פתרון תרגיל מספר 1־ AI

שם: מיכאל גרינבאום, **ת.ז:** 211747639 **שם:** הראל רותם, **ת.ז:** 318661741

26 באפריל 2020

1. **פתרון:**

(א) **צ"ל:** UCS הוא מקרה פרטי של A^*

הוכחה:

נשים לב כי UCS פותח נודים לפי priority queue ולפי פונקצית $g : V \to \mathbb{R}$ המוגדרת להיות מרחק המסלול מההתחלה ל־ v לכל $v \in V$.

נשים לב כי A^* פותח נודים לפי priority queue ולפי פונקצית $f : V \to \mathbb{R}$ המוגדרת $f(v) = g(v) + h(v)$ לכל $v \in V$ ליוריסטיקה h ו־ g המוגדרת מקודם.

לכן אם נבחר את $h(v) = 0$ לכל $v \in V$ (זאת יוריסטיקת האפס),

נקבל כי לכל $v \in V$ מתקיים

$$f(v) = g(v) + h(v) = g(v) + 0 = g(v)$$

כלומר במקרה ספציפי זה A^* פותח נודים לפי priority queue ולפי פונקצית $f : V \to \mathbb{R}$ המוגדרת $f(v) = g(v)$ לכל $v \in V$,

כלומר A^* פותח נודים לפי פונקציה g בדיוק כמו UCS,

ולכן במקרה זה נקבל ש־ A^* יתנהג בדיוק כמו UCS, ולכן UCS הוא מקרה פרטי של A^*

מ.ש.ל.א.☺

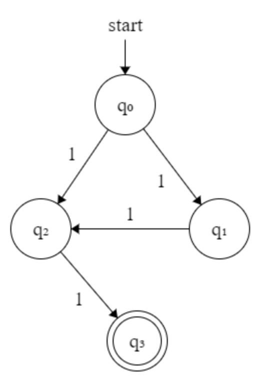
(ב) **צ"ל:** $ID - DFS$ לא בהכרח מוצא פתרון אופטימלי על גרף שכל המשקולות זהים

הוכחה:

נריץ את $ID - DFS$ כמו שראינו בהרצאה, מתחיל מגובה 0 ומתקדם בגובה 1 כל פעם,

נניח ש־ DFS פותח קודם את q_1 ולאחר מכן את q_2 (סדר החיפוש)

נסתכל על הדוגמא הבאה: (אפשר לפרש אוטומט כבעית חיפוש ☺)



נשים לב שבריצה הראשונה נמצא רק את q_0 ונסיים,

בריצה השנייה נפתח את q_0, q_1 ולאחר מכן את q_2 ונסיים,

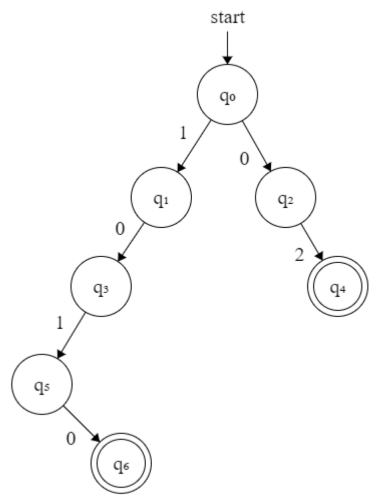
בריצה השלישית נפתח את q_0, q_1, q_2, לאחר מכן מכיוון שכבר פתחנו את q_2, לא נפתח אותו שוב ולכן לא נסתכל על המסלול q_0, q_2, q_3.

בריצה הרביעית נפתח את q_0, q_1, q_2, q_3 ונמצא פתרון באורך 3 כשקיים פתרון אופטימלי באורך 2, כלומר $ID-DFS$ לא בהכרח מוצא פתרון אופטימלי בגרף כשכל הצלעות בעלות אותו משקל

מ.ש.ל.ב.☺

2. **פתרון:**

(א) **צ"ל:** $h_1(n) \le h_2(n)$ וגם לא בהכרח פותחים אותם נודים
הוכחה:
נסתכל על הגרף הבא:



נגדיר $h_1(n) = 0$, זאת יוריסטיקה קונסיסטנטית וגם נשים לב שרצף הפתיחה שנקבל הוא q_0, q_2, q_1, q_3, q_4 ולא נפתח את q_5

