

הפקולטה להנדסת תעשייה וניהול
הטכניון

תאריך הבחינה: 15.02.2013
שם המרצה: פרופ/ח כרמל דומשלק

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

משך המבחן	3 שעות
סה"כ הניקוד במבחן	100 (+ 3 נק' בonus על עבודה עצמאית)
חומר עזר	מחשבון ללא יכולות תכנות
הוראות מיוחדות	את התשובות יש לספק אך ורק בטופס המבחן. מחברת הטיוטה לא תיבדק כלל!

מס' תעודת זהות

בonus

BFS (גרסת "חיפוש עץ")	GBFS (גרסת "חיפוש גרף")
<pre> 1 queue := new fifo-queue 2 queue.push-back(make-root-node(init())) 3 while not queue.empty(): 4 σ = queue.pop-front() 5 if is-goal(state(σ)): 6 return extract-solution(σ) 7 for each ⟨o, s⟩ ∈ succ(state(σ)): 8 σ' := make-node(σ, o, s) 9 queue.push-back(σ') 10 return unsolvable </pre> <p style="text-align: center;">ALG2</p>	<pre> 1 open := new min-heap ordered by (σ ↦ h(σ)) 2 open.insert(make-root-node(init())) 3 closed := ∅ 4 while not open.empty(): 5 σ = open.pop-min() 6 if state(σ) ∉ closed: 7 closed := closed ∪ {state(σ)} 8 if is-goal(state(σ)): 9 return extract-solution(σ) 10 for each ⟨o, s⟩ ∈ succ(state(σ)): 11 σ' := make-node(σ, o, s) 12 if h(σ') < ∞: 13 open.insert(σ') 14 return unsolvable </pre> <p style="text-align: center;">ALG1</p>
A* (גרסת "חיפוש גרף" עם פתיחת קדקודי חיפוש מחדש)	
<pre> 1 open := new min-heap ordered by (σ ↦ f(σ) = g(σ) + h(σ)) 2 open.insert(make-root-node(init())) 3 closed := ∅ 4 distance := ∅ 5 while not open.empty(): 6 σ = open.pop-min() 7 if state(σ) ∉ closed or g(σ) < distance(state(σ)): 8 closed := closed ∪ {state(σ)} 9 distance(state(σ)) := g(σ) 10 if is-goal(state(σ)): 11 return extract-solution(σ) 12 for each ⟨o, s⟩ ∈ succ(state(σ)): 13 σ' := make-node(σ, o, s) 14 if h(σ') < ∞: 15 open.insert(σ') 16 return unsolvable </pre> <p style="text-align: center;">ALG3</p>	

(א) מהן שורות הקוד באלגוריתם ALG1 שאם נעבירן (למקומות הנכונים של) האלגוריתם ALG2, נקבל BFS בגרסת "חיפוש גרף"? (יש לתת רק את רשימת מספרי השורות הרלוונטיות).

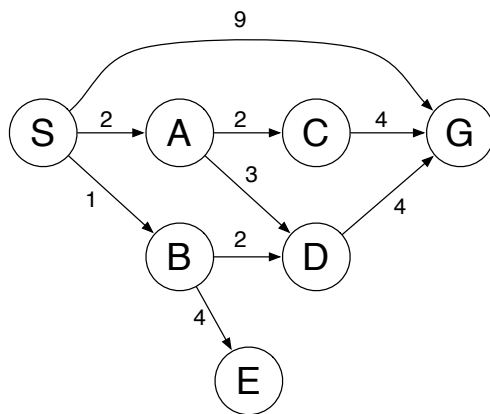
3,6,7

(ב) מהן שורות הקוד שהורדתן מאלגוריתם ALG3 תהפוך אותו ל-A* בגרסת "חיפוש עץ"? (יש לתת רק את רשימת מספרי השורות הרלוונטיות).

3,(4),7,8,(9)

* השורות שמספריהן מופיעות בסוגריים לא נדרשות לתשובה אך גם לא גרעו מנכונותה.

