

MI hetedik előadásához kapcsolódó kérdések, feladatok (szükséges mintaszám VKH)

1. A hipotézisterünk elemszáma: $|\mathbf{H}|=10^{29}$. Ha azt szeretnénk, hogy 90% biztonsággal állíthassuk, hogy az eszköz hibája nem lesz 1%-nál nagyobb, akkor a tanult mintaszámkorlát által adott érték 6908. Mikor fog inkább nőni a mintaszámkorlát értéke, ha az állításunk bizonyosságát növeljük 99%-ra, vagy ha az elérendő hibát csökkentjük 0,1%-ra? (Segítség: $\ln(10)=2,303$)
2. IGAZ/HAMIS?
 - a. A hipotézisteret két részre bontottuk a VKH megoldások vizsgálatánál. A $|\mathbf{H}|_{ROSSZ} \leq |\mathbf{H}|$ egyenlőtlenség mindig igaz.
 - b. A hipotézistérben a mindegyik mintánkkal konzisztens legegyszerűbb hipotézis megtalálása rendszerint kezelhetetlen komplexitású feladat.
 - c. Annak, hogy egy $\epsilon=0,1$ hibájú hipotézis konzisztens legyen az 100 véletlenszerűen választott tanítómintával tízszer akkora, mint hogy egy $\epsilon=0,01$ hibájú hipotézis legyen konzisztens ezekkel a mintákkal.
 - d. Az állapot hasznosságát csak akkor tudjuk értelmezni, ha az adott állapotból csak egy végállapotba lehet eljutni.
 - e. A szekvenciális probléma iteratív megoldása során gyakran az értékek még nem konvergálnak pontosan, de a stratégia már egyértelmű lehet.
3. Egy szekvenciális döntési problémában 4 állapot alkotja az állapotteret, S_3 a végállapot. Minden állapotban kétféle cselekvés (a_1 és a_2) közt választhatunk. Az alábbi baloldali táblázatban láthatók az a_1 -hez tartozó állapot-átmenet valószínűségek, a jobboldali táblázatban az a_2 cselekvéshez tartozók. A leszámítolási tényező $\gamma=0,5$.

a_1 esetén $T(s \rightarrow s')$		s'			
		S1	S2	S3	S4
s	S1	0,5	0	0,5	0
	S2	0	0,5	0	0,5
	S3	0	0	0	0
	S4	0	0	0,5	0,5

a_2 esetén $T(s \rightarrow s')$		s'			
		S1	S2	S3	S4
s	S1	0	0,5	0	0,5
	S2	0,2	0	0,8	0
	S3	0	0	0	0
	S4	1	0	0	0

Az egyes állapotokhoz tartozó jutalmak $R(S_1) = -0,5$; $R(S_2) = +0,4$; $R(S_3)=+1,2$; $R(S_4)=+0,5$. Értékiterációs algoritmust alkalmazunk, az eddigi iterációk eredményeképp a t . lépésben a becsült hasznosságok $U_t(S_1) = 0$; $U_t(S_2) = 1$; $U_t(S_3)= R(S_3)= 1,2$; $U_t(S_4) = 1$. Megegyeznek-e az iteráció során kapott értékek a valós hasznosságokkal? Előfordulhat-e nullánál nagyobb valószínűséggel, hogy a rendszer soha nem jut el a végállapotába (S_3 -ba), ha mindig az a_2 cselekvést választjuk?

4. p bizonyossággal tudjuk garantálni (pl. a hipotézisünkkel konzisztens mintahalmaz elemszámával), hogy a hipotézisünk ε -nál nem nagyobb hibát fog elkövetni, amikor új mintákat mutatunk neki. Ha az eszközünket ilyen hiba mellett N nyereséggel el tudjuk adni, de nem teljesítés esetén K kártérítést kell fizetnünk, akkor mekkora várható nyereségre számíthatunk az egyes eladott darabok után?
5. Ha egy adott állapotból p_1 valószínűséggel jutunk 3 lépésben a V_1 végállapotba, és ez alatt J_1 jutalmat gyűjtünk, illetve p_2 valószínűséggel jutunk 8 lépésben a V_2 végállapotba, és ez alatt J_2 jutalmat gyűjtünk, akkor mekkora lesz a vizsgált állapot hasznossága? (A leszámítolási tényező 1.)

MI nyolcadik előadásához kapcsolódó kérdések, feladatok (megerősítéses tanulás)

1. Az alábbi útvesztőben passzív megerősítéses tanulást végzünk. Az egyes mezőkben a felső szám az állapot sorszáma, az alsó, zárójelben lévő szám – ha van – az állapothoz rendelt jutalom értéke. Ahol nincs feltüntetve a jutalom értéke, ott nulla. Nem minden mezőben tüntettük fel az állapot sorszámát, értelemszerűen sorban növekednek a számok. A végállapot a szürke színnel is megjelölt 10-es.

