**תרגיל בית 1: שימוש באלגוריתמי חיפוש היוריסטיים לתכנון מסלולי חלוקה אופטימליים**

**מגישים:**

**אריאל ווייס  
תז: 204361315**

**דניאל בן חיון  
תז: 308021054**



## חלק א' – מבוא והנחיות (3 נק׳ יבש + 3 נק' בונוס)

1. יבש (1 נק'):

ראשית נסדר את הדירות – K! , לאחר מכן יש m אפשרויות לכל מעבדה בין דירות + אפשרות ללא מעבדה- ונשארו m אפשרויות למעבדה האחרונה, סה"כ:

1. יבש בונוס (3 נק'):

דומה לסידור הקודם, אך הפעם בין כל שתי דירות נחלק לשני מקרים- מקרה ראשון, לא עברנו במעבדה ולכך יש אפשרות אחת. מקרה שני, עברנו במעבדה שהתחילה רצף: כעת נבחר את מספר המעבדות (i) הנוספות שנעבור בהן ונסדר את המעבדות שנבחרו, ובנוסף נסדר את המעבדה שהתחילה את הרצף ברווחים בין המעבדות(2 אפשרויות, לשים או לא), כלומר:

סה"כ:

1. יבש (2 נק׳):

|  |  |  |  |
| --- | --- | --- | --- |
| Estimated calculation time | *#possiblePaths* |  |  |
| ~18.48 [secs] | ~22.04\*10^6 | 2 | 7 |
| ~3.85 [mins] | ~24.77\*10^7 | 3 | 7 |
| ~2.26 [hours] | ~79.27\*10^8 | 3 | 8 |
| ~19.56 [hours] | ~63.00\*10^9 | 4 | 8 |
| ~3.69 [days] | ~28.54\*10^10 | 3 | 9 |
| ~5.33 [months] | ~11.42\*10^12 | 3 | 10 |
| ~21.05 [years] | ~50.23\*10^13 | 3 | 11 |
| ~1.08 [thousand years] | ~24.11\*10^15 | 3 | 12 |
| ~22.41 [thousand years] | ~46.78\*10^16 | 4 | 12 |
| ~1.55 [million years] | ~30.41\*10^18 | 4 | 13 |

## חלק ג' – הגדרת מרחב החיפוש של בעיית מד״א (6 נק׳ יבש)

### תרגילים

לטובת הסעיפים בחלק זה הנח שלאו דווקא קיים פתרון ישיג במרחב.

1. יבש (1 נק׳):   
   במקרה המקסימלי, אנחנו נמצאים במצב ההתחלתי ונוכל להגיע לכל דירה או לכל מעבדה  
   כלומר   
   במקרה המינימלי, אנחנו במעבדה, ואם אנחנו במצב מטרה, אז הדרגה היא 0.
2. יבש (1 נק׳):   
   לא ייתכנו- כי כדי שיווצר מעגל, נצטרך לבקר באותה דירה פעמיים או באותה מעבדה פעמיים:  
   לא נבקר באותה דירה, כי מהגדרת האופרטור נוכל לבקר בדירה רק אם ואחרי ביקור אחד, di מתווסף לTaken.

ביקור לשם לקיחת מטושים יקרה פעם אחת במעבדה. בביקור חוזר, הגענו עם Transferred מסוים, ובהכרח המספר ישתנה בהגעה השנייה למעבדה ולכן זהו לא אותו מצב אף פעם.

1. יבש (1 נק׳):   
   מספר המצבים הוא אינסופי כיוון שיש אינסוף אפשרויות למספר המטושים (N טבעי). לא כל המצבים ישיגים, לדוגמא המצב בו מספר המטושים גדול יותר ממספר המטושים ההתחלתי + מספר המטושים הקיים בכל במעבדות.
2. יבש (1 נק׳):   
   ייתכנו בורות ישיגים, לדוגמא המצב בו מספר המטושים הכולל נמוך ממספר הדיירים הכולל (כלומר נתקע בשלב מסויים ללא מטושים אך עם דירות שלא ביקרנו בהן)
3. יבש (1 נק׳):

האורך המינימלי של מסלול: נניח שנתחיל עם מספר מטושים המספיק לבקר בכל הדירות, נשאר לבקר במעבדה האחרונה, לכן סה"כ: K+1