הסעיפים הבאים מתייחסים לבעיית חיפוש על גרף המצבים הנתון להלן. המצב ההתחלתי הוא S, ומצב המטרה היחיד הוא G. שימו לב שהסעיפים הבאים מתייחסים הן לאלגוריתמי "חיפוש עץ" והן לאלגוריתמי "חיפוש גרף". כמו כן, בכל האלגוריתמים, אם יש צורך בשבירת השוויון בבחירת המצבים לטיפול, אזי השוויון נשבר ע"ב שמות המצבים, לפי הסדר המילוני.



h_2

S	A	B	C	D	E	G
6	0	6	4	1	10	0

(ג) מהו המסלול שיוחזר ע"י BFS בגרסת "חיפוש עץ"?

S→G

(ד) מהו המסלול שיוחזר ע"י A* בגרסת "חיפוש עץ" עם הוריסטיקה $h_1(state) = 0$?

S→B→D→G

(ה) מהו המסלול שיוחזר ע"י A* בגרסת "חיפוש גרף" עם הוריסטיקה $h_1(state) = 0$?

S→B→D→G

(ו) מהם המצבים שנכנסים לרשימת closed בהפעלת A* בגרסת "חיפוש גרף" עם הוריסטיקה $h_1(state) = 0$? שימו לב: את רשימת המצבים יש לתת לפי הסדר בו הם נכנסים לרשימת closed.

S, B, A, D, C, E, G

(ז) מהו המסלול שיוחזר ע"י GBFS בגרסת "חיפוש עץ" עם הוריסטיקה h_2 שמתוארת לעיל?

S→G

(ח) מהו המסלול שיוחזר ע"י A* בגרסת "חיפוש עץ" עם הוריסטיקה h_2 ?

S→B→D→G

(ט) מהו המסלול שיוחזר ע"י A* בגרסת "חיפוש גרף" עם הוריסטיקה h_2 (שימו לב: בזהירות!)?

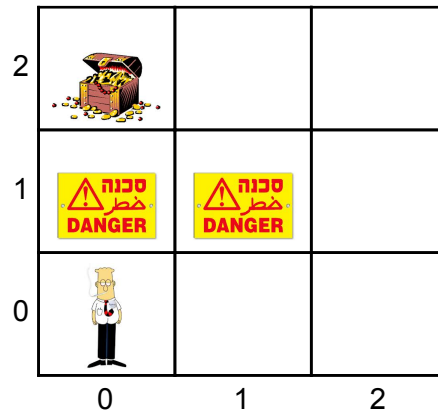
S→A→C→G

למה? כי ההוריסטיקה אינה עקבית (consistent).

הערה: במחשבה שניה, הגעתי למסקנה שהניסוח של השאלה היה מטעה: הכוונה הייתה ל-A* בגרסת "חיפוש גרף" ללא פתיחת מצבים מחדש (ולא ל-ALG3), אך זה לא הובהר טוב. לכן קיבלתי גם את התשובה (הלא מי יודע מה מענינת) S→B→D→G.

שאלה 2 [30 נק']

הסוכן שלנו מ"עולם הוומפוס (Wumpus)" החליט להעביר קצת זמן בהרפתקאה הרבה יותר רגועה על "אי האוצר". הסוכן התכונן היטב וצייד את עצמו במפת האי אשר מראה לו הן את מיקום האוצר והן את מיקומי הסכנות:



מכל תא לא מסומן (כלומר, תא שאין בו לא אוצר ולא סכנה), הסוכן יכול לבצע את התזוזות בארבעת הכיוונים הרגילים (N - צפון, S - דרום, W - מערב, E - מזרח), וכל הפעולות הללו הן דטרמיניסטיות. לעומת זאת, אם הסוכן נמצא בתא מסומן, אזי הפעולה היחידה שהוא יכול להפעיל היא "הטסה" (שנסמן אותה ב-X) שמעבירה את הסוכן מהאי למקום מבטחים בו ההרפתקאה נגמרת. הפרסים היחידים שהסוכן מקבל מזוהים עם התאים המסומנים: הסוכן מקבל פרס שלילי של -64 בתא עם סכנה ופרס חיובי של +128 בתא עם האוצר. שימו לב: אין לסוכן פעולות מהסוג "לא לעשות כלום".

(א) מהם הערכים האופטימליים, V^* , של כל תא על האי תחת קבלת החלטות עם אופק לא חסום ומקדם הפליית הפרסים העתידיים (discount factor) $\gamma = 0.5$? (ציינו את הערכים ישירות בטבלה להלן).

