

Práctica 3 - Competición entre agentes para el juego de Atari Go

Equipo CyBerry:

- Ascencio Rangel Luis Eduardo
- Mendoza Medina Alejandro
- González Cano Daphne Sofía

Estrategia del agente Tron.

Tron utiliza el algoritmo de Q learning con redes de aprendizaje prunde (Deep Q Learning) para aprender a jugar a AtariGo. Dada la naturaleza compleja del juego, la forma de exploración (mediante ϵ -greedy) de Tron es una combinación entre movimientos aleatorios, así como del uso de reglas heurísticas dependiendo de si se es ofensivo o defensivo.

De ser ofensivo:

1. Tron analiza todo el tablero y, tomando piedras enemigas como aliadas, busca aquellas piedras que tengan el menor número de libertades (intersecciones vacías), y en base a ello realiza su movimiento

De ser defensivo:

1. Tron analiza todo el tablero y en base a todas sus piedras, busca ampliar el número de libertades (acrecentando el número de piedras juntas, haciendo un grupo aliado) de aquellas que tengan entre 1 y 2 libertades. Si es un grupo de piedras, entonces el total de libertades es la suma de sus libertades individuales.
2. Tron, si no hay movimientos defensivos ni ofensivos, entonces buscará ampliar el número de libertades de cualquier grupo de piedras suyas.

De la misma manera, la función de recompensa utilizada para determinar la recompensa de cada estado está dada por las reglas heurísticas; es decir que, si tras un movimiento se llega a que coincide con alguno de los puntos especificados por la heurística, entonces se recompensa en base a ello.

Para el entrenamiento del agente, se especificó un valor de ϵ mínimo de 0.3, pero para la competencia se cambió a 0.7.

I fight for the users!

-Tron