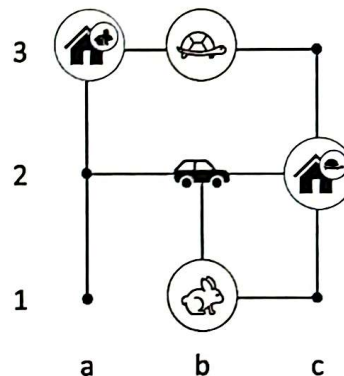


Odgovarjajte kratko in jedrnat, točno na zastavljena vprašanja. Vse odgovore pišite na črto pod vprašanji in izključno na ta list, ki ga edinega oddate na koncu Elektronske naprave niso dovoljene. Čas pisanja: 75 minut.

(1)

Slika na desni predstavlja situacijo v kateri sta se dva hišna ljubljénčka, želva in zajec, ponovno izgubila. Z avtomobilom ju želimo poiskati in vrniti domov; želva stanuje na polju c2, zajec pa na polju a3. V avtomobilu je zadosti prostora za oba, če bi ju želeli peljati skupaj, stepla se pa tudi ne bosta, saj sta prijatelja še od lanske avanture. Povezave predstavljajo cesto; direktne poti med a1 in b1 ni, prav tako je ni med b2 in b3. Svet opišemo z relacijo  $at(object, location)$ , kjer je *object* lahko avto, želva ali zajec, *location* pa je lahko polje na zemljevidu (npr. b2) ali za ljubljénčka tudi »in car«. Akcije v tem svetu so *left/right/up/down*, kar ustrezno spremeni pozicijo avtomobila, *pick(X)*, ki pobere ljubljénčka *X* in *drop(X)*, ki ljubljénčka *X* odloži. Seveda se mora avtomobil nahajati na ustrezni lokaciji, da lahko nekoga pobere, odloži pa ga načeloma lahko kjerkoli.



- (1) (a) Dopolnite spodnji STRIPS opis akcije  $pick(X)$ , torej akcije, ki pobere ljubljénčka *X*.  
 predpogoji:  $at(car, Loc), \dots at(X, Loc)$  učinki:  $\sim at(X, Loc), \dots at(X, "in car")$
- (1) (b) Regresirajte množico ciljev  $G = \{ at(zajec, a3), at(\text{želva}, c2) \}$  skozi akcijo  $A = drop(\text{želva})$ .  
 $RG = \{ at(zajec, a3), at(\text{želva}, "in car"), at(car, c2) \}$
- (1) (c) Kateri plan bo vrnil planer po principu regresiranja ciljev, če uporablja iskanje v širino? (lahko je več rešitev)  
 Najkrajši (11 korakov/akcij), recimo:  $down, pick(z), up, left, up, drop(z), right, pick(\text{ž}), right, down, drop(\text{ž})$
- (1) (d) Kako bi spremenili/dopolnili spodnji STRIPS opis akcije  $pick(X)$  v neljubem primeru, da se zajec in želva skregata in ju ne moremo peljati oba skupaj?  
 predpogoji:  $at(car, Loc), \dots at(X, Loc), empty$  učinki:  $\sim at(X, Loc), \dots at(X, "in car"), \sim empty$

(2)

Sedaj pa nam slika iz zgornje naloge (1) predstavlja navigacijski problem, ki ga rešujemo s pomočjo (hevrističnega) preiskovanja. Začetna pozicija avtomobila je b2, končni vozlišči pa sta obe hiški (a3 in c2). Skratka, naloga je avtomobil zapeljati pod streho. Premiki so takšni kot jih dovoljujejo povezave na sliki, npr. iz začetne pozicije se lahko premaknemo levo, desno ali navzdol. Vse povezave so dvosmerne. Vrstni red generiranja potez (premikov) je vedno levo/desno/gor/dol. (Želva in zajec sta medtem varno doma in pri tej nalogi ne igrata nobene vloge.)

Ob hevrističnem preiskovanju imamo na voljo hevristično funkcijo; le-ta je za vsako vozlišče (pozicijo) preprosto Manhattanska razdalja do najbližjega ciljnega vozlišča (pri tem vse povezave štejejo 1). Npr. za pozicijo b1 je  $h(b1) = 2$  (razdalja do končnega vozlišča c2). Cene povezav so naslednje: v drugi vrstici so ceste slabe (makadam), zato je cena teh povezav 5. Ostale ceste so asfaltirane, vendar pa se svet vzpenja navzgor. Posledično povezave navzgor stanejo 2. Vse ostale, vključno s povezavami navzdol, pa imajo ceno 1. (Pozor: dvosmerne povezave gor/dol nimajo enake cene.)

