

תרגיל בית 1

מרחבי חיפוש

מטרות התרגיל

- נתמודד עם בעיות פרקטיות ותיאורטיות של חיפוש במרחבי מצבים.
- נתרגל את הנלמד בהרצאות ובתרגולים.
- נתנסה בתכנות ב-python לפתרון בעיות פרקטיות.

הנחיות כלליות

- **תאריך הגשה 2/3 יום שבת, בשעה 23:59.** (לשנות מועד ולפתוח פיאצה)
- את המטלה יש להגיש בזוגות בלבד.
- יש להגיש מטלות מוקלדות בלבד בעברית או באנגלית. פתרונות בכתב יד לא ייבדקו.
- ניתן לשלוח שאלות בנוגע לתרגיל בפיאצה בלבד.
- המתרגל האחראית על תרגיל: **שאדי דיאב**.
- בקשות דחיה מוצדקות (מילואים, אשפוז וכו') יש לשלוח למתרגל האחראי (**ספיר טובול**) בלבד.
- במהלך התרגיל ייתכן שנעלה עדכונים, למסמך הנ"ל – תפורסם הודעה בהתאם.
- העדכונים הינם מחייבים, ועליכם להתעדכן עד מועד הגשת התרגיל.
- שימו לב, התרגיל מהווה כ- 15% מהציון הסופי במקצוע ולכן העתקות תטופלנה בחומרה!
- ציון המטלה יורכב מהגורמים הבאים:
 - 65% - המסמך היבש.
 - 35% - הקוד המוגש.
- אנו יודעים שעבור חלקכם זו התנסות ראשונה בכתיבת קוד בפיתון ותרגיל זה מתוכנן בהתאם לכך.
- שימו לב שלא יענו שאלות בסגנון: "איך מוצאים את עלות הפתרון שהוחזר?" / "איך ניגשים למפות הכבישים מתוך המימוש של הפונק' ההיא?" / "באיזה שדה שמור ה...?" וכדומה.
- אנחנו רוצים לעודד אתכם לעיין בקוד ולמצוא פרטים אלו בכוחות עצמכם. הכרת סביבת העבודה שסיפקנו לכם והתמצאות בה הן למעשה חלק מהתרגיל.
- בתרגילי הבית בקורס הרצת הניסויים עשויה לקחת זמן רב. לכן מומלץ מאוד להימנע מדחיית העבודה על התרגיל ו/או כתיבת הדו"ח לרגע האחרון. לא תינתנה דחיות על רקע זה.
- מסמך זה כתוב בלשון זכר מטעמי נוחות בלבד, אך מתייחס לנשים וגברים כאחד.

אנחנו קשובים לפניית שלכם במהלך התרגיל ומעדכנים את המסמך הזה בהתאם. גרסאות עדכניות של המסמך יועלו לאתר. **הבהרות ועדכונים שנוספים אחרי הפרסום הראשוני יסומנו כאן בצהוב.** בנוסף, לכל עדכון יהיה מספר גרסה כדי שתוכלו לעקוב. ייתכן שתפורסמה גרסאות רבות – אל תיבהלו מכך. השינויים בכל גרסה יכולים להיות קטנים.

הנחיות לחלק היבש

1. ככלל אצבע, בהינתן שאלה ראשית ספקו את התשובה המיידית ולאחר מכן תרחיבו ותסבירו. למשל, אם שואלים מה סיבוכיות הזמן של אלגוריתם BFS תשובה תהיה " $O(b^d)$ ", מכיוון שבקרה הכי גרוע נאחסן את כל עץ החיפוש של הבעיה ב-CLOSE".

הנחיות לחלק הרטוב

1. אנו מעודדים אתכם לעבור על הקבצים המצורפים ולהבין כיצד הסביבה בנויה ובאילו פונקציות תוכלו להשתמש במימוש שלכם.
2. הקוד שלכם ייבדק בקפדנות על ידי טסטים. הטסטים יבדקו את הפתרונות המוחזרים על ידי האלגוריתמים שלכם אל מול המימוש שלנו על פני בעיות שונות. אנו מצפים ממכם (אלא אם צוין אחרת)

להחזיר את אותם ערכים בדיוק. אנחנו נבדוק את המסלול המוחזר, מספר הצמתים שפתחו ואת עלות הפתרון המוחזר. הטסטים יהיו מוגבלים בזמן אך תקבלו זמן גדול מאוד לכל טסט.
3. ספקו קוד ברור ונקי לבידוק ידנית.

