

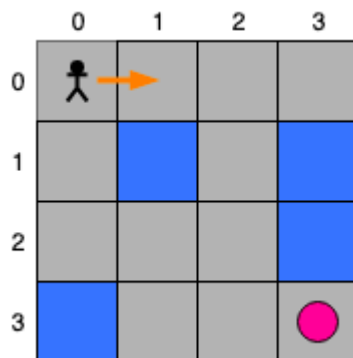
Inteligență Artificială

Laborator 9 și 10. Reinforcement Learning

Temă

Considerăm un agent care se poate deplasa într-un mediu (un grid de dimensiuni $n \times n$). Agentul se poate deplasa în direcția sus, jos, stânga sau dreapta. Agentul poate merge pe gheață. În unele locuri, gheața este subțire și se poate sparge. Dacă agentul ajunge într-o astfel de locație, atunci acesta moare.

Spre exemplu, putem avea următoarea configurație (4x4, iar pătratele albastre reprezintă gheața care se poate sparge):



Având un punct de start, scopul agentului este să ajungă la destinație. În exemplul de mai sus, punctul de start este pătratul stânga-sus, iar destinația este pătratul dreapta-jos. Când agentul ajunge la destinație, recompensa este egală cu 1, altfel este 0. Mediul este stochastic.

I. Implementarea algoritmului *Q-learning*

- (0.1p) inițializarea tabelului Q, a parametrilor algoritmului și a stării inițiale
- (0.1p) pentru o stare s , identifică starea următoare s' prin aplicarea unei acțiuni a
- (0.8p) implementează algoritmul Q-learning pentru a identifica drumul pe care trebuie să-l parcurgă agentul
 - selectează acțiunea cu cea mai mare valoare Q din starea s'
 - actualizează valorile Q
 - actualizează starea curentă
 - repetă

II. (0.6p) Utilizarea unei rețele neuronale pentru a aproxima funcția Q / Implementarea algoritmului Deep Q-learning

Observație: se poate folosi mediul Open AI gym <https://gym.openai.com/>

Pentru săptămâna 6-12 decembrie: punctele I.a, I.b

Pentru săptămâna 13-19 decembrie: punctul I.c

Pentru săptămâna 3-9 ianuarie: punctul II

Documentație:

Secțiunea 6.5 Q-learning: Off-policy TD Control
<http://incompleteideas.net/book/ebook/node65.html>