- (1) (a) Katero, če sploh, rešitev bo vrnilo preiskovanje v globino?  
 zacikla se: levo-desno-levo-...
- (1) (b) Katero, če sploh, rešitev bo vrnilo iterativno poglobljanje?  
 najkrajšo (po številu korakov), torej: desno
- (1) (c) Ali je dana hevristična funkcija dopustna (angl. *admissible*)? Odgovor na kratko utemeljite!  
 DA. Vrednosti potez (premikov) so vedno 1 pri Manhattanski razdalji; hevristične cene vseh povezav pa so  $\geq 1$ .
- (2) (d) Katero rešitev bo vrnila  $A^*$  in kakšna je cena rešitve, ki jo vrne IDA\*? («tie-break»: vrstni red generiranja vozlišč)  
 $A^*$ : dol-desno-gor ali b2-b1-c1-c2; IDA\*:  $1+1+2=4$

(3)

Na voljo imamo podatke naslednjih opazovanj: {2, 4, 4, 6, 7, 7, 12, 12, 13, 16, 17, 25}. Z uporabo nenadzorovanega učenja, konkretno gručenja k-means, želimo nad temi podatki zgraditi gruče.

(1) a) Začetna centroida sta postavljena na C1=0 in C2=18. V kateri gruči so posamezne točke po prvi iteraciji gručenja?  
C1: 2, 4, 4, 6, 7, 7      C2: 12, 12, 13, 16, 17, 25

(1) b) Kakšne so vrednosti centroidov po eni iteraciji pri treh centroidih z začetnimi vrednostmi C1=6, C2=16 in C3=26?  
C1: 5      C2: 14      C3: 25

(4)

Podane so matrike napovedi (angl. *confusion matrices*) za tri različne modele za napovedovanje prisotnosti poškodbe v proizvedenem izdelku: razred da = poškodba je prisotna, razred ne = izdelek je OK.

		kNN		Naivni Bayes (NB)		Naključni gozd (RF)	
		napovedan razred		napovedan razred		napovedan razred	
		da	ne	da	ne	da	ne
Pravi razred	da	1	24	16	9	20	5
	ne	5	70	8	67	15	60

(1) (a) Od najboljšega do najslabšega uredite vse tri modele po njihovi klasifikacijski točnosti na teh podatkih.  
(33%) (80%) (41%)  
NB > RF > kNN

(1) (b) Kateri model bi izbrali, če je škoda, da poškodovan izdelek date v prodajo, zelo visoka? Zakaj?  
RF, ker polovi največ poškodovanih izdelkov; ima največji recall za "da"

(1) (c) Recimo, da želimo kar najbolj prihraniti čas ročnega preverjanja izdelkov z morebitno napako. Želimo torej kar najboljšo natančnost modela pri napovedi »da«, torej da izdelek ima napako? Kolikšna je ta natančnost in kateri model je najboljši pri tem?

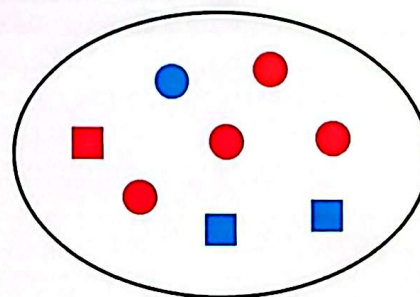
NB, natančnost za "da" je  $\frac{2}{3} \approx 67\%$ .

(1) (d) Kateri model ima najvišji priklic za razred »ne« in kolikšen je?

kNN; priklic za "ne" je  $\frac{70}{75}$

(5)

Imamo dvorazreden problem, ker nam razred predstavlja barva. Vsi primeri so opisani le z enim binarnim atributom »oblika«, ki ima vrednosti »krog« in »kvadrat«. Na sliki je podana učna množica.



(1) (a) Kakšna je gini nečistoča celotne učne množice? Rezultat lahko pustite v obliki ulomka.

$\frac{30}{64} = \frac{15}{32} = 0,46875$

(1) (b) Kakšna bo klasifikacijska točnost odločitvenega drevesa, zgrajenega na podani učni množici, če je naša testna množica {■, ●, ■, ■, ■, ●, ●, ■}?

$\frac{6}{8} = 75\%$

(1) (c) V kateri razred in s kakšno verjetnostjo napove model Naivnega Bayesa, zgrajen na dani učni množici, nov primer za katerega ne ve oblike?

v rdeč razred z verjetnostjo  $\frac{5}{8} = 62,5\%$