1	2	3	4
5	6	7	8
9		10 +12	11
12	13		14 (-5)
15 (-10)	16	17	18

Időbeli különbség tanulást végzünk az alábbi lépéssorozat mentén, a tanulás bátorsági faktora 0,1; a leszámítolási tényező 0,8:

$$1 \Rightarrow 5 \Rightarrow 9 \Rightarrow 12 \Rightarrow 13 \Rightarrow 16 \Rightarrow 17 \Rightarrow 18 \Rightarrow 14 \Rightarrow 11 \Rightarrow 8 \Rightarrow 11 \Rightarrow 14 \Rightarrow 11 \Rightarrow 8 \Rightarrow 7 \Rightarrow 10$$

Mi lesz a 11-es állapot hasznosságértékének új becslése a lépéssorozat után, ha a lépéssorozatot megelőzően a hasznosságbecslések a következők voltak:

$U_1 = 5$	$U_2 = 6$	$U_3 = 7$	$U_4 = 8$	$U_5 = 7$	$U_6 = 8$	$U_7 = 9$	$U_8 = 8$	$U_9 = -6$
$U_{10} = +12$	$U_{11} = +10$	$U_{12} = -9$	$U_{13} = -9$	$U_{14} = -4$	$U_{15} = -9$	$U_{16} = -7$	$U_{17} = -7$	$U_{18} = -7$

(Ne törődjön azzal, hogy kialakulhatott-e ez a hasznosság becslés!)

2. IGAZ/HAMIS?

- a. Mivel az időbeli különbség tanuláshoz nincs szükség az állapotátmenet-modellre, ezért Q-tanulásra is használható.
- b. A hasznosságfüggvény explicit reprezentációja jobb általánosító képességet tesz lehetővé, mint az implicit reprezentáció.
- c. Az adott állapothoz tartozó legnagyobb Q érték az állapot hasznosságát adja.
- d. A mohó felfedezést végző ágens könnyen egy szuboptimumba juthat.
- e. A hóbortos felfedezést végző ágens nagyon jól kiismeri környezetét, de nem aknázza ki ezt a tudását a jutalmak gyűjtésére.
- f. Az aktív megerősítéses tanulásban kettős célja van az ágensnek: egyrészt minél nagyobb jutalmakat akar gyűjteni, másrészt javítani akarja a megszerzett tudást, a jövőbeli szekvenciákhoz.

- g. Eljárásmód-iteráció esetén, ha a t -dik lépésben kiszámítottuk az $U_t(s)$ értékeket, akkor a következő eljárás mód becslés – $(\pi_{t+1}(s))$ – nem függ a leszámítolási tényezőtől.
3. Mi az ϵ -mohó felfedezési stratégia célja és lényege?
 4. Boltzmann felfedezési stratégiát alkalmazunk. Egy s állapotban a lehetséges cselekvések hasznossága $Q(a_1,s)=1$; $Q(a_2,s)=2$; $Q(a_3,s)=1$. Milyen valószínűséggel választjuk az a_1 , a_2 és a_3 cselekvést, ha a „hőmérséklet” $T=10^6$, ha $T=1$, illetve ha $T=10^{-6}$?
 5. Egy aktív megerősítéses tanulási problémában 3 állapot alkotja az állapotteret, S3 a végállapot. Minden állapotban kétféle cselekvés (a_1 és a_2) közt választhatunk. Az alábbi baloldali táblázatban láthatók az a_1 -hez tartozó állapot-átmenet valószínűségek, a jobboldali táblázatban az a_2 cselekvéshez tartozók. A leszámítolási tényező $\gamma=0,5$.

a ₁ esetén $T(s \rightarrow s')$		s'		
		S1	S2	S3
s	S1	0,5	0	0,5
	S2	0	0,5	0,5
	S3	0	0	0

a ₂ esetén $T(s \rightarrow s')$		s'		
		S1	S2	S3
s	S1	0	0,5	0,5
	S2	0,2	0	0,8
	S3	0	0	0

Az egyes állapotokhoz tartozó jutalmak $R(S_1) = -0,5$; $R(S_2) = +0,4$; $R(S_3)=+1,2$. Eljárásmód-iterációs algoritmust alkalmazunk, az eddigi iterációk eredményeképp a t . lépésben a becült eljárás mód $\pi_t(S_1) = a_2$; $\pi_t(S_2) = a_2$, a becült hasznosságok $U_t(S_1) = 0$; $U_t(S_2) = 1$; $U_t(S_3)= R(S_3)= 1,2$. (Ne törődjön vele, hogy kialakulhattak-e ezek az értékek!)

A következő iterációs lépés végére milyen eljárás módot és milyen hasznosságértékeket kapunk?

