

Požičovňa áut

Jakub spravuje činnosť dvoch pobočiek siete spoločnosti zaoberajúcej sa požičiavaním áut. Každý deň na každú pobočku zavíta niekoľko zákazníkov. Ak pobočka má voľné auto k dispozícii, tak ho požičia a Jakub získa kredit 10 EUR za sprostredkovanie požičania. Ak pobočka práve nemá voľné auto, tak sa požičanie nekoná a príležitosť na zárobok je stratená. Vrátené autá sa stávajú voľnými až nasledujúci deň po svojom vrátení.

Počet žiadaných aj počet vrátených áut má stochastický charakter – riadi sa Poissonovou diskretnou distribúciou pravdepodobnosti, kde pravdepodobnosť pre nejaké n je daná

$$\frac{z^n}{n!} \cdot e^{-z}$$

kde z je očakávaná hodnota. Pre prvú pobočku je $z=3$ pre počet záujemcov o požičanie a $z=3$ pre počet vrátených áut. Pre druhú pobočku je $z=4$ pre počet záujemcov o požičanie a $z=2$ pre počet vrátených áut.

Ak si zákazník požičia auto v niektorej Jakubovej pobočke, vrátiť ho môže v hociktovej pobočke danej siete, nielen v pobočke spravovanej Jakubom. A naopak, zákazník môže vrátiť v Jakubovej pobočke aj auto, ktoré si požičal v pobočke spravovanej niekým iným.

Každá pobočka má obmedzenú kapacitu – dokáže uskladniť maximálne 20 áut. Ak by sa v nejakej pobočke nahromadilo viac áut, prebytočné autá sú odsunuté do siete a strácajú sa z Jakubovho dohľadu.

Pre zabezpečovanie dostatku áut na pobočkách, Jakub môže v noci prevážať autá medzi pobočkami a tým zvýšiť počet áut na niektorej z pobočiek, pričom počet áut na druhej z pobočiek sa zodpovedajúco zníži. Každý takýto prevoz ho však stojí 2 EUR. Za noc takto môže presunúť maximálne 5 áut.

Sekvenčný rozhodovací problém

Systém je v nejakom **počiatočnom** stave S_0 . Agent prijme rozhodnutie a vykoná **akciu** A_0 . Následkom toho agent získa **odmenu** R_1 a systém sa preklolí do stavu S_1 . Agent prijme rozhodnutie a vykoná akciu A_1 . Následkom toho agent získa odmenu R_2 a systém sa preklolí do stavu S_2 . A to pokračuje určitú dobu až sa systém preklolí do terminálneho stavu S_T .

Agent vykonáva výber akcií A_0, A_1, \dots, A_{T-1} pričom výber akcie A_t nezávisí na čase t ale iba na tom, v akom stave S_t bol systém v čase t (teda v čase výberu jednej z možných realizovateľných akcií). Akcie vyberá na základe nejakej svojej **politiky**. Jeden spôsob vyjadrenia agentovej politiky je založený na existencii tzv. **hodnotovej funkcie** Q ktorá je zobrazením

$$|S| \times |A| \rightarrow \mathcal{R}$$

kde S je množina všetkých stavov a A je množina všetkých akcií. Táto funkcia teda vlastne ohodnocuje každú dvojicu (stav, akcia) nejakou číselnou hodnotou. Ak je systém aktuálne v nejakom stave S_i , tak agent vyberá akciu

$$a = \operatorname{argmax}_{j=1, \dots, |A|} Q(S_i, a_j)$$

teda tú akciu, ktorá v kombinácii s daným stavom dosahuje najvyššiu hodnotu hodnotovej funkcie.

Za svoj výber akcií podľa nejakej politiky agent získava odmeny (alebo pokuty ak sa jedná o záporné hodnoty) R_1, R_2, \dots, R_T . Ich suma

$$\sum_{i=1}^T R_i$$

sa označuje ako **kumulatívna odmena** a vlastne ohodnocuje agentovu politiku. Politika je tým úspešnejšia, čím agent získava vyššie odmeny a teda čím získa lepšiu kumulatívnu odmenu.

Úloha

Vašou úlohou je nájsť pre Jakuba takú politiku nočného presúvania áut medzi oboma jeho prevádzkami, ktorá mu zabezpečí čo najväčší zárobok. Predpokladajte, že so svojimi prevádzkami začne pôsobiť na začiatku roku a berte do úvahy obdobie ročného pôsobenia (vrátane víkendov a sviatkov).

Vstupmi sú štyri rôzne počiatočné distribúcie áut v prevádzkach pri zahájení činnosti:

- po 10 áut v každej z prevádzok
- 5 áut v prvej prevádzke a 15 áut v druhej prevádzke
- 15 áut v prvej prevádzke a 5 áut v druhej prevádzke
- obe prevádzky prázdne

Ako svoje riešenie dodajte

- súbor so zdrojovým textom vášho algoritmu (predpokladá sa, že všetky štyri problémy budete riešiť tým istým algoritmom)

- súbor typu *.pdf s popisom riešenia (ako a prečo ste volili reprezentáciu, určovanie vhodnosti, generovanie iniciálnej populácie, genetické operátory, selekciu, náhradu, návrh hodnôt parametrov, ukončenie behu algoritmu), dosiahnutých výsledkov (priemerný ročný zárobok Jakuba pri použití vami vytvorenej politiky - pre každú z počiatočných distribúcií áut v prevádzkach. Keďže sa jedná o stochastický systém, je potrebné simulácie opakovať a výsledky spriemerniť.) a svojich skúseností s tvorbou riešenia (s čím ste mali problém, čo bolo jednoduché, ...)

- súbor s vašimi riešeniami (4x – jeden súbor pre riešenie každého z problémov). V tomto súbore bude uvedená vytvorená hodnotová funkcia v tvare

$$\text{hodnota} = \text{stav} : \text{akcia}$$

pričom

- hodnota je reálne číslo,
- každá kombinácia stav-akcia je na samostatnom riadku