נגדיר $h_2(n) = \begin{cases} 2 & n = q_2 \\ 0 & n \neq q_2 \end{cases}$, עבור כל נוד q_2 זה קונסיסטנטי, ועבור q_2 מתקיים $2 = h_2(q_2) \le h_2(q_4) + 2 = 2$ ולכן גם היא קונסיסטנטית.
וגם נשים לב כי $h_1(n) \le h_2(n)$ לכל $n \in V$,
עתה נשים לב שנתחיל את הריצה מ־ q_0, והתור יהיה $q_1 : 1 + 0 = 1, q_2 : 0 + 2 = 2$
עתה נפתח את q_1 והתור יהיה $q_2 : 2, q_3 : 1 + 0 = 1$,
עתה נפתח את q_3 והתור יהיה $q_2 : 2, q_5 : 2 + 0 = 2$,
ואז נפתח את q_2 והתור יהיה $q_5 : 2, q_4 : 2 + 0 = 2$,
ואז נפתח את q_5 והתור יהיה $q_4 : 2, q_6 : 2 + 0 = 2$,
ואז נפתח את q_4 ונסיים ואז חקרנו את $q_0, q_1, q_3, q_2, q_5, q_4$,
נשים לב ש־ h_2 פתחה את נוד q_5 כש־ h_1 לא פתחה אותו,
ולכן הטענה שכל נוד שנפתח על ידי h_2 ייפתח על ידי h_1 לא נכונה

מ.ש.ל.א.☺

(ב) **פתרון:**

i. **צ"ל:** מספר החלקים שלא במקומם אדמיסבילי
הוכחה:
נשים לב שלכל חלק שלא במקומו, נצטרך לבצע לפחות החלפה 1 עם הריבוע הריק כדי להביא אותו למקומו, לכן אם נסמן את מספר המהלכים שצריך לבצע עבור החלק ה־ i ב־ s_i אזי מתקיים

$$\sum_{i=1}^{8} s_i \ge \sum_{i=1}^{8} \begin{cases} 1 & s_i \neq 0 \\ 0 & s_i = 0 \end{cases} = \#\text{number of pieces not in place}$$

לכן מספר החלקים שלא במקומם הינו אדמיסבילי

מ.ש.ל.ב.1.☺

ii. **צ"ל:** סכום מרחקי מנהטן של החלקים אדמיסבילי

הוכחה:

לכל חלק i, נסמן את המרחק מנהטן שלו ממקומו ב־ d_i עם x_i מרחק בציר ה־ x ו־ y_i מרחק בציר ה־ y למקומו.
כלומר $d_i = x_i + y_i$
נשים לב שלכל חלק שלא במקומו, נצטרך לבצע לפחות x_i החלפות בציר ה־x עם הריבוע הריק ו־ y_i החלפות בציר ה־ y עם הריבוע הריקכדי להביא אותו למקומו,
לכן אם נסמן את מספר המהלכים שצריך לבצע עבור החלק ה־ i ב־ s_i,
נשים לב שלכל $s_i = 0$, מתקיים שהחלק במקומו ולכן $x_i = y_i = 0$ אזי מתקיים

$$\sum_{i=1}^{8} s_i \geq \sum_{i=1}^{8} \begin{cases} x_i + y_i & s_i \neq 0 \\ 0 & s_i = 0 \end{cases} = \sum_{i=1}^{8} \begin{cases} x_i + y_i & s_i \neq 0 \\ x_i + y_i & s_i = 0 \end{cases}$$
$$= \sum_{i=1}^{8} [x_i + y_i] = \sum_{i=1}^{8} d_i = \#\text{sum of manhatan distances}$$

לכן סכום מרחקי מנהטן של החלקים הינו אדמיסבילי

מ.ש.ל.ב.2.☺

iii. **צ"ל:** כמות החלקים שלא בשורה הנכונה + כמות חלקים שלא עמודה הנכונה

הוכחה:

לכל חלק i, נסמן את המרחק מנהטן שלו ממקומו ב־ d_i עם x_i מרחק בציר ה־ x ו־ y_i מרחק בציר ה־ y למקומו.
כלומר $d_i = x_i + y_i$
נשים לב שלכל חלק i שלא בעמודה הנכונה מתקיים $x_i > 0$, וכל חלק שלא בשורה הנכונה מתקיים $y_i > 0$, נסמן ב־ m_i אינדיקטור של האם החלק ה־ i בעמודה הלא נכונה וב ־ n_i אינדיקטור של האם החלק ה־ i בשורה הלא נכונה.
נשים לב שאם $m_i = 0$ אזי $x_i = 0$ ובאופן דומה $n_i = 0$ אזי $y_i = 0$
לכן אם נסמן את מספר המהלכים שצריך לבצע עבור החלק ה־ i ב־ s_i,
לכן

$$\sum_{i=1}^{8} m_i + \sum_{i=1}^{8} n_i = \sum_{i=1}^{8} \begin{cases} 1 & m_i \neq 0 \\ 0 & m_i = 0 \end{cases} + \sum_{i=1}^{8} \begin{cases} 1 & n_i \neq 0 \\ 0 & n_i = 0 \end{cases}$$
$$\leq \sum_{i=1}^{8} \begin{cases} x_i & m_i \neq 0 \\ 0 & m_i = 0 \end{cases} + \sum_{i=1}^{8} \begin{cases} y_i & n_i \neq 0 \\ 0 & n_i = 0 \end{cases} = \sum_{i=1}^{8} \begin{cases} x_i & m_i \neq 0 \\ x_i & m_i = 0 \end{cases} + \sum_{i=1}^{8} \begin{cases} y_i & n_i \neq 0 \\ y_i & n_i = 0 \end{cases}$$
$$= \sum_{i=1}^{8} [x_i + y_i] = \sum_{i=1}^{8} d_i = \#\text{sum of manhatan distances}$$

לכן מהסעיף הקודם נקבל כי

$$\#\text{number of tiles out ot row} + \#\text{number of tiles out ot col} = \sum_{i=1}^{8} m_i + \sum_{i=1}^{8} n_i \leq \sum_{i=1}^{8} d_i \leq \sum_{i=1}^{8} s_i$$

כמות החלקים שלא בשורה הנכונה + כמות חלקים שלא עמודה הנכונה הינו אדמיסבילי

מ.ש.ל.ב.3.☺

iv. **צ"ל:** פתרון אופטימלי ל־ $blank - Swap$ הוא אדמיסבילי

הוכחה:

נסמן ב־ $a_1, \ldots, a_n$ רשימה של פעולות אופטימליות לפתרון הלוח.
עתה נשים לב כי פעולות $a_1, \ldots, a_n$ מחליפות $tile$ עם החלק הריק ולכן מהוות פתרון ל־ $blank - Swap$,
ולכן מאופטימליות $blank - Swap$ נקבל כי $blank - Swap \leq n$,
כלומר פתרון אופטימלי ל־ $blank - Swap$ הוא אדמיסבילי

מ.ש.ל.ב.4.☺

3. **פתרון:**

(א) **צ"ל:** פונקציה שמביעה את דרישות השאלה

הוכחה:

יהי $G=\langle V,E\rangle$ גרף לא מכוון, ויהיו $U_1,U_2,U\subseteq V$,
תחילה נגדיר $f:2^V\times 2^V\to\mathbb{R}$ באופן הבא

$$f\left(U_1,U_2\right)=\left|\left\{\{u,v\}=e\in E\mid u\in U_1\wedge v\in U_2\right\}\right|+\left|\left|U_1\right|-\left|U_2\right|\right|$$

הגדרנו את f בתור כמות הצלעות בחתך + הפרש גדלי הקבוצות,
נגדיר פונקצית אופטימיזציה $g:2^V\to\mathbb{R}$ באופן הבא:

$$g\left(U\right)=f\left(U,V\setminus U\right)$$

נרצה למזער את g, כלומר למצוא $U'\subseteq V$ כך ש־

$$g\left(U'\right)=\min_{U\subseteq V}\left\{g\left(U\right)\right\}$$

מ.ש.ל.א.☺

(ב) **צ"ל:** gradient ascent

הוכחה:

תחילה נאתחל קבוצה $U\subseteq V$ ובהתחלה נגדיר $U=V$,
נגדיר

i. <u>מצב</u>: קבוצה $U\subseteq V$

הערה: לא צריך להגדיר זוג קבוצות כי הקבוצה השנייה מוגדרת על ידי $V\setminus U$ לכל מצב (U)

ii. <u>שכנים</u>: יהי (U) מצב, השכנים שלו הינם כל מצב (W) כך שמתקיים 1 מ2 הדברים הבאים:

א'. $\exists v\in U\setminus W$ שעבורו מתקיים $W=U\setminus\{v\}$

ב'. $\exists v\in W\setminus U$ שעבורו מתקיים $U=W\setminus\{v\}$

כלומר הזזה של קודקוד 1 מקבוצה אחת לשנייה.

