

# Učenie sa podporškom

Reinforcement Learning

# Učenje sa podrškom

- Optimizacija
- Odložene posledice (Delayed consequences)
- Istraživanje
- Generalizacija

# Optimizacija

- Cilj je da se pronađe optimalan način za donošenje odluka
  - Davanje najboljih ishoda ili bar veoma dobrih ishoda
- Eksplicitna notacija za upotrebu odluka
- Primer: pronalaženje minimalne razdaljine između dva data grada ako je poznata mreža puteva

# Odložene posledice (Delayed consequences)

- Odluke sada mogu da utiču na rezultat mnogo kasnije ...
  - Štednja za penziju
  - Pronalaženje knjige ili ključa u računarskoj igri
- Uvodi dva izazova
  - Prilikom planiranja: odluke podrazumevaju rezonovanje o ne samo neposrednoj koristi od odluke već i njenim dugoročnim posledicama
  - Prilikom učenja: kreditni zadatak je težak (šta prouzrokuje kasnije visoke ili niske nagrade?)

# Istraživanje

- Učenje o svetu donošenjem odluka
  - Agent kao naučnik
  - Naučite da vozite bicikl tako što ćete pokušati (i ne uspeti)
  - Pronalaženje knjige ili ključa u igrici
- Cenzurisani podaci
  - Dobijate samo nagradu (labelu) za donetu odluku
  - Ne znam šta bi se desilo da smo uzeli crvenu pilulu umesto plave pilule (Matriks)
- Odluke utiču na ono o čemu učimo
  - Ako odlučimo da idemo na ETF umesto na FON, dobićemo drugačija kasnija iskustva...

# Generalizacija

- Smernice mapiraju prošla iskustava u akcije
- Zašto jednostavno definišemo ranije smernice?
- Atari igrice –DeepMind rešenje
- $(256^{100 \times 200})^3$  mogućih slika

# Postupak

- Istražiti svet
- Korišćenje iskustva za donošenje budućih odluka

# Dodatni problemi

- Odakle dolaze nagrade?
  - A šta će se desiti ako pogrešimo?
- Robustnost / Osetljivost na rizik
- Nismo sami...
  - RL sa više agenata