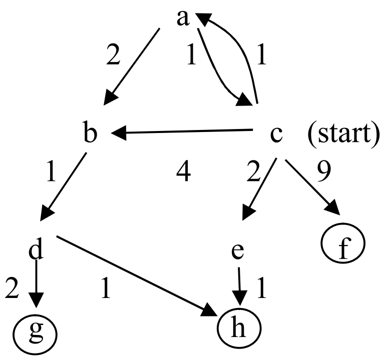
Osnove umetne inteligence, izpit 25. januar 2017

Navodila:
Čas: 80 min. Uporaba literature, zapiskov in elektronskih naprav ni dovoljena.
Točkovanje: vse naloge so enakovredne.
Ustni izpiti: Torek, 31. jan. ob 10h

1. Dan je spodnji prostor stanj:



Naj bo **c** začetno vozlišče preiskovanja. **f**, **g** in **h** so ciljna vozlišča. Algoritmi preiskovanja naj generirajo naslednike vozlišč po abecednem vrstnem redu. Npr. vrstni red naslednikov vozlišča **c** je: **a**, **b**, **e**, **f**.

Vsi preiskovalni algoritmi naj razpoznavajo cikle in generirano vozlišče, ki sklene cikel, takoj zavržejo. Vendar pa obravnavajo graf kot drevo. Torej, če pridejo do kakega vozlišča N po različnih poteh, naredijo kopijo N' vozlišča N in obravnavajo N', kot da bi bilo novo vozlišče. Če imata dve vozlišči enako f-oceno, se najprej razvije tisto vozlišče, ki je bilo prej generirano.

Hevristične ocene h vozlišč so dane takole:

X	a	b	c	d	e	f	g	h
h(X)	5	1	1	2	7	0	0	0

(a) Katero rešitveno pot vrne iskanje v globino, katero pa iterativno poglabljanje?
Globinsko: c,a,b,d,g; Iter. pogl.: c,f

(b) Katero rešitveno pot vrne algoritem A*?
c,a,b,d,h

(c) V kakšnem vrstnem redu A* razvija vozlišča pri vprašanju (b). Zapiši vrstni red teh vozlišč in njihove f-vrednosti.
c (f(c)=1), b(5), a(6), b'(4), d(6), [h(5)]

(d) Ali je zgoraj podana h-ocena vozlišča a optimistična? Pojasni odovor.
Ne; h(a) = 5 > 4 = h*(a)

(e) Katero rešitveno pot vrne algoritem IDA*?
c,a,b,d,g

(f) Zapiši zaporedne vrednosti spremenljivke f-limit med izvajanjem IDA* v vprašanju (d) ter vrstni red obiskanih vozlišč za vsako vrednost f-limit. Predpostavi, da je v prvi iteraciji f-limit=0.

f-limit	obiskana vozlišča
0	c
1	c,a,b,e,f
5	c,a,b,d,e,f (ali c,a,b,e,f,d)
6	c,a,b,d,g (ali c,a,b,e,f,b,d,g,h)

(g) Ali je funkcija f za ta prostor stanj in to hevristično funkcijo monotona? Na kratko utemelji odgovor.
Ne. Pojasnilo: b je naslednik od a, f(a)=6, f(b) = 4. Torej ima b manjši f kot b-jev predhodnik a, kar ne zadošča pogoju o monotonosti.

2. V problemu učenja odločitvenih dreves iz primerov naj ima razred tri možne vrednosti: R1, R2 in R3. Imamo atribut X s tremi možnimi vrednostmi: V1, V2 in V3. Iz ogromne množice učnih primerov so bile zanesljivo ocenjene naslednje verjetnosti:

$p(R1) = 1/2$, $p(R2) = 1/4$
$p(V1) = 1/2$, $p(V2) = 1/4$

(a) Koliko sta verjetnosti p(R3) in p(V3)?
$p(R3) = ¼$, $p(V3) = ¼$

(b) Kako je definirana "privzeta klasifikacijska točnost" (angl. "default accuracy")? Kolikšna je za ta problem učenja privzeta klasifikacijska točnost?
Privzeta klasifikacijska tocnost je definirana kot verjetnost vecinskega razreda v celotnem prostoru primerov. Za nas primer je to $p(R1) = ½$.

Znane so tudi tele pogojne verjetnosti:
$p(R1 \mid V1) = 1/2$, $p(R1 \mid V2) = 1$, $p(R1 \mid V3) = 0$,
$p(R2 \mid V1) = 1/4$, $p(R2 \mid V2) = 0$, $p(R2 \mid V3) = 1/2$
Iz teh podatkov je možno določiti tudi druge pogojne verjetnosti, kot npr. $p(R3 \mid V1)$ in $p(R3 \mid V2)$, ...

