	Fakülte	Mühendislik	Sınav	Final	Numara	
	Bölüm	Bilgisayar Müh.	Tarih	24/01/2024	Ad-Soyad	
	Ders	BMB4019	Süre	60 dk.	İmza	

# SORULAR

- 1) Bir robot, 5 durumlu ortamda (S1,S2,S3,S4,S5) bulunmaktadır. 5 numaralı durum "Terminal" durumu olup, eylemini sonlandırmaktadır. Robotun ödül durumları aşağıdaki gibi olmaktadır.

$$R(S5)=1$$

$$R(Si)=0 \text{ (i=1,2,3,4)}$$

Robot, döngü-1 (episode-1)'de ve döngü-2 (episode-2)'de  $S1 \rightarrow S2 \rightarrow S3 \rightarrow S4 \rightarrow S5$  durumlarını takip etmektedir. TD'den (temporal difference) yararlanarak, başlangıç V değerlerinin 0 olduğunu gözeterek, bu robotun iki döngü sonucunda ulaşacağı V değerlerini her bir durum için bulunuz (düşüm/indirim faktörü ( $\gamma$ )=0.9, öğrenme oranı( $\alpha$ )=0.1). (30p)

Episode-1:

$$V(s1)=V(s1)+\alpha \cdot \delta=0+0.1 \cdot 0=0$$

$$V(s2)=V(s2)+\alpha \cdot \delta=0+0.1 \cdot 0=0$$

$$V(s3)=V(s3)+\alpha \cdot \delta=0+0.1 \cdot 0=0$$

$$V(s4)=V(s4)+\alpha \cdot \delta=0+0.1 \cdot 0=0$$

$$V(s5)=0$$

Episode-2:

$$V(s4)=V(s4)+\alpha \cdot \delta=0+0.1 \cdot 0.9=0.09$$

$$V(s3)=V(s3)+\alpha \cdot \delta=0+0.1 \cdot 0.081=0.0081$$

$$V(s2)=V(s2)+\alpha \cdot \delta=0+0.1 \cdot 0.00729=0.000729$$

$$V(s1)=V(s1)+\alpha \cdot \delta=0+0.1 \cdot 0.0006561=0.00006561$$

Son değerler:

$$V(s1)=0.00006561, V(s2)=0.000729, V(s3)=0.0081, V(s4)=0.09, V(s5)=1$$

- 2) Replay buffer kapasitesi 6 olan bir robot, ortamda aşağıdaki TD hata değerlerine ulaşmaktadır.

$$\delta 1=0.8, \delta 2=1.2, \delta 3=0.5, \delta 4=2.0, \delta 5=0.3, \delta 6=1.5$$

Aşağıdaki sabitleri kullanmaktadır:

$$\text{Öncelik derecesi için güç katsayısı: } \alpha=0.6$$

$$\text{Küçük sabit } (\epsilon) \text{ öncelik değeri: } 0.01$$

$$\text{Örnekleme için IS ağırlıklandırma derecesi: } \beta=0.4$$

S4 ve S4 durumları için IS ağırlıklarını hesaplayınız ( $w4$  ve  $w6$ ). (40p)

$$p1=0.8+0.01=0.81, p2=1.2+0.01=1.21, p3=0.5+0.01=0.51,$$

$$p4=2.0+0.01=2.01, p5=0.3+0.01=0.31, p6=1.5+0.01=1.51$$

$$\text{Toplam öncelik}=0.868+1.126+0.650+1.515+0.505+1.293=5.957$$

Normalize olasılıklar:

$$P(1)=0.868/5.957=0.146, P(2)=1.126/5.957=0.189, P(3)=0.650/5.957=0.109$$


$$P(4)=1.515/5.957=0.254, P(5)=0.505/5.957=0.085, P(6)=1.293/5.957=0.217$$

$$w4=(0.656)^{0.4}=0.84 \quad w6=(0.768)^{0.4}=0.89$$

- 3) Bir robot 8 durumlu bir yüzeyde hareket etmek istemektedir.

