

Veštačka inteligencija i igre: rezime

Andrej Ivašković

NEDELJA INFORMATIKE, 30. mart 2015.

1 Agenti i okruženja

- **Agent** je bilo koji uređaj koji reaguje na nadražaje koje prima od **okruženja**.
- Okruženje može da bude:
 - **determinističko**: ako je okruženje u stanju s , a agent izvrši akciju a , tada postoji i jedinstveno je određeno naredno stanje $\mathcal{S}(s, a)$;
 - **stohastičko**.

2 Igre sa pretragom

- U igrama sa malim brojem stanja (npr. iks-oks) je dovoljno napraviti stablo igre, pri čemu ono može da se obiđe ili metodama za obilazak stabla (*post-order*) ili obilaskom grafa (*pretraga po dubini*).
- **Minimax algoritam** se primenjuje u igrama u kojima igrači treba da postignu izvesnu prednost (npr. poeni). Dva igrača, Max i Min, biraju poteze na sledeći način:
 - Max će odabrati onaj potez takav da ostvari najveću razliku pri optimalnoj igri Min-a;
 - Min će odabrati onaj potez takav da ostvari najmanju razliku pri optimalnoj igri Max-a.
- Ukoliko nije izvodljivo napraviti celo stablo, uvodi se ograničenje za dubinu stabla koje se pravi (broj poteza unapred koji razmatramo) i koristi se **funkcija procene**. Ovo se primenjuje u npr. šahu.

- **Algoritam α - β reza** se koristi za otpisivanje delova stabla koji ne vode ka traženim ekstremima u minimax algoritmu. Zasniva se na pretrazi po dubini, ali se u svakom trenutku pamte α (najveća vrednost koju može da postigne Max) i β (najmanja vrednost koju može da postigne Min).
- Rešiti go je aktuelan zadatak!

3 Q -učenje

- Uvodi se **funkcija nagrade** $\mathcal{R} : S \times A \rightarrow \mathbb{R}$.
- Tražimo **funkciju polise** $p : S \rightarrow A$ koja određuje akciju koja će se sprovesti u slučaju da je okruženje u nekom stanju.
- Uvodimo ocenu za neku datu polisu p . Ako je startno stanje s'_1 , a $s'_k = \mathcal{S}(s'_{k-1}, p(s'_{k-1}))$, pa onda definišemo i $r_k = \mathcal{R}(s'_k, p(s'_k))$, za neko odabrano γ imamo ocenu polise:

$$V^p(s) = \sum_{k=1}^{\infty} \gamma^{k-1} r_k$$

- Za određivanje optimalne polise je dovoljno naći $\max_{\alpha} \{Q(s, \alpha)\}$, gde je:

$$Q(s, a) = \mathcal{R}(s, a) + \gamma V_{\text{opt}}(\mathcal{S}(s, a))$$

- Određivanje Q se vrši tako što najpre imamo procenu Q' koja je na početku tabela popunjena nasumično odabranim vrednostima, nakon čega:
 1. Trenutno stanje je s , isprobaj neku akciju a (odabir a je magija verovatnoće!).
 2. Uradi a , sleduje nagrada $\mathcal{R}(s, a)$.
 3. Posmatraj novo stanje $\mathcal{S}(s, a)$.
 4. Ažuriranje:

$$Q'(s, a) \leftarrow \mathcal{R}(s, a) + \gamma \max_{\alpha} \{Q'(s, a), \alpha\}$$

5. Idi na 1.

4 Relevantni linkovi

- xkcd.com/832/
- xkcd.com/1002/
- www.youtube.com/watch?v=YgYSv2KSyWg