האורך המקסימלי: נניח שנתחיל בביקור בכל המעבדות ללקיחת מטושים ולאחר מכן נבקר בכל הדירות כשאחרי כל דירה נבקר במעבדה, לכן סה"כ: M + 2\*K

1. יבש (1 נק׳):

כש-

## חלק ד' – מתחילים לתכנת

### תרגילים

1. רטוב

## חלק ה' – אלגוריתם A\* (2 נק׳ יבש)

### תרגילים

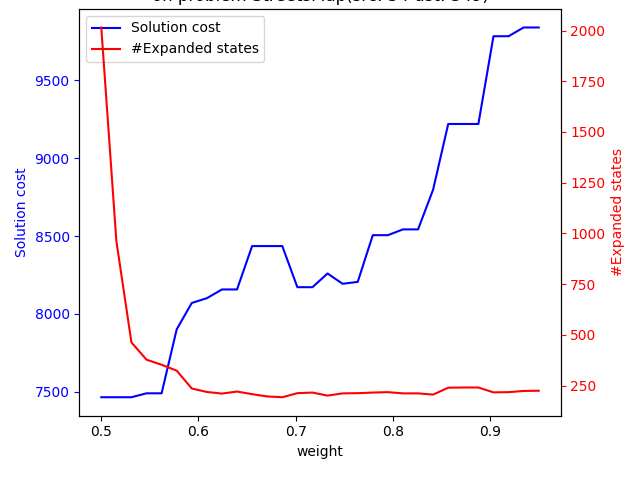
1. רטוב
2. רטוב
3. רטוב
4. יבש (1 נק')

מספר מצבי פיתוח ( nullhuristic) : 17354

מספר מצבי פיתוח ( AirDistHeuristic ) : 2015

החיסכון: 88 אחוז

1. רטוב
2. יבש (1 נק׳):



**הסבר**: הגרף מתאר את מספר הצמתים שפותחו<אדום> ועלות הפתרון <כחול> (של ריצת A\* עם AirDistHeuristic) כתלות בערכי w שונים בתחום [0.5,0.95].

**ערך W**: נרצה ערכי w עבורם נפתח כמה שפחות מצבים ונשלם כמה שפחות על הפתרון, מהגרף ניתן לראות כי נקודה זו מתקבלת באזור שבין 0.58~ לבין 0.52~, כלומר זה האזור העדיף בשבילנו.  
נשים לב שהכלל אכן בא לידי ביטוי בגרף באופן כללי אך יש מקטעים בהם הוא אינו נכון, למשל בקטע 0.68~ עד 0.71~ שם יש ירידה בעלות הפתרון אך עלייה בערכי w.

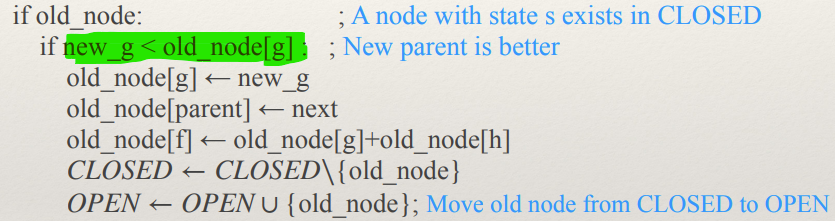
## חלק ו׳ – מימוש בעיית מד״א (15 נק׳ יבש)

1. רטוב
2. רטוב
3. שאלה יבש (2 נק'):

במרחב "העל" יהיו הרבה יותר מצבים אפשריים, מסלולים יתנפחו והcache יהיה מנופח בהתאמה. כיוון שאנחנו בודקים עלויות בcache בפתרון בעיית המרחק, cache גדול יותר אומר יותר זמן לחפש בו, ודבר זה יפגע לנו ביעילות זמן ריצה של הפתרון.

