., Mohiuddin, A., Sepassi, R., Tucker, G. & Michalewski, H. 2019, ‘Model-Based Reinforcement Learning for Atari’, in ICLR, arXiv.org,*[*https://arxiv.org/pdf/1903.00374.pdf*](https://arxiv.org/pdf/1903.00374.pdf)*>.*