הפקולטה להנדסת תעשיה וניהול
הטכניון

תאריך הבחינה: 11.03.2012
שם המרצה: פרופ/ח כרמל דומשלק

יסודות בינה מלאכותית ויישומיה
מבחן מועד ב', סמסטר א'

משך המבחן	3 שעות
סה"כ הניקוד במבחן	101
חומר עזר	דף נוסחאות דו-צדדי + מחשבון ללא יכולות תכנות
הוראות מיוחדות	את התשובות יש לספק **אך ורק** בטופס המבחן. מחברת הטיוטה לא תיבדק כלל! יש להגיש את דף הנוסחאות יחד עם המבחן.

שאלה 1 [16 נק' (10 + 6)]

(א) להלן חמישה בסיסי ידע בתחשיב הפסוקים. לגבי כל אחד מהם, בדקו האם הוא ניתן לייצוג שקול ע"י בסיס ידע Horn, ואם כן, הציגו את אותו בסיס ידע Horn השקול.
הערה: יש להראות רק את התשובות הסופיות, ולא את פיתוחי הביניים.

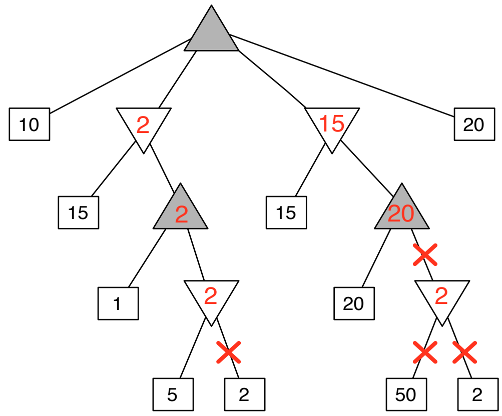
בסיס ידע	ניתן לייצוג שקול בלוגיקת Horn	בסיס ידע שקול בלוגיקת Horn
$(A \wedge B) \rightarrow C$	כן / לא	$(A \wedge B) \rightarrow C$
$(\neg A \wedge \neg B) \rightarrow C$	כן / לא	
$(A \vee B) \rightarrow C$	כן / לא	$(A \rightarrow C) \wedge (B \rightarrow C)$
$(A \vee B) \rightarrow C \wedge D$	כן / לא	$(A \rightarrow C) \wedge (B \rightarrow C) \wedge (A \rightarrow D) \wedge (B \rightarrow D)$
$(A \vee B) \rightarrow C \vee D$	כן / לא	

(ב) בתחשיב הפסוקים מעל הפרופוזיציות $\{A, B, C, D, E\}$, ציינו את מספר המודלים (השמות מספקות) לביטוי $A \leftrightarrow B \leftrightarrow C$.

16

שאלה 2 [24 נק']

להלן עץ minimax רגיל, כמו שהכרנו בכיתה. (קדקודי החלטה כהים הם קדקודי max.)



(א) מהו הערך minimax של השורש?

20

(ב) על הציור של העץ עצמו, סמנו את כל הקדקודים (הן פנימיים והן עלים) אשר לא יבוקרו תחת $\alpha-\beta$, בהנחה שהבנים של כל קדקוד נסרקים משמאל לימין.

(ג) האם יש סדר סריקה אחר של הבנים של השורש אשר מוביל לקיטום גדול יותר של קדקודים ע"י $\alpha-\beta$? אם תשובתכם חיובית, ציינו את הסדר הרלוונטי של הקדקדים.

כן / לא	למשל, קודם 20, אחר כך 15, 10, 2.

להבדיל מהסעיפים הקודמים, כאן נבחן "עצי max", בהם כל הקדקודים בעץ הם קדקודי max. (עצי חיפוש כאלה מתאימים לפתרון בעיות עם מספר סוכנים שלא מתחרים זה בזה אלא מונעים ע"י מטרה משותפת.)

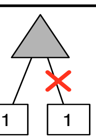
(ד) בהנחה שהערכים של כל העלים בעץ הם סופיים אבל לא ידוע מראש חסם על הערכים האלה, האם קיטום קדקודים בסגנון $\alpha-\beta$ אפשרי בעצי max? ספקו דוגמא או נמקו בקצרה מדוע לא.

התשובה היא לא. בכל שלב של החיפוש, הענף הבא של העץ יכול להכיל ערך גבוה מכל הערכים שנראו עד כה.

(ה) בהנחה שידוע שכל הערכים של העלים בעץ הם חיוביים, האם קיטום קדקודים בסגנון $\alpha-\beta$ אפשרי בעצי max? ספקו דוגמא או נמקו בקצרה מדוע לא.

התשובה היא לא. חסם תחתון על הערכים בעצי max לא משנה את הטיעון שבתשובה לסעיף ד'.

(ו) בהנחה שידוע שכל הערכים של העלים בעץ הם בתחום [0,1], האם קיטום קדקודים בסגנון $\alpha-\beta$ אפשרי בעצי max? ספקו דוגמא או נמקו בקצרה מדוע לא.



שאלה 3 [15 נק']

כל סעיפי השאלה הזו מתיחסים לבעיית התכנון הדטרמיניסטי (המאויירת להלן) שבה סוכן בודד צריך להזיז ארבעה רובוטים מקונפיגורציה התחלתית מפינה אחת של המבוך $N \times M$ לקונפיגורצית מטרה בפינה אחרת של המבוך. בכל יחידת זמן, כל אחד מהרובוטים יכול להיות מוזז תא אחד בכל אחד מהכיוונים או להשאר במקום: {צפון, דרום, מזרח, מערב, במקום}. אחרי שההזזה ה*סימולטנית* של כל הרובוטים מבוצעת, שני רובוטים לא יכולים להימצא באותו תא של המבוך, אך במהלך ההזזה, שני רובוטים סמוכים יכולים להתחלף במיקומיהם. להמחשה, אחד הבנים של המצב ההתחלתי מוצג במרכז האיור. העלות של כל הזזה סימולטנית של הרובוטים היא 1, לא משנה כמה רובוטים אכן שינו את מיקומיהם באותה הזזה.

#	#	#	#	#	#	#	#	#	#
#									#
#		#		#		#	#		#
#		#				#			#
#		#			#	#			#
#		#	#		#				#
#			#				C	D	#
#			#		#	#	A	D	#
#	#	#	#	#	#	#	#	#	#

מצב מטרה (יחיד)

#	#	#	#	#	#	#	#	#	#
#									#
#		#		#		#	#		#
#		#				#			#
#		#			#	#			#
#	B	#	#		#				#
#	D	C	#						#
#	A		#		#	#			#
#	#	#	#	#	#	#	#	#	#

אחד הבנים של מצב התחלה

#	#	#	#	#	#	#	#	#	#
#									#
#		#		#		#	#		#
#		#				#			#
#		#			#	#			#
#		#	#		#				#
#	B	D	#						#
#	A	C	#		#	#			#
#	#	#	#	#	#	#	#	#	#

מצב התחלה

(א) מהן התכונות של בעיית תכנון זו(לסוג כזה של מבוכים, לא למבוך הספציפי שבאיור)?
הערה: יש לספק חסמים כמה שיותר הדוקים (קטנים).

גודל מקסימלי של מרחב המצבים	$n^2(n^2-1)(n^2-2)(n^2-3)$
דרגת סיעוף מקסימלית	5^4

(ב) הציעו היוריסטיקה קבילה לא טריוויאלית, *ספציפית לבעייה שבשאלה*. (היוריסטיקה טריוויאלית היא, למשל, היוריסטיקה שמחזירה מספר קבוע.)

למשל, מרחק מנהטן מקסימלי מאיזשהו רובוט בקבוצה למיקום היעד שלו.

(ג) הקיפו את <u>כל</u> האלגוריתמים להלן אשר מובטח שיחזירו תכנית *למשפחה הזאת של בעיות תכנון*. • DFS. • Random walk • A* עם היוריסטיקה לא קבילה. • GBFS עם היוריסטיקה קבילה. • Hill climbing עם היוריסטיקה קבועה לכל המצבים שהם לא מצב המטרה.	**(ד)** אם h_1 ו-h_2 הן היוריסטיקות קבילות, לאילו מההיוריסטיקות הבאות מובטחת קבילות? • בחירה אקראית בהתפלגות אחידה מ-$\{h_1, h_2\}$. • h_1/h_2 • $\min(h_1, h_2)$ • $\max(h_1, h_2) + \min(h_1, h_2)$ • $(h_1 + h_2)/2$

(ה) מתי בשימוש באלגוריתם A* נעדיף היוריסטיקה שפותחת יותר קדקודים על היוריסטיקה שפותחת פחות קדקודים, אם בכלל?
<u>תשובה מנומקת במשפט אחד בלבד.</u>

העדפה כזו יכולה להיות הגיונית בסיטואציות מסויימות בהן (א') אנו מוגבלים בסה"כ זמן חיפוש ו-(ב) חישוב ההיוריסטיקה המדוייקת יותר (שמביאה לפתיחת פחות קדקודים) הינו הרבה יותר איטי מזה של ההיוריסטיקה המדוייקת פחות.

<u>שאלה 4 [10 נק']</u>

הוכיחו שפונקצית הערכה h_{FF} אינה קבילה, כלומר היחס $h_{FF}(s) \leq h^*(s)$ לא בהכרח מתקיים.

כל דוגמא נגדית מהווה הוכחה אך הכי קל היה להראות דוגמא נגדית דרך בעיית תכנון שכבר חסרת אפקטי הורדה. למשל:

$$
\begin{aligned}
P &= \{p_1, \ldots, p_n\} \\
I &= \emptyset \\
G &= P \\
A &= \{a_1, \ldots, a_n, a_{n+1}\} \\
a_i &= \begin{cases} \{\emptyset, \{p_i\}, \emptyset\}, & 1 \leq i \leq n \\ \{\{p_1\}, \{p_2, \ldots, p_n\}, \emptyset\}, & i = n+1 \end{cases}
\end{aligned}
$$

שאלה 5 [16 נק']

בשאלה זו אתם מתבקשים לקודד ב-STRIPS את בעיית התכנון "דילמת הנהר". (אל תנסו להזכר. את השם הזה לבעייה המצאתי כרגע.) הבעייה היא כדלהלן.

איש אחד צריך להעביר מהגדה המזרחית של הנהר לגדה המערבית שלו זאב, עז, וכרוב, אך לסירה של האיש יש מקום רק לאיש עצמו ועוד אחד משלושת האובייקטים האלה. אם האיש לוקח אתו את הכרוב, הזאב אוכל את העז. אם האיש לוקח אתו את הזאב, העז אוכל את הכרוב. במילים אחרות, רק כשהאיש נוכח, בטוחים הכרוב והעז מהחומדים אותם.

אטומים	{manE,manW,wolfE,wolfW,goatE,goatW,cabbageE,cabbageW}
מצב התחלתי	{manE,wolfE,goatE,cabbageE}
מטרה	{wolfW,goatW,cabbageW}

שם הפעולה	תנאי קדם	אפקט הוספה	אפקט הורדה
moveEW	manE	manW	manE
moveCabbageEW1	manE,cabbageE,goatW	cabbageW,manW	cabbageE,manE
moveCabbageEW2	manE,cabbageE,wolfW	cabbageW,manW	cabbageE,manE
moveCabbageEW3	manE,cabbageE	cabbageW,manW	cabbageE,manE,goatE
moveWolfEW1	manE,wolfE,cabbageW	manW,wolfW	manE,wolfE
moveWolfEW2	manE,wolfE,goatW	manW,wolfW	manE,wolfE
moveWolfEW3	manE,wolfE	manW,wolfW	manE,wolfE,cabbageE
moveGoatEW	manE,goatE	manW,goatW	manE,goatE
+ אופרטורים סימטריים מ-W ל-E.			

שאלה 6 [20 נק']

שואב האבק הרובוטי (שא"ר) שלנו נועד לניקיון הבית והוא מתפקד על סוללה נטענת. בכול נקודה בזמן השא"ר יכול לבצע אחת משלוש הפעולות הבאות: ניקיון הבית (clean), המתנה (wait), או טעינת הסוללה (recharge). החיישנים של שא"ר מאפשרים לו לחוש את מצב הסוללה רק בצורה מאד גסה, המאפשרת להבדיל רק בין שתי רמות הטעינה: גבוהה (high) ונמוכה (low). בעצם, המצב כולו של השא"ר הפשוט שלנו מתואר ע"י משתנה בינארי שתופס את מצב הסוללה high/low.