מבוא ורקע

התרגיל מתפרש על פני מסמך זה והמחברת המצורפת. מומלץ לענות על השאלות לפי הסדר במסמך זה. במטלה זו נעסוק בהפעלת אלגוריתמי חיפוש על מרחבי מצבים לבעיות ניווט. מומלץ לחזור על שקפי ההרצאות והתרגולים הרלוונטיים לפני תחילת העבודה על התרגיל.

סיפור מסגרת

לקאקרוטו וגוהאן יש 5 כדורי דרקון וחסר להם שני כדורים, והם ממש צריכים אותם כדי להזמין הדרקון שן-ראן ולבקש ממנו להחזיר את החברים שלהם לחיים, לכן הם הלכו לכוכב לכת נאמיק כדי לחפש כדורי הדרקון, קאקרוטו הציע שיחפשו על הכדור דרך ה ג'י.פי.אס שלהם אבל גוהאן מסביר לקאקרוטו שיש לו חברים שלוקחים הסמסטר את קורס "מבוא לבינה מלאכותית". גוהאן מבקש ממכם לעזור לו לתכנן את המסלול הטוב ביותר כדי לאסוף כדורי הדרקון ולהגיע לקאקרוטו שמחקה לו.



שאלה 1 – מבוא (8 נק'):

השאלות בחלק זה מתבססות על הלוח "8x8" שמופיע במחברת אלא אם נכתב אחרת:

S	F	F	F	F	F	F	F
F	F	F	F	T	A	L	
T	F	F	H	F	T	F	
F	F	F	F	H	T	F	
F	A	F	H	F	F	F	
F	H	H	F	F	F	H	F
D	F	T	F	H	D	T	L
F	L	F	H	F	F	F	G

1. **רטוב:** עבורו על המחברת עד שאתם מגיעים לחלק של BFS-G ועצרו שם.
2. יבש (1 נק'): תחילה נרצה להגדיר את מרחב החיפוש כפי שנלמד בתרגול. הגדר את $(S, 0, I, G)$ עבור סביבת כדורי הדרקון. כאשר S זה מרחב המצבים, 0 , זה מרחב האופרטורים, I , זה המצב ההתחלתי ו G הוא קבוצת מצבי המטרה. מה גודל מרחב המצבים S ? הסבירו.
3. יבש (1 נק'): מה תחזיר לנו הפונקציה Domain על אופרטור 2 (UP)?
4. יבש (1 נק'): מה תחזיר לנו הפונקציה Succ על המצב ההתחלתי 0?

5. יבש (1 נק'): האם קיימים מעגלים במרחב החיפוש שלנו?
6. יבש (1 נק'): מה הוא מקדם הסיעוף בבעיה?
7. יבש (1 נק'): במקרה הגרוע ביותר, כמה פעולות ידרשו לסוכן כללי להגיע למצב הסופי?
8. יבש (1 נק'): במקרה הטוב ביותר, כמה פעולות ידרשו לסוכן כללי להגיע למצב הסופי?
9. יבש (1 נק'): עבור לוח כללי, המסלול הקל ביותר הוא המסלול שמגיע למצב מטרה שהכי קרוב למצב ההתחלתי (במונחים של Manhattan distance) אם כן, הוכיחו. אם לא, ספקו דוגמא נגדית.

שאלה 2 – Breadth First Search-G (7 נק'):

השאלות בחלק זה מתבססות על הלוח "8x8" שמופיע במחברת אלא אם נכתב אחרת.

