

שאלה 1 [27 נק']

במשחק הפאזל שלנו יש לוח של 1×7 , שלושה כלים שחורים, ושלושה כלים לבנים. המצב ההתחלתי הוא:



והמטרה היא שכל הכלים הלבנים יהיו משמאלם של כל הכלים השחורים; המיקום הסופי של הכלים השחורים לא מעניין. שני סוגים של פעולות היום חוקיים במשחק זה:

1. אפשר להעביר כל כלי למשבצת הסמוכה אם אותה משבצת פנויה. מחיר פעולה כזאת הוא 1.
2. אפשר להקפיץ כלי מעל אחד או שני כלים אחרים לתוך משבצת פנויה. מחיר פעולה כזאת שווה למספר הכלים שמדולגים (כלומר 1 או 2).

את התשובות יש לספק אך ורק בטבלה שנמצאת בסוף השאלה!

(א) מהי דרגת הסיעוף (branching factor) של בעיית החיפוש הזאת?

(ב) האם מרחב החיפוש מכיל מעגלים?

(ג) האם החיפוש לרוחב (BFS) הינו אופטימלי לבעיה הזאת?

(ד) האם החיפוש A^* בגרסת "חיפוש עץ" עם היוריסטיקה $h(state) = 0$ הינו אופטימלי לבעיה הזאת?

(ה) האם ההיוריסטיקה h_1 אשר שווה למספר הכלים השחורים משמאלו של הכלי הלבן השמאלי ביותר הינה קבילה?

(ו) האם ההיוריסטיקה h_2 אשר שווה למספר הכלים השחורים משמאלו של הכלי הלבן הימני ביותר הינה קבילה?

(ז) האם ההיוריסטיקה h_3 אשר שווה למספר הכלים השחורים משמאלו של הכלי הלבן הימני ביותר ועוד מספר הכלים הלבנים מימינו של הכלי השחור השמאלי ביותר הינה קבילה?

(ח) האם ההיוריסטיקה $h_4(state) = \frac{1}{2} \cdot h_3(state)$ הינה קבילה?

(ט) האם ההיוריסטיקה $h_5(state) = 2 \cdot h_3(state)$ הינה קבילה?

תשובות:

(א)	(ב)	(ג)	(ד)	(ה)	(ו)	(ז)	(ח)	(ט)
6	כן	לא	כן	כן	כן	לא	כן	לא

שאלה 2 [20 נק']

אתם בונים רובוט שבכל צעד יכול לנוע מהר (FAST) או לאט (SLOW). באופן כללי, תנועה מהירה מזכה את הרובוט בפרס +2, ותנועה איטית מזכה אותו בפרס של רק +1. יחד עם זאת, הרובוט צריך גם להתחשב בחום המנועים שלו, שיכול להיות תקין (OK) או גבוה (HOT). תנועה איטית נוטה להוריד את חום המנועים, ותנועה מהירה נוטה להעלות אותו. אם חום המנועים גבוה, אז קיימת סכנה של התחממות יתר, שבמקרה כזה הרובוט חייב לעצור, להתקרר, ולבצע תיקונים (הערה: זאת הסיטואציה שמתוארת בשורה האחרונה של הטבלה להלן).

פורמלית, פונקציית המעבר ופונקציית הפרס של ה-MDP שלנו הינם כלהלן:

s	a	s'	$P(s' s, a)$	$R(s, a, s')$
OK	SLOW	OK	1.0	+1
OK	FAST	OK	0.5	+2
OK	FAST	HOT	0.5	+2
HOT	SLOW	OK	1.0	+1
HOT	FAST	HOT	0.5	+2
HOT	FAST	OK	0.5	-10

בהמשך להערה הקודמת, שימו לב שהתיקונים יקרים, אך מובטח שחום מנועי הרובוט יהיה תקין לאחר מכן.

(א) בצעו שתי איטרציות של האלגוריתם value iteration עם מקדם הפליית הפרסים העתידיים $\gamma = 0.8$, ודווחו את התוצאות בטבלה להלן. שימו לב: אין צורך לחשב $V_2(\text{HOT})$.

s	V_0	V_1	V_2
OK	0	2	3.2
HOT	0	1	×

$$V_{i+1}(s) = \max_a \sum_{s'} P(s' | s, a) [R(s, a, s') + \gamma V_i(s')]$$

$$V_1(\text{OK}) = \max [(1(1 + \gamma \cdot 0)), (0.5(2 + \gamma \cdot 0) + 0.5(2 + \gamma \cdot 0))] = \max(1, 2) = 2$$

$$V_1(\text{HOT}) = \max [(1(1 + \gamma \cdot 0)), (0.5(2 + \gamma \cdot 0) + 0.5(-10 + \gamma \cdot 0))] = \max(1, -4) = 1$$

$$V_2(\text{OK}) = \max [(1(1 + \gamma \cdot 2)), (0.5(2 + \gamma \cdot 2) + 0.5(2 + \gamma \cdot 1))] = \max(2.6, 3.2) = 3.2$$

(ב) בצעו למידת Q-learning עם מקדם הפליית הפרסים העתידיים $\gamma = 0.8$, ומקדם "קצב למידה" (learning rate) $\alpha = 0.5$, על בסיס סדרת ההתנסויות הבאה:

OK, FAST, HOT, reward +2, (calculate Q_1)
 HOT, FAST, OK, reward -10, (calculate Q_2)
 OK, SLOW, OK, reward +1, (calculate Q_3)

הערות:

- הלמידה מסתיימת עם פונקציית Q_3 , אך עליכם לדווח גם את שתי פונקציות הביניים אשר מתקבלות אחרי ביצוע הפעולה הראשונה וביצוע הפעולה השנייה.