- אם הרובוט מבצע פעולה recharge, הסוללה חוזרת לרמת הטעינה high, והשא"ר לא מקבל שום תגמול מיידי.
- אם הרובוט ממתין (כלומר, מבצע פעולה wait), אז מצב טעינה של הסוללה נשאר כפי שהיה, והשא"ר מקבל תגמול מיידי R_{wait}.
- אם הרובוט מבצע פעולה clean, אזי התגמול המיידי על הפעולה תלוי במצב הסוללה לפני תחילת ביצוע הפעולה. אם מצב הסוללה high, פעולת clean משנה אותו ל-low בהסתברות 1/3. אם מצב הסוללה low, אז הסוללה נגמרת בהסתברות 1/2. אם הסוללה לא נגמרת, השא"ר מקבל תגמול מיידי R_{clean}. אם הסוללה כן נגמרת, אנחנו חייבים לאסוף את הרובוט ולטעון אותו בעצמנו. במקרה כזה, הסוללה חוזרת למצב high, והשא"ר מקבל תגמול מיידי של 10-.

מבחינת הביצועים, אנחנו רוצה שהשא"ר שלנו יהיה מונע ע"י ערכים תחת אופק אינסופי, עם מקדם הפליית עתיד (discount factor) $\gamma = 0.9$.

<u>הערה:</u> שימו לב שהתגמולים של הרובוט תלויים לא רק במצב הנוכחי, אלא בשלשה: מצב נוכחי, פעולה, ומצב הבא.

(א) בהנחה ש: $0 \leq R_{\text{wait}} \leq R_{\text{clean}}$, אלו (אולי כמה!) מהאופציות להלן יכולות להוות מדיניות אופטימלית לשא"ר? נמקו בקצרה מאד את תשובותיכם.

תמיד לבצע clean, לא משנה מהו המצב של הסוללה.

כן / לא	נימוק: כאשר R_{clean} יהיה מספיק גדול.

תמיד לבצע recharge, לא משנה מהו המצב של הסוללה.

כן / לא	נימוק: אם $R_{\text{wait}} = R_{\text{clean}} = 0$

בצע clean אם מצב הסוללה הוא high. אחרת (אם מצב הסוללה הוא low), בצע recharge.

כן / לא	נימוק: לערכים סבירים של התגמולים, זאת תהיה המדינות האופטימלית שהיינו חושבים עליה גם בלי המתמטיקה.

בצע recharge אם מצב הסוללה הוא high. אחרת (אם מצב הסוללה הוא low), בצע clean.

כן / לא	נימוק: כי אם ביצוע clean הוא אופטימלי במצב סוללה low, אז קל וחומר הוא אופטימלי במצב סוללה high.

(ב) רשמו את משוואת בלמן, המבטאת את הערך האופטימלי של המצב low בשפת הערכים האופטימליים של כלל המצבים. במילים אחרות, אם

$$V^*(low) = f\left(V^*(low), V^*(high), R_{\text{wait}}, R_{\text{clean}}\right),$$

אז מהי אותה פונקציה f?

$$V^*(low) = \max\left[\gamma V^*(high), R_{\text{wait}} + \gamma V^*(low), \frac{1}{2}(R_{\text{clean}} + \gamma V^*(low)) + \frac{1}{2}(-10 + \gamma V^*(high))\right]$$

(ג) נניח ש: $R_{\text{wait}} = 1, R_{\text{clean}} = 3$. אם הערכים של כל המצבים האפשריים של השא"ר מאותחלים ל-0, מהם יהיו אחרי איטרציה אחת של אלגוריתם value iteration?

$$V(low) = 1, V(high) = 3$$

(ד) תחת אותה ההנחה של $R_{\text{wait}} = 1, R_{\text{clean}} = 3$ כמו בסעיף הקודם, נבחן למידת Q-learning עם מקדם "קצב למידה" (learning rate) $\alpha = 0.2$. הלמידה מתחילה מכל הערכים $Q(s,a)$ מאותחלים ל-0, ומתבססת על סידרת התנסויות של מצבים, פעולות, ותגמולים כדלהלן:

high, clean, +3, high, clean, +3, low, clean, +3, low, clean, -10

t

הציגו את פונקציית $Q(s,a)$ במלואה, כפי שהיא מתקבלת לאחר כל אחד מארבעת העידכונים שלה בלמידת Q-learning. את התשובות יש לספק בטבלא שלהלן.

	$Q(low, recharge)$	$Q(low, wait)$	$Q(low, clean)$	$Q(high, recharge)$	$Q(high, wait)$	$Q(high, clean)$
אתחול	0	0	0	0	0	0
עדכון 1	0	0	0	0	0	0.6
עידכון 2	0	0	0	0	0	1.08
עידכון 3	0	0	0.6	0	0	1.08
עידכון 4	0	0	-1.33	0	0	1.08