Creating Adaptive Agents through Interaction

Ashley Edwards    Himanshu Sahni

Abstract
Software and user interfaces are designed with the preferences of the end-users in mind. However, these modules are not easily modified. Users often think of changes to products long before they are implemented, but changing these products requires constant updates and resources.

We propose a method for creating an agent that is capable of learning and adapting from interactions with an end user. In our approach, we use Interactive Reinforcement Learning to train a bipedal agent to learn how to run and remain balanced. This will give the user a unique experience with the game, while allowing the agent to adapt to the users’ preferences.

Intro
• Reinforcement Learning is a machine learning paradigm where an agent interacts with its environment and modifies its actions to receive an optimal payoff

Approach
In our approach, we implemented a Reinforcement Learning agent that aimed to learn how to play the game QWOP. In this game, users control the calf and thigh muscles of an agent using the keyboard letters Q, W, O, and P. An optimal configuration of these actions allows the agent to run and remain balanced.

Simulation
• We created a simulation of QWOP using JBox2D and placed it online
• The agent had four available actions: Q and W for controlling the left and right thighs and O and P for controlling the calves

Reinforcement Learning
We had two modes for training our Reinforcement Learning agent:
• Sarsa Learning (Critic)
  • During this phase, the agent received a positive reward for remaining balanced and walking forward and a negative reward for falling
  • Users of the system could act as the Critic by giving the agent a positive or negative reward for their actions
  • Rewards from users were weighed more heavily than those from the environment
• Interactive Learning (Actor)
  • During this phase, the users became the Actor and took over as the agent
  • The agent continues to learn during this phase, however, it is now learning based on the user’s preferences
  • This gives each user a unique experience with the agent

Simulation
• We created a simulation of QWOP using JBox2D and placed it online
• The agent had four available actions: Q and W for controlling the left and right thighs and O and P for controlling the calves

Reinforcement Learning
We had two modes for training our Reinforcement Learning agent:
• Sarsa Learning (Critic)
  • During this phase, the agent received a positive reward for remaining balanced and walking forward and a negative reward for falling
  • Users of the system could act as the Critic by giving the agent a positive or negative reward for their actions
  • Rewards from users were weighed more heavily than those from the environment
• Interactive Learning (Actor)
  • During this phase, the users became the Actor and took over as the agent
  • The agent continues to learn during this phase, however, it is now learning based on the user’s preferences
  • This gives each user a unique experience with the agent

Conclusion and Future Work
In conclusion, we have created an agent that is able to learn actions by interacting with an end-user. This gives the user a unique experience with the game, while allowing the agent to adapt to the users’ preferences.

We are currently testing to see if the agent learns the task of running better with interaction or without it. In the future, we will allow multiple users to train the same agent. This will likely lead to a faster learning experience. However, the agent’s performance will be biased by the current user of the system.

References: