Sammabine Admettre

CS370

6-2 Assignment:

Cart pole Revisited

Using improved algorithms, you can solve the cart pole issue. In control theory, the cart pole is a well-studied device (Cart Pole, n.d.). The objective is to get an upside-down pendant mounted on a cart pole to balance upright and vertically. This is a difficult task since the straight fixed point is inherently shaky (Cart Pole, n.d.). Reinforcement learning algorithms are known as policy gradient algorithms, and they are a subset of reinforcement algorithms. The reinforcement algorithm requires both an input state and an output state to function (Cart Pole, n.d.). The artificial intelligence manages the output state.

In the center, there is just one function that accepts the state as input and returns the action as output. There is nothing else within the heart. When these two jobs or activities are accomplished, the event is said to be a program. This plan allows the agent to see what actions will occur in either the input or output states. When you look at the policy, it's almost like a roadmap. The cart pole challenge is an open AI job for training and testing algorithms in a laboratory setting (Cart Pole, n.d.). In the case of a cart pole, the cart is the driver. Following that, it was governed by two occurrences, POSTIVE and NEGATIVE, which moved to the left and right in their respective paths (Cart Pole, n.d.). Because the pole is standing straight, a one-point reward will be awarded. Making sure the pole does not fall to either side is critical for achieving the aim.

The policy gradient is distinct from value-based techniques. Policy gradient is a sort of reinforcement learning in which parameterized policies are used to estimate the anticipated return via a process known as gradient descent (Singh, n.d.). When you adopt a value-based strategy, such as Q-learning, you learn by completing one of a large number of activities in order to get the best outcomes (Singh, n.d.). Finding the highest payout for an action Q-learning is the same as determining the likelihood that the action will be performed (Singh, n.d.). When compared side by side, the two methodologies seem to be identical and result in the same policy. It improves its performance based on the same factors by using what it has learned from its surroundings or the comments it has received (Singh, n.d.). We will be talking about reinforcement learning in this article, focusing on deep Q-learning, which is a well-known method that is used a lot in RL (Singh, n.d.).

Reference

*Cart Pole*. (n.d.). Ethanweber.me. Retrieved June 17, 2024, from <https://ethanweber.me/cartpole.html#:~:text=The%20cart%20pole%20is%20a>

Singh, S. (n.d.). *A Comprehensive Guide to Neural Networks in Deep Q-learning*. Www.turing.com. https://www.turing.com/kb/how-are-neural-networks-used-in-deep-q-learning

‌

‌