(c) Izračunaj verjetnosti $p(R3 \mid V1)$, $p(R3 \mid V2)$ ter $p(R3 \mid V3)$.
$p(R3 \mid V1) = 1/4$, $p(R3 \mid V2) = 0$, ter $p(R3 \mid V3) = ½$

(d) Koliko je informacijska vsebina I (entropija) na en primer v celotni množici primerov?
$I = -(½ \log ½ + 2*(1/4 \log ¼)) = ½ + ½ * 2 = 1.5$

(e) Naj odločitveno drevo vsebuje le koren, v katerem je atribut X, ter tri liste. Nariši to drevo. Za vsak list drevesa ustrezno določi večinski razred oz. verjetnostno porazdelitev razredov v listu.
X
veja1(V1,0.5) → R1(0.5)
veja2(V2,0.25) → R1 (1)
veja3(V3,0.25) → R2,R3 (oba verjetnost 0.5)

(f) Koliko je klasifikacijska točnost tega drevesa?
Tocnost drevesa = ½*1/2 + ¼ * 1 + ¼ * ½ = 5/8

(g) Koliko je informacijski dobitek Gain(X) atributa X?
Gain(X) = I – Ires = 1.5 – (½*1.5 + ¼*0 + ¼*1) = 0.5

3.

(a) Naslednja trditev se nanaša na hevristično preiskovanje prostora stanj: "Če algoritem A* uporablja hevristično funkcijo h, ki je monotona in optimistična, potem zagotovo najde optimalno rešitev." Komentiraj pravilnost in smiselnost te trditve. Predlagaj izboljšano trditev.
Trditev omenja monotonost h, ki ni niti definirana (definiran je pojem monotonosti f) niti ni potrebna za popolnost A*. Boljse: ""Če algoritem A* uporablja optimisticno hevristično funkcijo h, potem zagotovo najde optimalno rešitev."

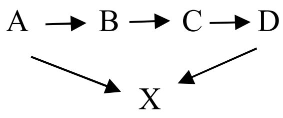
(b) Primerjaj časovno učinkovitost algoritmov A* in IDA* (merjeno s številom generiranih vozlišč, števši tudi morebitna ponovno generirana vozlišča). Navedi in na kratko razloži primer posebej neugodnih okoliščin za algoritem IDA*, ki ta algoritem lahko zelo upočasnijo.
Oba A* in IDA* imata v principu eksponentno casovno zahtevnost. IDA* je sicer zaradi ponovnega generiranja istih vozlisc pocasnejsi, vendar je njegova zahtevnosti se vedno priblizno istega reda kot A*, ce imajo mnoga vozlisca enako oceno f. Zelo neugoden primer pa nastane, ce imajo (skoraj) vsa vozlisca razlicne f vrednosti. V tem primeru povecevanje f-limit zahteva veliko stevilo iteracij ...

(c) V planiranju po principu sredstev in ciljev se uporablja t.i. regresiranje ciljev skozi akcijo. Naj bo G množica ciljev, A akcija in RG množica regresirani ciljev. Zapiši z uporabo relacij CAN, ADD in DEL, kako za dani G in A določimo RG. Zapiši tudi formalni pogoj, po katerem regresijski planer vnaprej prepozna, da za dane cilje G neka akcija B ni neposredno uporabna.

RG=CAN(A) U G \ ADD(A)

Pogoj: $DEL(B) \cap G = \emptyset$

4. Dana je bayesovska mreža z naslednjo strukturo:



(a) Koliko verjetnosti (oz. pogojnih verjetnosti, to je številskih podatkov) mora biti podanih za to strukturo?
11 podatkov

(b) Naj bodo vse povezave v mreži "pozitivno vzročne", torej vzročni dogodek povečuje verjetnost posledice (npr. puščica od A do B nakazuje, da če se zgodi A, postane B bolj verjeten). Brez računanja ugotovi, katera pogojna verjetnost je večja: P(D|C) ali P(D|AC)?
Obe verjetnosti sta enaki, saj ce je podan C, D ni odvisen od A.

V naslednjih vprašanjih c, d in e z upoštevanjem strukture mreže čim bolj poenostavi pogojni del v danih pogojnih verjetnostih:

(c) P(D | A C) = P(D|C)

(d) P(D | A B C) = P(D|C)

(e) P(D | A C X) = P(D|ACX)

(f) Zapiši vse množice E, ki d-ločujejo dogodka A in D.

{B}, {C}, {B,C}