Rozhodování, markovské rozhodovací procesy

Neřešené úlohy
Shromáždil: Jiří Kléma, klema@fel.cvut.cz

LS 2013/2014

Cíle materiálu:

Text poskytuje řešené úlohy jako podpůrný výukový materiál ke cvičením v předmětu A4B33ZUI.

1 Jednotlivá rozhodnutí, bayesovské rozhodování

Příklad 1. *(AIMA, 16.10): Jdete si koupit ojeté auto do bazaru. V úvahu připadá, že si auto před nákupem prověříte testem (kopnete do pneumatik, zavezete ho ke kamarádovi mechanikovi) a teprve pak se rozhodnete. Každé auto může být buď v dobrém nebo špatném stavu (s_+ a s_-). Nechť se rozhodujete o konkrétním autě a_1, jeho bazarová cena je 30000 Kč, tržní cena a_1 v dobrém stavu je 40000 Kč. Případná oprava auta (přechod ze špatného do dobrého stavu) stojí 14000 Kč. Odhadujete, že auto je v dobrém stavu a_{1+} s pstí 70%. Před nákupem můžete provést jeden konkrétní test t_1 za cenu 1000 Kč. Test určí, v jakém stavu auto je, ale s neurčitostí: $Pr(t_{1+}(a_1)|a_{1+}) = 0.8$ a $Pr(t_{1+}(a_1)|a_{1-}) = 0.35$.*

Vypočtěte střední čistý zisk pokud koupíte a_1 bez t_1.

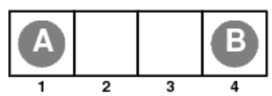
Použijte Bayesův teorém k určení psti, že auto je v dobrém stavu pro oba výsledky testu.

Najděte optimální rozhodnutí o nákupu pro oba výsledky testu.

Určete VPI testu t_1. Navrhněte optimální strategii pro potenciálního kupce auta a_1.

2 Markovské rozhodovací procesy

Příklad 2. *Uvažujme hru dvou hráčů na herním plánu o čtyřech polích. Každý z hráčů má jeden kámen a jeho cílem je dostat svůj kámen na opačnou stranu herního plánu (hráč A se z pole 1 musí dostat na pole 4, hráč B z pole 4 na pole 1). Vítězí ten hráč, kterému se to podaří prvnímu. Hráč A táhne jako první. Přípustné akce jsou pohyby vlevo a vpravo na sousední pole, nelze zůstat stát a vzdát se tahu, nelze táhnout mimo herní plán. Pokud je na vedlejším poli soupeřův kámen, je výsledkem pohybu přeskočení kamene (příklad: je-li A na pozici 3 a B na pozici 2 je výsledkem pohybu A vlevo posun A na pozici 1).*



Který z hráčů vyhraje? Naznačte klasické řešení problému pomocí prohledávání stavového prostoru.

Lze tuto úlohu formulovat a řešit jako MDP? Je to výhodné? Jak by se úloha musela změnit, aby to výhodné bylo?

Problém formulujte jako MDP. Nechť je $V_A(s)$ je hodnota stavu pokud je na tahu hráč A, $V_B(s)$ je hodnota stavu pokud je na tahu hráč B. Odměna ve stavu s nechť je $R(s)$, pro terminální stavy vítězné pro A je 1, pro terminální stavy vítězné pro B je -1. Nakreslete graf stavového prostoru. Zapište Bellmanovy rovnice pro oba hráče a aplikujte tyto rovnice v rámci hodnotové iterace. Formulujte ukončovací podmínku iterace.

Příklad 3. Za dveřmi je tygr (řešení jako POMDP). *Stojíte v místnosti, z níž vedou dvoje dveře. Víte, že za jedněmi dveřmi je hladový tygr, druhé dveře garantují bezpečný odchod z místnosti. Tygr občas zařve. Řev je slyšet, ale z jednoho poslechu není úplně zřejmé, odkud řev vychází. Chcete se z místnosti bezpečně a rychle dostat, v každém okamžiku se můžete rozhodnout mezi třemi volbami: otevřít dveře vlevo, otevřít dveře vpravo nebo počkat až tygr znovu zařve. Pracujte s následujícími ohodnoceními: otevření nesprávných dveří odpovídá ztrátě 100, otevření správných dveří zisku 10, čekání je spojeno se ztátou 1, při každém zařvání se zmýlíte v odhadu směru v 15% případů (ukážete na jedny ze dveří, ale správně to bude jen v 85%*

situací), pokud bereme v úvahu celou sekvenci náslechů, omyly jsou vzájemně nezávislé, tygr mezi řvaním svoji pozici nemění.

Problém formalizujte jako částečně pozorovatelný markovský rozhodovací proces.

Nalezněte optimální plán délky 1 jako funkci belief. Tj. navrhněte optimální akci v závislosti na tom jakou pravděpodobnost přiřazujete skrytým stavům. V jakých bodech belief prostoru se bude rozhodnutí měnit?

Kolik je podmíněných plánů délky 2? Určete užitek alespoň jednoho z nich (opět půjde o funkci belief). Bude některý z plánů čistě dominován plány jinými?

Kolikrát je třeba na začátku hry slyšet řev ze stejné strany předtím než se vyplatí otevřít jedny ze dveří? Zdůvodněte.