

Sztuczna inteligencja

Ćwiczenia 5

Zajęcia w przedostatnim tygodniu

Zadanie 1. Wycieczki górskie przebiegają na trasie między punktem A i B (to może być ten sam punkt, ważne, że po dotarciu do niego nigdzie już dalej nie idziemy, bo nam się nie chce). Przebiegają po sieci szlaków turystycznych, przy czym mapa podaje średni czas przejścia między każdymi dwoma węzłami sieci szlaków. Każdy odcinek łączy się z pewnymi wrażeniami pozytywnymi (bo ładne widoki) oraz z negatywnymi (bo trochę męczy, bo jakieś lawiny, możliwe oblodzenie). Oba te rodzaje wrażeń zależą od pogody, która, jak wiadomo, w górach może się dość szybko zmienić. Ponadto widoki cieszą (w obrębie danej wycieczki) tylko za pierwszym przejściem danego odcinka. Wycieczka sama w sobie (niezależnie od wrażeń) daje pewną satysfakcję, jeżeli ma określoną długość (nie jest za krótka i nie jest za długa). Wraz ze wzrostem czasu wycieczki pojawia się dodatkowy jej koszt (związany z wycieńczeniem, rosnącym ryzykiem itd). Po jakimś czasie wycieczka kończy się obligatoryjnie stanem, w którym turysta nie może się samodzielnie przemieszczać i czeka na helikopter (to oczywiście nie jest pożądana sytuacja).

Opisz możliwie dokładnie, jak takie wycieczki modelować jako proces decyzyjny Markowa (MDP). Co warto dodać do powyższego opisu? W jaki sposób taki model może pomóc turyście/przewodnikowi?

Zadanie 2. W jaki sposób idee uczenia ze wzmocnieniem mogą pomóc w zadaniu układania planu lekcji?

Zadanie 3. Rozważmy następującą ideę algorytmu uczenia ze wzmocnieniem:

1. Wybieramy jakąś politykę, liczymy wartość stanów dla tej polityki.
2. Następnie modyfikujemy politykę w sposób uwzględniający nowo policzone wartości
3. Czynności powtarzamy tak długo, jak długo ma to sens.

Opisz, jak zamienić te idee na algorytm. Jaki inny algorytm, omawiany na wykładzie, rozwiązuje ten sam problem?

Zadanie 4. Rozważamy grę w kółko i krzyżyk z losowym przeciwnikiem jako zadanie uczenia ze wzmocnieniem. Odpowiedz na następujące pytania:

- a) W MiniMaxowym rozwiązaniu kółka i krzyżyka sensowne jest uwzględnianie symetrii (dlaczego?). A w tym przypadku?
- b) Porównajmy gracza zachłannego (który zawsze wybiera ruch w danym momencie optymalny) z graczem stosującym strategię ϵ -zachłanną. W jakich sytuacjach spodziewasz się, że pierwszy osiągnie lepsze wyniki, a w jakich – że drugi?
- c) Jak powinien zmieniać się ϵ w trakcie uczenia?

Zadanie 5. Rozważamy procesy MDP. Przyjmujemy, że v_* oraz q_* to odpowiednio wartości stanów i akcji dla optymalnej polityki.

- a) Pokaż, jak wyrazić v_* za pomocą q_* .
- b) Pokaż, jak wyrazić q_* za pomocą v_* oraz $p(s', r|s, a)$.
- c) Pokaż, jak wyrazić π_* (optymalną politykę) za pomocą q_* .
- d) Pokaż, jak wyrazić π_* za pomocą v_* oraz $p(s', r|s, a)$.

Zadanie 6. Jakie warunki muszą być spełnione (tzn. w jaki sposób należy wybierać węzły do rozwijania), żeby Monte Carlo Tree Search w granicy dążyło do strategii wyznaczonej przez MinMax-a (odpowiedź uzasadnij). Jakie praktyczne zalecenia mogą wynikać z tych rozważań?

Zadanie 7. Zaproponuj algorytm, który dla formuły F w logice zdaniowej znajduje równoważną formułę F' w koniunkcyjnej postaci normalnej (tzn. taką, że $F \leftrightarrow F'$ jest tautologią). Pokaż przykładową formułę, dla której Twój algorytm spowoduje wykładniczy wzrost długości formuły.

Zadanie 8. Pokaż, że dla danej formuły zdaniowej F o długości N istnieje formuła zdaniowa w koniunkcyjnej postaci normalnej F' o długości $O(N)$, spełnialna wtedy i tylko wtedy, gdy F jest spełnialna.

Zadanie 9. Mamy n zmiennych, które mogą przyjmować wartości 0 lub 1 (czyli są *zmiennymi logicznymi*). Opisz sposób konstruowania formuły rachunku zdań (w koniunkcyjnej postaci normalnej), która jest równoważna zdaniu $x_1 + \dots + x_n = K$.

Zadania do poczytania

Poniższe zadania to materiał nadprogramowy, którego najprawdopodobniej nie będzie na wykładzie, a na pewno nie będzie na egzaminie (wyjątkiem jest zadanie o grach z jednym ruchem, o których trochę było na wykładzie).

Wymagają one przeczytania (niezbyt trudnego) tekstu i opowiedzenia koleżankom i kolegom o nim. Wszystkie numery rozdziałów odnoszą się do podręcznika *Artificial Intelligence Modern Approach 3rd edition* (znajduje się on również na SKOSie).

Zadanie 10. ★ Przeczytaj rozdział 16.2 THE BASIS OF UTILITY THEORY. Wyjaśnij pojęcia: loteria, relacja preferencji oraz funkcja użyteczności. Dlaczego liniowe przekształcenia funkcji użyteczności mają szczególne znaczenie?

Zadanie 11. ★ Przeczytaj podrozdział 16.3.2 The utility of money. Dlaczego rozsądne jest przyjąć, że użyteczność pieniędzy **nie** jest po prostu funkcją identyczności (z ew. zmianą jednostek)? Jak wyglądałby podatek liniowy płacony nie od dochodu, lecz od użyteczności dochodu?

Zadanie 12. ★ Przeczytaj podrozdział 16.3.4 Human judgment and irrationality i wyjaśnij, na czym polega paradoks Allais i jakie jego rozwiązania są proponowane. Analogicznie dla paradoksu Ellsberga.

Zadanie 13. ★ Przeczytaj podrozdział 16.6 THE VALUE OF INFORMATION. Wyjaśnij przykład z wartością praw wydobywania ropy w sektorach o nieznanym rzeczywistym zawartości tego surowca.

Zadanie 14. Przeczytaj rozdział 17.5.1 Single-move games i następnie:

- a) Przypomnij, jak rozwiązuje się grę *Two-finger Morra*,
- b) Powiedz, co to jest Dylemat więźnia, (★)
- c) Powiedz, co to jest punkt równowagi Nasha.

Zadanie 15. ★ Co to jest nadracjonalność (superrationality)? Tego nie ma w AIMA, ale łatwo znajdziesz odpowiednie informacje w Internecie.