

Aprendizaje Automatico

Trabajo Práctico 2

QLearning

20 de noviembre de 2016

Integrante	LU	Correo electrónico	Carrera
Martin Baigorria	575/14	martinbaigorria@gmail.com	computación (licenciatura)
Damián Furman	936/11	damian.a.furman@gmail.com	computación (licenciatura)
Germán Abrevaya	-	germanabrevaya@gmail.com	física (doctorado)

Reservado para la cátedra

Instancia	Docente	Nota
Primera entrega		
Segunda entrega		

Índice

1. Introducción	3
-----------------	---

1. Introducción

Hablar un poco de QLearning y cadenas de Markov.

Hablar un poco sobre la cantidad de estados posibles que hay en 4 en linea. La cadena de Markov asociada tiene 1 estado por posible configuracion del tablero, y una transicion por cada accion posible a otro estado.

¿Cuál es el espacio de estados? ¿cuán rápido se puede explorar? ¿cómo cambia la inicialización de Q con respecto a la velocidad de aprendizaje? ¿qué importancia tiene la temperatura y la velocidad con que se enfría el sistema si se decide usar la distribution Boltzmann? ¿qué efecto tiene cambiar la tasa de aprendizaje?

Implementar el alpha dinamico, a medida que pasa el tiempo alpha cae. Faltan cosas, pero la base del codigo esta.

Jugar con todos los parametros de los players, y el tablero.