

## PLAGIARISM SCAN REPORT

Words 822 Date February 02,2020

Characters 5173 Exclude Url

0%

Plagiarism

100%

Unique

0

Plagiarized  
Sentences

45

Unique Sentences

## Content Checked For Plagiarism

Reinforcement Learning Agents The attainment of knowledge and wisdom is a continuous process for humanity. It commences at infancy and continues till the end of life. A child learns by observing the actions of those around him. Adults acclimatize themselves with new cultures by interaction. Such behavioral learning uses reinforcement procedures to communicate notions of 'right and 'wrong' – analogous to what might be good and bad demeanor for a child, or culturally appropriate and culturally insensitive quips for an adult. These learning techniques form the underlying theory of reinforcement learning. With the aim to maximize some form of cumulative reward, reinforcement learning is an approach to artificial intelligence that emphasizes decision making based on past experiences. It is one of the three basic machine learning paradigms, alongside supervised learning and unsupervised learning. Its key features include trial and error, and delayed gratification, which focus on finding a balance between utilization of knowledge and assessment of raw data. Such a problem solving approach is especially useful in the world of video gaming. Consider a classic game of Pac-Man. The titular character, Pac-Man, attempts to eat the maximum amount of Pac-Dots before being eaten by the ghosts (opponents). Near the corners of the labyrinth are four energizers that allow Pac-Man to eat the ghosts and receive bonus points. In this scenario, the enemies turn deep blue and move away from Pac-Man. So while eating a blue ghost rewards bonus points, it increases the risk as well, since the period for which the ghost will remain blue (and harmless) is unknown. Consequently, when devising an optimal strategy to play Pac-Man, the reward earned for eating a ghost is discounted. A discount rate called gamma is defined, having a value between 0 and 1. It is inversely proportional to the reduction of reward. A large gamma implies that the learning agent is more concerned about the long term reward. On the other hand, a small gamma value means the agent has greater interest in short term rewards. Trial and error calculations, based on simulations, are then used to perfect the value of gamma. A Closer Look at Reinforcement Learning for Video Games DeepMind, a UK based AI company, has designed a learning agent which can recreate human-level performance in multi-player video games based on the 'Capture the Flag' model. The game begins with a team of players on either side of a playground. Both teams must protect their flag from the enemy. The first team to steal the flag of their opponent and bring it back to their base wins the game. If a player gets tagged by the opposition, they are out of the game. Using a digital version of the same, scientists at DeepMind have trained an entire population of agents, who learn by interacting with a three-dimensional environment via first-person perspective. There is no centralized entity and every player acts independently based on their own observations. Researchers have trained teams of up to 30 agents in parallel, all playing against each other thousands of times. Initially, all agents wander aimlessly, until one of them discovers some useful information and begins to take control of the flag and scoring points. At that point, there is an evolutionary pressure on the population. The agents continue playing till they achieve satisfactory performance. Genetic algorithms are used to ensure that the entire population evolves, and weaker agents are removed. Furthermore, the original group of trained agents is 'bred' together to produce 'child' agents, and as the number of generations increase, only the strongest traits persist. Unlike human children, AI agents inherit all knowledge their 'parents' possess. Getting to the subtleties of how the learning process actually takes place, consider a typical simulation. An agent plays for five minutes and in the process, performs thousands of actions. Scientists then need to devise a way to associate, the analysis done by the agent about their environment, with the rules of the game. So, internal rewards may be linked with events such as retrieving the enemy flag, or a teammate tagging an opponent. Also, agents are allowed to independently evolve the rewards assigned to each event, which helps delegate responsibility within the team. AI agents trained in this manner can play with artificial as well as human teammates. They offer a certain advantage over human players, in that they are free from any human bias and only focus on winning the game. Conclusion While reinforcement learning is no novel innovation, it is certainly years away from reaching its full potential. A decade long trend within artificial intelligence and machine learning, it forms the core of

from reaching its full potential. A decade-long trend within artificial intelligence and machine learning, it forms the core of many computational processes which mimic the biological learning system. It has provided psychological models of animal learning that match empirical data better than traditional ones, and has helped map an influential model of the brain reward system. And if the past is any evidence, future developments in this area will certainly be worth anticipating.

Sources	Similarity
---------	------------