1. **רטוב:** ממשו את אלג' BFS-G (על גרף) במחברת ע"פ ההנחיות המופיעות שם.
2. יבש (1 נק'): מה צריך להיות התנאי על גרף החיפוש (לא בהכרח בבעיית כדורי הדרכון) כך ש-BFS על גרף ו-BFS על עץ ייצרו ויפתחו צמתים זהים באותו הסדר?
3. יבש (2 נק'): עבור הלוח "4x4" שמופיע במחברת, ציירו את גרף המצבים.
4. יבש (2 נק'): נתון לוח בגודל $N \times N$. הציעו דרך להשתמש באלגוריתם BFS-G כך שיחזיר פתרון אופטימלי (עלות מינימלית) והסבירו.
☐ רמז: עליכם לספק פונקציה $T: G \rightarrow G'$ המקבלת את גרף המצבים G ויוצרת גרף חדש G' ובעזרתה למצוא את המסלול האופטימלי בגרף G .
5. יבש (2 נק'): נתון לוח בגודל $N \times N$, ללא חורים, המכיל $N^2 - 2$ משבצות רגילות (F,T,A,L) מצב התחלתי בפינה השמאלית עליונה ומצב מטרה בפינה הימנית תחתונה. כמה צמתים יפותחו וייווצרו במהלך חיפוש BFS-G? הסבירו?

שאלה 3 – Depth First Search-G (6 נק'):

1. יבש (1 נק'): עבור בעיית כדורי הדרכון עם לוח $N \times N$, האם האלגוריתם שלם? האם הוא קביל?
2. יבש (1 נק'): האם אלגוריתם DFS (על עץ), עבור בעיית כדורי הדרכון על לוח $N \times N$, היה מוצא פתרון כלשהו? אם כן, מה המסלול שיתקבל? אם לא, כיצד האלגוריתם היה פועל?
3. יבש (2 נק'): נתון לוח בגודל $N \times N$, ללא חורים, המכיל $N^2 - 2$ משבצות רגילות (F,T,A,L) מצב התחלתי בפינה השמאלית עליונה ומצב מטרה בפינה הימנית תחתונה (תניחו כי שני כדורי הדרכון הם בפינה ימנית תחתונה). כמה צמתים יפותחו וייווצרו במהלך חיפוש DFS-G? הסבירו?
4. יבש (2 נק'): נתון לוח בגודל $N \times N$, ללא חורים, המכיל $N^2 - 2$ משבצות רגילות (F,T,A,L) מצב התחלתי בפינה השמאלית עליונה ומצב מטרה בפינה הימנית תחתונה (תניחו כי שני כדורי הדרכון הם בפינה ימנית תחתונה). כמה צמתים יפותחו וייווצרו במהלך חיפוש DFS-G backtracking? הסבירו?

שאלה 4 – ID-DFS (6 נק'):

1.

a. (1 נק') האם האלגוריתם שלם? אם כן, הוכיחו. אם לא, ספקו דוגמה נגדית.

b. (1 נק') נניח כי עלות כל פעולה היא 1, האם האלגוריתם קביל? אם כן, הוכיחו. אם לא, הסבירו.

2. הניחו כי יש לנו ידע מקדים על חסם עליון למרחק למצב מטרה, נסמן D. בת (Beth) הציעה את האלגוריתם חיפוש הבא:

```
function ReverseDFS (problem, D):  
    L ← D  
    result ← failure  
    While Not Interrupted:  
        new_result ← DFS-L (problem, L)  
        if new_result = failure:  
            break  
        L ← L - 1  
        result ← new_result  
  
    return result
```

3. בשאלות הבאות הניחו כי יש מספיק זמן לסיים האיטרציה הראשונה.

a. (1 נק') ספקו דוגמה בה ReverseDFS עדיף על ID-DFS ודוגמה בה ID-DFS עדיף על ReverseDFS. הדוגמאות יכולות להיות כלליות ולא בהכרח מסביבת התרגיל.

b. (2 נק') הציעו כיצד ניצן לייעל את האלגוריתם. רמז: האם אתם יכולים לחשוב על צעד עדכון עדיף ל L? נשנה את האלגוריתם בצורה הבאה :

שאלה 6 – UCS (4 נק'):

השאלות בחלק זה מתבססות על הלוח "8x8" שמופיע במחברת אלא אם נכתב אחרת.

1. יבש (1 נק'): עבור אילו בעיות חיפוש אלגוריתם UCS ואלגוריתם BFS יפעלו באותו האופן? הסבירו.

2. יבש (1 נק'): האם בבעיית החיפוש שלנו, עבור לוח NxN, האלגוריתם הוא שלם? האם הוא קביל?

3. יבש (2 נק'): שאדי טעה במימוש של אלגוריתם UCS ובטעות בדק בעת יצירת הצומת האם היא צומת מטרה במקום בפיתוח שלה. הביאו דוגמה לגרף חיפוש שעבורו שאדי יחזיר בכל זאת את המסלול הקל ביותר, ודוגמה לגרף חיפוש שעבורו שאדי לא יחזיר את המסלול הקל ביותר. עבור כל דוגמה הסבירו מה המסלול והעלות ש-UCS השגוי החזיר, ומה המסלול והעלות שהאלגוריתם הנכון היה מחזיר. נדגיש שגרף החיפוש לא בהכרח צריך לייצג את בעיית כדור הדרכון. אתם יכולים לתת דוגמה לגרף שמייצג בעיית חיפוש אחרת. הגרף צריך להכיל קשתות מכוונות ואת העלות של כל קשת.

שאלה 7 – יוריסטיקות (8 נק'):

יהי מרחב חיפוש (S, O, I, G) , נסתכל על בעיית הניווט לכדור דרקון יחיד. המטרה היא למצוא מסלול זול ביותר מהמוצא I ליעד יחיד G . פונק' העלות מוגדרת כאורך הכביש המחבר בין שתי נקודות, **יהיה חסם תחתון $\delta > 0$ על אורך הכבישים**. ניתן להניח כי העולם שטוח. מלבד זאת, לא ניתן להניח דבר נוסף על מרחב החיפוש.

הגדרה: יוריסטיקה h היא ϵ -קבילה אם קיים $\epsilon \geq 1$ כך שלכל מצב $s \in S$ מתקיים $h(s) \leq \epsilon \times h^*(s)$. נזכיר כי $h^*(s)$ הינה פונקציית המחיר המסלול האופטימלי מ- s לצומת היעד.

עבור כל אחת מהיוריסטיקות הבאות קבעו האם קיים $\epsilon \geq 1$ כך שהיוריסטיקה תהיה ϵ -קבילה. אם כן מצאו את ה- ϵ ההדוק ביותר המקיים את זאת. נמקו היטב.

$$1. \text{ יבש (1 נק') : מרחק מנהטן : } h_{MD}(p) = |P - G|_1 = |G_x - P_x| + |G_y - P_y|$$

$$2. \text{ יבש (1 נק') : } h(p) = |P - G| = \min \{G_x - P_x, G_y - P_y\}$$

$$3. \text{ יבש (1 נק') : } h(p) = |P - G|_3 = \sqrt[3]{|G_x - P_x|^3 + |G_y - P_y|^3} : L^3$$

$$4. \text{ יבש (1 נק') : נתונות יוריסטיקות } h_1, h_2 \text{ שהן } \epsilon_1, \epsilon_2 \text{ קבילות בהתאמה וכי } \epsilon_1, \epsilon_2 \text{ הם האפסילונים ההדוקים ביותר. הראו כי } h_3 = h_1 + h_2 \text{ היא } \epsilon_3 \text{-קבילה, מצאו את } \epsilon_3 \text{ ההדוק ביותר והוכיחו.}$$

נגדיר יוריסטיקה חדשה :

$$\square \quad D \text{ היא קבוצת כדורי הדרקון, } D = \{d1, d2\}.$$

$$h_{MSAP}(s) = \min\{h_{Manhattan}(s, g) | g \in G \cup D\}$$

הערה: בנוסחת המרחק מתייחסים למיקום של צומת.