- Durumlar (S): S1, S2, S3, S4, S5, S6, S7, S8
- Eylemler (A): A1, A2 (Her durumda iki eylem seçeneği var.)
- İndirim Faktörü ( $\gamma$ ): 1

"Geçiş Ödülleri" tablosu aşağıda verilmektedir (Örneğin,  $R(S1 \rightarrow S2)=5$ ). "Terminal" durumuna ulaştığında robot daha fazla eylem gerçekleştiremez ve hareket sonlanır (robot artık başka bir eylem yapamaz). Başlangıç durumu S1'dir ve geçiş olasılıkları görmezden gelinmiştir.

	<b>Fakülte</b> <i>Mühendislik</i>	<b>Sınav</b> <i>Final</i>	<b>Numara</b>	
	<b>Bölüm</b> <i>Bilgisayar Müh.</i>	<b>Tarih</b> <i>24/01/2024</i>	<b>Ad-Soyad</b>	
	<b>Ders</b> <i>BMB4019</i>	<b>Süre</b> <i>60 dk.</i>	<b>İmza</b>	

Durum	Eylem	Sonraki Durum	Ödül
S1	A1	S2	5
S1	A2	S3	2
S2	A1	S4	0
S2	A2	S5	-1
S3	A1	S5	0
S3	A2	S6	1
S4	A1	S7	10
S4	A2	S8	-10
S5	A1	S7	-5
S5	A2	S8	5
S6	A1	S7	0
S6	A2	S8	0
S7	-	Terminal	0
S8	-	Terminal	0

- a) Q değerleri için **denklemleri yazınız** ? (10P)  
b) V değerleri için **denklemleri yazınız** ? (10P)

- a)  $Q(S1,A1)=R(S1,A1)+\max Q(S1,A1')$   
 $Q(S1,A2)=R(S1,A2)+\max Q(S1,A2')$   
 $Q(S2,A1)=R(S2,A1)+\max Q(S2,A2')$   
 $Q(S2,A2)=R(S2,A2)+\max Q(S2,A2')$   
 $Q(S3,A1)=R(S3,A1)+\max Q(S3,A1')$   
 $Q(S3,A2)=R(S3,A2)+\max Q(S3,A2')$   
 $Q(S4,A1)=R(S4,A1)+\max Q(S4,A1')$   
 $Q(S4,A2)=R(S4,A2)+\max Q(S4,A2')$   
 $Q(S5,A1)=R(S5,A1)+\max Q(S5,A1')$   
 $Q(S5,A2)=R(S5,A2)+\max Q(S5,A2')$   
 $Q(S6,A1)=R(S6,A1)+\max Q(S6,A1')$   
 $Q(S6,A2)=R(S6,A2)+\max Q(S6,A2')$   
 $Q(S7,A1)=R(S7,A1')=0$   
 $Q(S7,A2)=R(S7,A2')=0$   
 $Q(S8,A1)=R(S8,A1')=0$   
 $Q(S8,A2)=R(S8,A2')=0$

- b)  $V(S1)=\max\{R(S1,A1)+V(S2), R(S1,A2)+V(S3)\}$   
 $V(S2)=\max\{R(S2,A1)+V(S4), R(S2,A2)+V(S5)\}$   
 $V(S3)=\max\{R(S3,A1)+V(S5), R(S3,A2)+V(S6)\}$   
 $V(S4)=\max\{R(S4,A1)+V(S7), R(S4,A2)+V(S8)\}$   
 $V(S5)=\max\{R(S5,A1)+V(S7), R(S5,A2)+V(S8)\}$   
 $V(S6)=\max\{R(S6,A1)+V(S7), R(S6,A2)+V(S8)\}$   
 $V(S7)=R(S7,A1)$   
 $V(S8)=R(S8,A2)$

4) Proje