

## PLAGIARISM SCAN REPORT

Words 999 Date January 28,2020

Characters 5771 Exclude Url

8%

Plagiarism

92%

Unique

4

Plagiarized  
Sentences

47

Unique Sentences

## Content Checked For Plagiarism

The Reinforcement learning agents During our childhood, we learn by interacting with our environment. When an infant watches their parent and learns by observing every action taken by people around, such interaction are undoubtedly a major source of knowledge about their environment and themselves. The idea of behavior learning is an underlying theory of reinforcement learning. Reinforcement Learning is an approach to natural and artificial intelligence that emphasizes learning and planning from sample experience. The characteristics trial and error, searching and delayed reward are the two important distinguishing features of reinforcement learning. In reinforcement learning, we using ideas of dynamic system theory. The world of the game provides the perfect mini-universe in which our AI agent facing a real problem and achieves its goals extend AI agent must be able to take action that affects the state. The agents have a goal relating to the state of the environment. Let say in the game of Pac-man your agent is this Pac-man and your opponent is the ghosts. Your goal is to eat the maximum amount of Pac dots before being eaten by ghosts. As we know for certain period enemies turn deep blue, move away from Pac-man and bonus reward earned for eating blue ghost but it is dangerous because the ghost turns blue for a certain period. As a consequence, the reward earned for eating a ghost, even if it is bigger will be discounted. We are not sure we will be able to eat it. To discount the reward we proceed like this we define a discount rate called gamma. It must be between 0 and 1. \* The larger the gamma, the smaller the discount. This means the learning agent cares more about the long term reward. \* On the other hand the smaller the gamma the bigger the discount. This means our agent cares more about the short term reward. Our discounted total reward is To be simple, each reward will be discounted by gamma to an exponent of the time step. As the time step increases, ghosts get closer to us, so the further reward is led and less probable to happen. Using the new development in reinforcement learning deep mind design a game capture the flag. The rules are the first team to steal the flag of their opponent and bring it back to their base. If you get tagged by the opponent and bring it back to their base. If you get tagged by the opposition then you are out of the game. Deepmind research scientists train the whole population of AI agents into a digital version of the game. In this game AI agent sees their first-person point of view so, the AI agent has to look around and move through this 3-D world from the first-person perspective and interact with these other things. Here there is no centralized entity, every player act independently only thing they analyze is their observation. Research scientists at deep mind train the whole population of teammates, the entire classroom of 30 agents in parallel they are all playing within and against each other around 1000 of 1000 times each agent can learn from their own experience. When they start AI agents were bouncing around the place without a clue and one of them will discover something and will start taking control of the flag and scoring points. At that point, there is an evolutionary pressure on this population. Deep mind research scientists not only letting AI agents play on and on forever they are also using a genetic algorithm the way to make sure the whole culture of the population evolves. Some of the weaker one was removed from the population. The original classroom of agents breeds together ever and has kids of their own and as you go down the generation the strongest trait survive. Unlike human children, AI agents will inherit the whole knowledge been to gain from their parents. For example, AI agents play 5 minutes and do 1000of actions you just have to win or lose. Somehow deep mind research scientist has to learn what to do with the analysis done by AI agent about their environment. So the idea of internal reward were events in the game such as picking up a flag or dropping a flag or your teammate tagging an opponent or an opponent tagging you and all these sorts of things. Also, research scientist allows the agents to individually evolve their internal reward which is the reward they assign to each event due to this some agents care a lot about holding on the flag and some agents care about teammate tagging, someone. But after some round of games, the AI agent team is getting good at the game. Training this way these agents have an advantage they can play within themself and more interestingly they can play with people. So we can drop people into these games and have an AI team mate. When people play with the AI agents some time agents completely ignore

people into these games and have an AI team mate. when people play with the AI agents some time agents completely ignore human players treat them that they completely use less. People pay a lot of attention to others even in-game scenarios human will fixate with other players but these agents are trained completely unbiased without these human biases. In some games, your AI opponent will run passed near you and not even try to tag you because they are so fixated on getting the flag so quickly as possible because that going to maximize their number of capture and win the game. Reinforcement learning is a part of a decade-long trend within artificial intelligence and Machine Learning. Reinforcement learning is the closest to the kind of learning that humans do and many of the core algorithms of reinforcement learning were originally inspired by the biological learning system. Reinforcement learning has also given back both through, a psychological model of animal learning that better matches some of the empirical data and though an influential model of part of the brain reward system.

Sources	Similarity
<p>Ch 12:Reinforcement learning Complete Guide #towardsAGICompare text</p> <p>the larger the gamma, the smaller the discount (so we get decent future rewards). this means the agent focuses more about the long term reward.value function = immediate reward + value function of next state (ignore gamma). this is called bellman equation.</p> <p><a href="https://medium.com/deep-math-machine-learning-ai/ch-12-reinforcement-learning-complete-guide-towardsagi-ceeda325c5d53">https://medium.com/deep-math-machine-learning-ai/ch-12-reinforcement-learning-complete-guide-towardsagi-ceeda325c5d53</a></p>	10%
<p>An introduction to Reinforcement LearningCompare text</p> <p>this means our agent cares more about the short term reward (the nearest cheese). our discounted cumulative expected rewards isby running more and more episodes, the agent will learn to play better and better. temporal difference learning : learning at each time step.</p> <p><a href="https://www.freecodecamp.org/news/an-introduction-to-reinforcement-learning-4339519de419/">https://www.freecodecamp.org/news/an-introduction-to-reinforcement-learning-4339519de419/</a></p>	5%
<p>1. Introduction to Reinforcement LearningCompare text</p> <p>environment gives some reward <math>r_1</math> to the rl agent. it probably gives a +1 because the agent is not dead yet. this rl loop continues until we are dead or we reach our destination, and it continuously outputs a sequence ofthis means the learning agent cares more about the long term reward.</p> <p><a href="https://www.nishkamipatel.com/post/reinforcement-learning-1">https://www.nishkamipatel.com/post/reinforcement-learning-1</a></p>	5%