שימו לב שבמקרה זה אנחנו לוקחים את המינימום על פני כל צמתי היעד.

$$5. \text{ יבש (1 נק') : האם היוריסטיקה } h_{MSAP} \text{ קבילה על כל לוח? אם כן הסבר, אם לא הבא קודמה נגדית.}$$

$$6. \text{ יבש (1 נק') : האם היוריסטיקה } h_{MSAP} \text{ עקבית על כל לוח? אם כן הסבר, אם לא הבא דוגמה נגדית. (לחשוב אם היא עקבית ולתקן בהתאם)}$$

נגדיר יוריסטיקה חדשה :

$$\square \quad D \text{ היא קבוצת כדורי הדרקון, } D = \{d1, d2\}.$$

$$h_{new}(s) = \max\{h_{Manhattan}(s, g) | g \in G \cup D\}$$

$$7. \text{ יבש (1 נק') : האם היוריסטיקה } h_{new} \text{ קבילה על כל לוח? אם כן הסבר, אם לא הבא קודמה נגדית.}$$

$$8. \text{ יבש (1 נק') : האם היוריסטיקה } h_{new} \text{ עקבית על כל לוח? אם כן הסבר, אם לא הבא דוגמה נגדית.}$$

שאלה 8 – Greedy Best First Search (3 נק'):

השאלות בחלק זה מתבססות על הלוח "8x8" שמופיע במחברת אלא אם נכתב אחרת.

1. יבש (1 נק'): האם האלגוריתם שלם? האם הוא קביל?
2. יבש (2 נק'): תנו יתרון וחיסרון של אלגוריתם Greedy Best first Search לעומת Beam Search.

שאלה 9 – W-A* (2 נק'):

השאלות בחלק זה מתבססות על הלוח "8x8" שמופיע במחברת.

1. **רטוב:** ממשו את החלקים החסרים באלג' W-A* בקובץ ע"פ ההנחיות המופיעות שם. עליכם להשתמש ביוריסטיקה h_{MSAP} .
2. (יבש 2 נק') בהינתן $w_1 < w_2 \leq 1$, נסמן את המסלולים המחוזרים על ידי W-A* תחת הפורמולציה $f = g + w \cdot h$ ב p_1, p_2 עבור w_1, w_2 בהתאמה. אזי $cost(p_1) < cost(p_2)$ עבור:
a. יוריסטיקה קבילה h . אם כן הסבירו. אם לא, ספקו דוגמה נגדית.
b. יוריסטיקה כללית (לא בהכרח קבילה) h . אם כן הסבירו. אם לא, ספקו דוגמה נגדית.

שאלה 10 – IDA* (2 נק'):

1. יבש (1 נק'): ספקו יתרון וחיסרון של IDA* ביחס ל-A*. באילו מקרים הייתם מעדיפים להשתמש בכל אחד מהם?
2. יבש (1 נק'): ספק המחשה שלב אחר שלב של אלגוריתם IDA* על הלוח (4x4) שמופיע במחברת, המראה כיצד החיפוש מתקדם באמצעות העמקה איטרטיבית?

שאלה 11 – A* epsilon (6 נק'):

1. **רטוב:** ממשו את החלקים החסרים באלג' W-A* בקובץ ע"פ ההנחיות המופיעות שם. עליכם להשתמש ביוריסטיקה h_{MSAP} .
2. יבש (2 נק'): תנו יתרון וחיסרון של A*-epsilon לעומת A*.
3. יבש (3 נק'): תנו הצעה ליוריסטיקה כדי לבחור את הצומת הבאה לפיתוח מתוך FOCAL. תארו את היוריסטיקה והציגו השוואה בין השימוש ביוריסטיקה זו לעומת השימוש ב- $g(v)$, מבחינת מספר פיתוחים, מסלול שנבחר ועלות המסלול שנבחר.
4. יבש (1 נק'): אם נגדיר שאפסילון שווה לאינסוף איך תהיה ההתנהגות של האלגוריתם עם סביבת כדורי הדרכון.