- אין להעתיק משלב לשלב את הערכים אשר נשארים ללא שינוי.

s	a	Q_0	Q_1	Q_2	Q_3
OK	SLOW	0			0.9
OK	FAST	0	1.0		
HOT	SLOW	0			
HOT	FAST	0		-4.6	

$$\begin{aligned}
 Q(s, a) &\leftarrow Q(s, a) + 0.5 \left[R(s, a, s') + 0.8 \max_{a'} Q(s', a') - Q(s, a) \right] \\
 Q(\text{OK}, \text{FAST}) &\leftarrow 0.0 + 0.5 \left[2.0 + 0.8 \max_{a'} Q(\text{HOT}, a') - 0.0 \right] = 1 \\
 Q(\text{HOT}, \text{FAST}) &\leftarrow 0.0 + 0.5 \left[-10 + 0.8 \max_{a'} Q(\text{OK}, a') - 0.0 \right] = -4.6 \\
 Q(\text{OK}, \text{SLOW}) &\leftarrow 0.0 + 0.5 \left[1.0 + 0.8 \max_{a'} Q(\text{OK}, a') - 0.0 \right] = 0.9
 \end{aligned}$$

(ג) מסתבר שאחרי 10 תקלות, הרובוט כבר אינו ניתן לתיקון ולכן מושבת. מהו המספר הקרוב ביותר למספר המינימלי של מצבים שבעזרתם נוכל לייצג את ה-MDP החדש?

- (a) 2
- (b) 10
- (c) 100
- (d) 2^{10}
- (e) 10^2

(ד) אנחנו שוקלים להשתמש במנגנון למידה אשר מבצע עדכונים שמעריבים שתי פעולות עוקבות (ולא פעולה בודדת, כמו בדרך כלל). איזו מהנוסחאות הבאות משקפת נכון את הערך של מצב s תחת מדיניות π , בהינתן ש:

$$V^\pi(s) = R(s) + \sum_{s'} P(s'|s, \pi(s)) \gamma V^\pi(s')$$

$$V^\pi(s) = R(s) + \sum_{s'} P(s'|s, \pi(s)) \sum_{s''} P(s''|s', \pi(s')) \gamma^2 V^\pi(s'') \quad (\text{a})$$

$$V^\pi(s) = R(s) + \sum_{s'} P(s'|s, \pi(s)) \left[\gamma R(s') + \sum_{s''} P(s''|s', \pi(s')) \gamma^2 V^\pi(s'') \right] \quad (\text{b})$$

$$V^\pi(s) = R(s) + \sum_{s'} P(s'|s, \pi(s)) \left[\gamma V^\pi(s') + \sum_{s''} P(s''|s', \pi(s')) \gamma^2 V^\pi(s'') \right] \quad (\text{c})$$

$$V^\pi(s) = \sum_{s'} P(s'|s, \pi(s)) \left[\gamma R(s') + V^\pi(s') + \sum_{s''} P(s''|s', \pi(s')) \gamma^2 V^\pi(s'') \right] \quad (\text{d})$$

$$V^\pi(s) = \sum_{s'} P(s'|s, \pi(s)) \left[R(s) + \frac{1}{2} \gamma V^\pi(s') + \frac{1}{2} (\gamma R(s') + \sum_{s''} P(s''|s', \pi(s')) \gamma^2 V^\pi(s'')) \right] \quad (\text{e})$$

שאלה 3 [27 נק']

(א) בהינתן בסיס הידע $KB = (A \vee B) \wedge (\neg C \vee \neg D \vee E)$,

• האם $KB \models (A \vee B)$?

תשובה (כן / לא)	נימוק
כן	בסיס הידע הוא קוניונקציה בין השאילתא לנוסחא נוספת, ולכן כל המודלים של בסיס הידע הם בפרט המודלים של השאילתא.

• האם $KB \models (A \vee B) \wedge (\neg D \vee E)$?

תשובה (כן / לא)	נימוק
לא	$A = 1, B = 1, C = 0, D = 1, E = 0$ הינו מודל של בסיס הידע אך אינו מודל של השאילתא (השאילתא מתקבלת מבסיס הידע ע"י הורדה של $\neg C$ מתוך הדיסיונקציה השנייה, מה שהופך אותה למאולצת יותר).

• האם $KB \models (A \vee B \vee C) \wedge ((B \wedge C \wedge D) \rightarrow E)$?