2	128	64	32
1	-64	-64	16
0	2	4	8
	0	1	2

(ב) מהי המדיניות האופטימלית תחת אותה קבלת החלטות עם אופק לא חסום ומקדם הפלייט הפרסים העתידיים $\gamma = 0.5$? (צינו את פעולות המדיניות ישירות בטבלה להלן).

2	X	W	W
1	X	X	N
0	E	E	N
	0	1	2

בסעיפים ג' ו-ד', נקרא בשם π_0 למדיניות שבתשובה לסעיף ב'. הסוכן שלנו חושד שהמפה שלו לא הכי מעודכנת, ולכן הוא מחליט לבצע קצת למידת Q-learning במטרה להכיר את האי. מכיוון שהסוכן מאמין ש- π_0 היא די קרובה למדיניות האופטימלית, הוא מחליט לבצע Q-learning תוך התנהגות לפי מדיניות ε -אקראית לפי π_0 : בהסתברות ε הסוכן בוחר מהפעולות שעומדות לרשותו בצורה אקראית לפי התפלגות אחידה, ובהסתברות $(1 - \varepsilon)$ הוא מבצע את הפעולה המוכתבת ע"י π_0 . נקרא למדיניות הזאת π_ε .

להבדיל ממדיניות דטרמיניסטיות כמו π_0 , מדיניות π_ε היא דוגמא למדיניות סטוכסטית אשר מגרילה את המלצתה מתוך התפלגות מעל פעולות. למדיניות סטוכסטית π , נסמן ב- $\pi(s, a)$ את ההסתברות של בחירת הפעולה a במצב s .

(ג) נסחו את משוואת בלמן להערכת המדיניות הנתונה π , המעודכנת ממה שלמדנו בכיתה להעריך מדיניות סטוכסטיות. (יש לתת תשובה כללית, ותשובה ספציפית למדיניות π_ε לעיל).

$$V^\pi(s) = R(s) + \gamma \sum_a \pi(s, a) \sum_{s'} T(s, a, s') V^\pi(s')$$

$$V^{\pi_\varepsilon}(s) = R(s) + \gamma \left[(1 - \varepsilon) V^{\pi_\varepsilon}(\text{succ}(s, \pi_0(s))) + \frac{\varepsilon}{|A(s)|} \sum_{a \in A(s)} V^{\pi_\varepsilon}(\text{succ}(s, a)) \right]$$

(ד) בהנתן $\varepsilon = 0.1$, איזה מהיחסים להלן מתקיים לכל המצבים אם המפה של הסוכן נכונה?

- i. $V^{\pi_0} \geq V^{\pi_\varepsilon}$
- ii. $V^{\pi_0} = V^{\pi_\varepsilon}$
- iii. $V^{\pi_0} \leq V^{\pi_\varepsilon}$

בסעיפים ה' ו-ו', מסתבר שהמפה של הסוכן היא לרוב נכונה, למעט זה שחלק מהתאים שסומנו במפה כמסוכנים הינם בפועל תאים רגילים! לכן, בתחילת ה-Q-learning, הסוכן חווה את ההתנסויות הבאות:

[(0,0), 0, N, (0,1), 0, N, (0,2), 128, X, done]
 [(0,0), 0, N, (0,1), 0, N, (0,2), 128, X, done]
 [(0,0), 0, N, (0,1), 0, E, (1,1), -64, X, done]

(ה) מה יהיו ערכי ה-Q שהסוכן ילמד בתום כל האפיזודות הללו בהנחה שהוא איתחל את כל ערכי ה-Q ל-0, ומשתמש במקדם "קצב למידה" $\alpha = 1$ (learning rate).
 (שימו לב: יש צורך לציין רק ערכי ה-Q השונים מ-0.)

$Q((0, 2), X)$	=	128
$Q((0, 1), N)$	=	64
$Q((0, 0), N)$	=	32
$Q((1, 1), X)$	=	-64
$Q((0, 1), E)$	=	0

(ו) עם הידע שיש לכם על מבנה האי ובהנתן האפיזודות שהסוכן חווה, מהם הערכים האופטימליים האמיתיים, V^* , של כל תא על האי תחת קבלת החלטות עם אופק לא חסום ומקדם הפליית הפרסים העתידיים $\gamma = 0.5$? (ציינו את הערכים ישירות בטבלה להלן.)