נגדיר $h\left(U\right)=|E|+|V|-g\left(U\right)$, לפי g שהוגדר בסעיף הקודם (חיפוש מינימום של g שקול לחיפוש מקסימום של h), נריץ gradient ascent על h כאשר:

i. "עלייה" מוגדר כ־ $dh>0$

ii. "מישור" מוגדר כ־ $dh=0$

iii. "ירידה" מוגדר כ־ $dh<0$.

מ.ש.ל.ב.☺

(ג) **צ"ל:** מה עדיף במקרה שלנו, gradient ascent או simulated annealing?

הוכחה:

נשים לב שבבעיה שלנו יש מצבים שיכולים להוריד את הניקוד ולהעלות לאחר מכן הרבה יותר,
וגם נשים לב כי gradient ascent מוצא נקודת מקסימום לוקאלית כש־ simulated annealing יכול למצוא נקודת מקסימום יותר טובה בגלל שהוא לא תמיד בוחר את האופטימלי אלא מוכן לחקור את המרחב בניסיון למצוא פתרון יותר טוב ואולי אף למצוא נקודת מקסימום גלובלית.
מסיבה זאת, simulated annealing רוב הסיכויים יהיה יותר טוב לבעיה הזאת כי הוא רוב הסיכויים יימצא פתרון יותר מ־ gradient ascent לבעיה שלנו מהנימוק לעיל.

מ.ש.ל.ג.☺

(ד) **צ"ל:** המרת הבעיה לבעיה גנטית

הוכחה:

נסמן $V=\{v_1,\dots,v_n\}$,

תחילה נגדיר המרה של הבעיה לגנטית בעזרת וקטור אינדיקטורים,

עבור מצב (U) נגדיר וקטור אינדיקטורים גנטי מתאים $\begin{bmatrix} s_1 \\ \vdots \\ s_n \end{bmatrix}$ כאשר $s_i=1 \Leftrightarrow v_i\in U$ ו־ $s_i=0 \Leftrightarrow v_i\notin U$

הפונקצית $fitness$ שלנו תהיה הפונקצית h שהוגדרה בסעיף ב', שאותה אנחנו רוצים למקסם כאשר מפעילים את h על המצב המתאים לוקטור

מ.ש.ל.ד.☺

(ה) **צ"ל:** הגדרת פונקצית מוטציה וקרוסאובר

הוכחה:

תחילה נגדיר פונקצית מוטציה וקרוסאובר כלליים.

פונקצית מוטציה: פונקציה זאת אחראית על שמירה על גיוון הוקטורים באלגוריתם הגנטי

פונקצית קרוסאובר: פונקציה זאת אחראיות על לשלב מידע של שני הורים על מנת להפיק פתרונות חדשים ולעודד את הוקטורים היותר טובים לשרוד.

עתה נגדיר האלגוריתם הגנטי לסעיף הקודם,

נגדיר פונקצית עזר $q(U)=\frac{g(U)}{|E|+|V|}$

פונקצית מוטציה: יהי v וקטור גנטי ו־ U המצב המתאים לו, לכל אינדקס $i\in[n]$, בהסתברות $q(U)$ נחליף את v_i ל־ $\sim v_i$. ונחזיר את הוקטור שהתקבל.

פונקצית קרוסאובר: יהיו v_1,v_2 וקטורים גנטיים ו־ U_1,U_2 המצבים המתאימים לו, לכל אינדקס $i\in[n]$, בהסתברות $\frac{e^{q(U_1)}}{e^{q(U_1)}+e^{q(U_2)}}$ נגדיר את המקדם ה־ i להיות $v_{1,i}$, אחרת $v_{2,i}$. ונחזיר את הוקטור שהתקבל.

מ.ש.ל.ה.☺

(ו) **צ"ל:** מה עדיף במקרה שלנו, שיפורים איטרטיבים או אלגוריתם גנטי?

הוכחה:

באופן דומה לסעיף ג', שיפורים איטרטיבים ימצאו מקסימום לוקאלי ולא גלובלי,

לעומת זאת, אלגוריתם גנטי מהווה חלק עם הסתברות בעקבות החלק ה$mutation$ ולכן יכול לצאת מהמקסימום הלוקאלי ולמצוא ערך יותר ואולי אף מקסימום גלובאלי, על אף שיש אפשרות שהוא דווקא יימצא מקסימום לוקאלי פחות טוב מהמקורי.

מסיבה זאת, אלגוריתם גנטי עדיף על שיפורים איטרטיבים לבעיה שלנו שכוללת מקסימומים לוקאליים.

מ.ש.ל.ו.☺

4. **פתרון:**

(א) **צ"ל:** מה המשתנה הראשון שיבחר ומדוע?

הוכחה:

לפי היוריסטיקה הראשונה, צריך לבחור אחד מהמשתנים עם הכי קצת השמות, מהיות וזאת השמה ראשונה אנחנו נשארים עם הקבוצה $\{A,B,C,D,E,F\}$,

עתה צריך לבחור מהקבוצה הזאת עם היוריסטיקה השנייה את אלה שבעלי הדרגה הכי גבוהה ונשאר עם $\{B,D,F\}$,

עתה מההגדרה שאם נשארו יותר מאפשרות אחת, לבחור לפי סדר אלף בית, נקבל כי המשתנה הראשון שנחפש לו השמה הוא B.

מ.ש.ל.א.☺

(ב) **צ"ל:** מה ההשמה למשתנה הראשון ומדוע?

הוכחה:

עתה ההיוריסטיקה השלישית לא תשפיע כי כל השמה תגביל את שאר המשתנים באותה הכמות, לכן נשאר עם הקבוצת ההשמות $\{Red,Green,Blue\}$.

עתה מההגדרה שאם נשארו יותר מאפשרות אחת לבחור לפי הסדר $red,green,blue$ נקבל כי ההשמה למשתנה הראשון היא Red.

עתה מ־ $Forward-Checking$, לשכנים של B שהם $\{A,C,D,F\}$ יישארו רק האפשרויות $\{green,blue\}$

מ.ש.ל.ב.☺

(ג) **צ"ל:** מה המשתנה השני שיבחר ומדוע?
הוכחה:
לפי היוריסטיקה הראשונה, צריך לבחור אחד מהמשתנים עם הכי קצת השמות, מההשמה הקודמת אנחנו יודעים כי הקבוצה עם הכי מעט אפשרויות היא $\{A, C, D, F\}$,עתה צריך לבחור מהקבוצה הזאת עם היוריסטיקה השנייה את אלה שבעלי הדרגה הכי גבוהה ונשאר עם $\{D, F\}$,
עתה מההגדרה שאם נשארו יותר מאפשרות אחת, לבחור לפי סדר אלף בית, נקבל כי המשתנה השני שנחפש לו השמה הוא D.

מ.ש.ל.ג.☺

(ד) **צ"ל:** מה ההשמה למשתנה השני ומדוע?
הוכחה:
עתה ההיוריסטיקה השלישית לא תשפיע כי כל השמה תגביל את שאר המשתנים באותה הכמות, לכן נשאר עם הקבוצת ההשמות $\{Green, Blue\}$.
עתה מההגדרה שאם נשארו יותר מאפשרות אחת לבחור לפי הסדר $red, green, blue$ נקבל כי ההשמה למשתנה הראשון היא $Green$.
עתה מ־ $Forward - Checking$, לשכנים של D שהם $\{B, C, E, F\}$ נצטרך למחוק את האפשרות ל$Green$ ונקבל את התנאים הבאים:

$$A : \{Green, Blue\}$$
$$C : \{Blue\}$$
$$F : \{Blue\}$$
$$E : \{Red, Blue\}$$

מ.ש.ל.ד.☺

(ה) **בונוס:** נוכל להשלים את הצביעה ולהסיק מיידית כי C, F נצבעים ב־ $Blue$, לאחר מכן מ־ $Forward - Checking$ ל־ F נצטרך למחוק את האפשרות $Blue$ ל־ A, E, ואז נקבל כי A ייצבע כירוק ו־ E באדום ☺