תשובה (כן / לא)	נימוק
כן	$((B \wedge C \wedge D) \rightarrow E)$ שקול לוגית ל- $(\neg B \vee \neg C \vee \neg D \vee E)$, ולכן שתי הפסוקיות של השאילתא מהוות החלשות של האילוצים על מרחב המודלים המושרים ע"י שתי הפסוקיות של בסיס הידע.

(ב) נבחן את לוגיקת תחשיב הפסוקים מעל קבוצת המשתנים הבולאניים A,B,C,D. כמה מודלים יש לכל אחת מהנוסחאות הבאות?

נוסחה	$B \vee C$	$\neg A \vee \neg B \vee \neg C \vee \neg D$	$(A \rightarrow B) \wedge A \wedge \neg B \wedge C \wedge D$
תשובה	12	15	0

(ג) לפי דעותיהם של הפרשנים הפוליטיים, "מפלגה חדשה" (N) "יכולה להיכנס לכנסת" (E) אם ורק אם היא "דוגלת בקידום שיוויון בנטל" (A). אילו מהנוסחאות הבאות מייצגות את הטענה הזו במדויק?

נוסחה	תשובה	נימוק (אם התשובתכם היא "לא")
$(N \wedge E) \leftrightarrow A$	לא	מפלגה ש"דוגלת בקידום שיוויון בנטל" לא חייבת להיות "מפלגה חדשה"
$N \rightarrow (E \leftrightarrow A)$	כן	
$N \rightarrow ((A \rightarrow E) \vee \neg E)$	לא	הנוסחא היא טאוטולוגיה (נכונה תחת כל השמה), והביטוי שלנו ממש לא כזה.

שאלה 4 [14 נק']

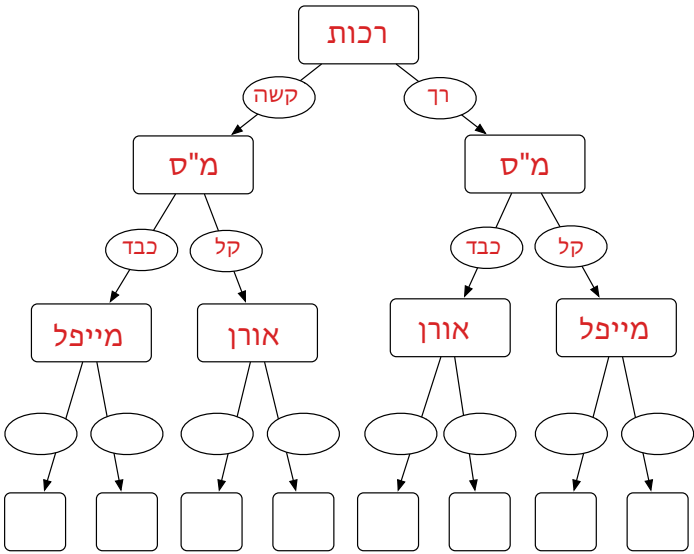
רובוט שעובד בנגריה צריך ללמוד להבדיל בין קרשים מעץ אורן לקרשים מעץ מייפל. לצורך הלמידה נבחר המודל של עצי החלטה, ולרובוט הוצגו שמונה דוגמאות כלהלן:

דוגמא	משקל סגולי	טקסטורה	רכות	סוג העץ
1	כבד	חלק	קשה	מייפל
2	כבד	מחוספס	קשה	מייפל
3	כבד	חלק	קשה	מייפל
4	קל	מחוספס	רך	מייפל
5	קל	מחוספס	קשה	אורן
6	כבד	חלק	רך	אורן
7	כבד	מחוספס	רך	אורן
8	כבד	חלק	רך	אורן

(א) איזו תכונה תיבחר להיות שורש העץ ע"י האלגוריתם לבניית עצי החלטה שבוחר בצורה חמדנית את התכונה אשר ערכיה ממקסמים את הורדת האנטרופיה (כלומר האלגוריתם שנלמד בכיתה)?

- (a) משקל סגולי
- (b) טקסטורה
- (c) **רכות**
- (d) האלגוריתם יצטרך לבחור בצורה שרירותית בין שתי תכונות
- (e) שלוש התכונות שקולות מבחינת הורדת האנטרופיה

(ב) הציגו את עץ ההחלטה אשר יכול להיבנות ע"י האלגוריתם מסעיף א'. המלבנים מיועדים לתכונות והחלטות, האליפסות מיועדות לערכי התכונות. (הערה: הקודקודים המיותרים חייבים להישאר ריקים!)



שאלה 5 [12] נק'

ידוע לנו שתכנון STRIPS הינה בעיה NP-קשה. הוכיחו או הפריכו שתכנון STRIPS המוגבל לבעיות עם אטום מטרה יחיד הינה בעיה NP-קשה.

כל בעיית STRIPS $\Pi = \langle P, A, I, G \rangle$ ניתנת לתרגום בזמן פולינומיאלי לבעיית STRIPS שקולה $\Pi = \langle P', A', I', \{g\} \rangle$ עם אטום מטרה יחיד (ואת התרגום הזה ראינו בכיתה):

$$P' = P \cup \{g\}$$

$$I' = I$$

$$A' = A \cup \{a_g = \langle G, \{g\}, \emptyset \rangle\}$$