2	128	64	32
1	64	-64	16
0	32	16	8
	0	1	2

שאלה 3 [15 נק']

לוגיקת "תחשיב הפסוקים עם נשמה" מרחיבה את לוגיקת תחשיב הפסוקים עם קשר \odot אשר מוגדר כדלקמן:

α	β	$\alpha \odot \beta$
T	T	F
T	F	T
F	T	T
F	F	T

(א) האם כלל היסק $\beta \models \neg(\alpha \odot \beta)$ הינו נאות ללוגיקה שלנו? נמקו (הוכיחו) את תשובתכם.

אכן, כלל היסק זה נאות ללוגיקה שלנו.
הוכחה בצורה ישירה: טבלת אמת שהמודלים של $\neg(\alpha \odot \beta)$ מוכלים בקבוצת המודלים של β .

(ב) בהנתן בסיס הידע

$$KB = (p \odot q) \wedge (p \vee q) \wedge (p \vee r) \wedge (s \vee q) \wedge (s \vee \neg r) \wedge (\neg r)$$

הוכיחו בעזרת רזולוציה ש: $KB \models (s)$.

דבר ראשון, שימו לב שהקשר החדש שלנו לא באמת מרחיב את יכולת הביטוי של תחשיב הפסוקים, ואפשר להעלים אותו כלל ע"י שימוש בשקילות לוגית:

$$(\alpha \odot \beta) \equiv \neg(\alpha \wedge \beta) \equiv (\neg\alpha \vee \neg\beta)$$

לכן, בסיס הידע לעיל ניתן לייצוג CNF:

$$KB = (\neg p \vee \neg q) \wedge (p \vee q) \wedge (p \vee r) \wedge (s \vee q) \wedge (s \vee \neg r) \wedge (\neg r)$$

ומכאן:

$$\frac{(p \vee r) \quad (\neg r)}{(p)} \Rightarrow \frac{(\neg p \vee \neg q) \quad (p)}{(\neg q)} \Rightarrow \frac{(s \vee q) \quad (\neg q)}{(s)}$$

(ג) ספקו ביטוי שקול לוגית ל- $(p \vee q)$ אשר מורכב רק מסימנים p, q , ו- \odot (וסוגריים, לצורך הצגה ברורה).

(רמז 1: שני סימני \odot לא יספיקו. רמז 2: מה זה $(p \odot p)$?)

$(p \odot p) \odot (q \odot q)$

שאלה 4 [10 נק']

(א) נתונה בעיית תכנון STRIPS מעל N אטומים ו- K פעולות, ולכל פעולה בדיוק M אטומים בתנאי קדם. כמה צלעות יש בגרף המצבים המושרה ע"י הבעיה הזו? (הערה: יכולים להיות מספר צלעות בין זוג קדקדים).

$$K2^{N-M} \quad (a)$$

$$K2^M \quad (b)$$

$$2^K 2^{N-M} \quad (c)$$

$$2^K 2^M \quad (d)$$

(e) התשובה תלויה באפקטים של האופרטורים

(ב) הוכיחו שפונקציית הערכה h_{\max} לבעיות STRIPS אינה קבילה, כלומר היחס $h_{\max}(s) \leq h^*(s)$ לא בהכרח מתקיים.

להוכחת הטענה מספיק להוכיח ש: $h_{\max}(s) \leq h^+(s)$. תהי $\pi = \langle a_1^+, \dots, a_m^+ \rangle$ תכנית מוחלשת אופטימלית מ- s . לפי הגדרת גרף הנגישות (relaxed planning graph - RPG) מ- s , לכל $1 \leq i \leq m$, a_i^+ שייכת לשכבת הפעולות ה- i של הגרף, ואיחוד אפקטי ההוספה של a_1^+, \dots, a_i^+ שייך לשכבת האטומים ה- i של הגרף. לפי כך, מכון ש- π הינה תכנית מוחלשת, כל אטומי המטרה נמצאים בשכבת האטומים ה- m של הגרף. אך לפי הגדרה, $h_{\max}(s)$ שווה למספר שכבת האטומים הראשונה של גרף הנגישות אשר מכילה את כל אטומי המטרה, ולכן $h_{\max}(s) \leq m = h^+(s)$.