שאלה 12 – Benchmarking (2 נק')::

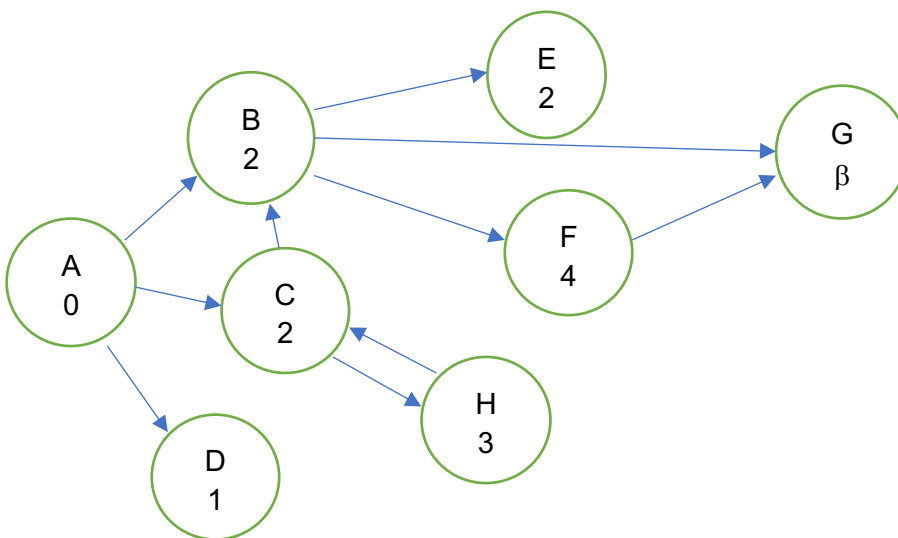
בשאלה זאת נשווה בין אלגוריתמי חיפוש שונים על בעיות שונות. הריצו את החלק הרלוונטי במחברת ותיראו שנוצר קובץ csv. (ניתן לפתוח עם Excel).

1. **רטוב:** הריצו את החלק הרלוונטי במחברת ותיראו שנוצר קובץ csv. (ניתן לפתוח עם Excel).

2. יבש (2 נק'): הסבירו את התוצאות. האם הן תואמות לציפיות שלכם? האם התוצאות היו משתנות עם יוריסטיקה יותר מידועת? נתחו והסבירו את התוצאות במונחים של מספר פיתוחים, מסלול מוחזר ומחיר הפתרון. שימו לב שבסעיף זה אין תשובה נכונה או לא נכונה אבל נדרש ממכם לספק הסבר מפורט ומבוסס.

שאלה 13 – Local Search (5 נק')::

בהינתן מרחב המצבים הבא, כאשר a הינו המצב ההתחלתי, $U: S \rightarrow \mathbb{R}^+$ הינה פונקציית ערך והערך עבור כל מצב מצוין בצומת. המטרה שלנו היא למצוא מצב שממקסם את ערך U .



נשתמש באלגוריתם Stochastic Hill Climbing.

כמו כן ידוע כי $\beta > 3$.

1. יבש (1 נק'): מה ההסתברויות למעבר מהצב ההתחלתי לכל אחד מהמצבים b, c, d . רשמו את

$$p(d|a), p(b|a), p(c|a)$$

2. יבש (1 נק'): מה הוא מספר הצעדים המקסימלי שהאלגוריתם יכול לבצע? צעד מוגדר כמעבר בין מצבים.

3. יבש (1 נק'): בהיתן שבצעד הראשון האלגוריתם עבר למצב c . האם האלגוריתם יתכנס למקסימום הגלובלי?

4. יבש (1 נק'): מה ההסתברות שהאלגוריתם יתכנס לפתרון לא אופטימלי (שאינו מקסימום גלובלי)? יש שני מקרים:

5. יבש (1 נק'): עבור אילו ערכים של β ההסתברות להגיע מהמצב ההתחלתי למקסימום הגלובלי תוך בדיוק 3 צעדים גדול מ $\frac{1}{5}$?

הוראות הגשה:

עליכם להגיש קובץ יחד בשם `AI1_<id1>_<id2>.zip` (בלי הסוגריים המשולשים) המכיל:

1. קובץ בשם `AI1_<id1>_<id2>.pdf` שמכיל את התשובות לחלק היבש.

2. קובץ בשם `Algorithms.py` המכיל את המימוש לאלגוריתמי החיפוש.