1. שאלה יבש (3 נק׳):   
   (i)   
     
   (ii)

שורה זו לא מספיקה -היא מונעת שינוי של שדות שניגשים אליהם by value אבל לא מונעת שינוי של שדות שניגשים אליהם by reference . אל הסטים ניגשים by reference ולכן הם מוגדרים להיות FrozenSet, וזהו הדבר הנוסף שמבטיח שלא יהיה ניתן לשנות בטעות את האובייקט.  
(iii)

כן יכול להווצר מצב כזה. במקרה בו בעת פיתוח צומת נגלה מצב שכבר פיתחנו אשר ערך הפונקציית g שלו "זולה" יותר, ואז נכניס את הצומת לתור ה OPEN שוב למרות שכבר פיתחנו את המצב הזה , רק שהפעם נעדכן את ערכי g ו f לערכים החדשים שגילינו . זה קורה פה:  


(iv)

הצורך בהגדרת MDAState כ- frozen נובע מזה שבעת פיתוח צמתים עוקבים, כאשר סורקים צומת מסוים לא נרצה לאפשר למצב העוקב שמפותח לשנות את הערכים של מצב האבא שלו, מכוון שכל צומת עוקב הוא חלק ממסלול שונה בסריקה. שינויים כאלה עלולים לגרום להגדרת מצבים שגויים ולשגיאות במהלך החיפוש. לדוגמא במתודה cost\_with\_state\_expand של המחלקה MDAProblem :

המתודה מחזירה **צמתים** **עקבים** עבור צומת מסוים עם שימוש ב yield. כאשר מעצבעים פיתוח של עוקב הנוצר מביקור במעבדה L כלשהי ( הפעלת אופרטור ביקור) ניתן לשנות ישירות את הרשימה visited\_labs באופן הבא : להוסיף את המעבדה L לרשימה visited\_lbas ואז לבצע השמה של המעבדה L ל visited\_labs במצב החדש שפיתחנו.

state\_to\_expand.visited\_labs.append(L)

new\_state.visited\_lab=L

כתוצאה מכך גם האבא וגם המצב העוקב חולקים מצביע לאותה רשימה visited\_labs בעוד שהרשימה לא נכונה עבור האבא. בפעם הבאה שנפתח עוקב לאותו מצב השגיאה תגרום לפיתוח מצבים לא נכונים

עתה, כדי להריץ את A\* על הבעיה, יש ראשית להגדיר (ולממש) היוריסטיקות עבור הבעיה.

1. רטוב
2. רטוב
3. יבש (4 נק׳): הוכח/הפרך:   
   היוריסטיקה קבילה. הוכחה:

ע"פ הגדרה h יוריסטיקה קבילה אם מתקיים:

נסמן h'= MDAMaxAirDistHeuristic . h' נותנת לכל צומת ערך שמציין את המרחק האווירי המקסימלי בין 2 נקודות בהן המסלול צריך לעבור (שעדיין לא עבר בהן). כלומר הערך מציין את המרחק המקסימלי בין 2 דירות שעדיין לא ביקרו בהן. נבחן את המקרה בו נשארה דירה אחת בלבד שעדיין לא ביקרו בה. מתקיים: h'(s)≤〖h'〗^\* (s) . מאחר והמרחק האווירי הוא יוריסטיקה קבילה ( הוכח בהרצאה). לפי א"ש המשולש כל הוספה של דירה תאריך את המסלול ולכן מתקיים לכל s :

h'(s)≤h^('\*) (s)

ולכן קיבלנו כי h' קבילה ע"פ הגדרה.

1. רטוב
2. רטוב
3. יבש (4 נק׳): הוכח/הפרך:

לא קבילה. דוגמא נגדית: עבור k=3 ומיקומי הדירות הם : והמיקום ההתחלתי של האמבולנס הוא (0,0) (יש מספיק מטושים התחלתיים ובכל דירה דייר אחד) נקבל:

H(s)=sum((0,0)=>(1,0)=>(3,0)=>(-2,0))=1+2+5=8

אבל קיים מסלול קצר יותר:

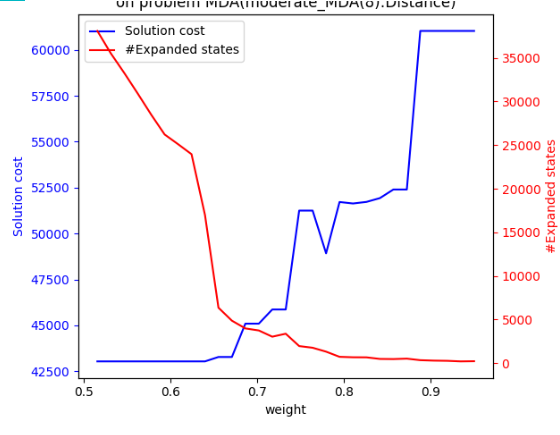
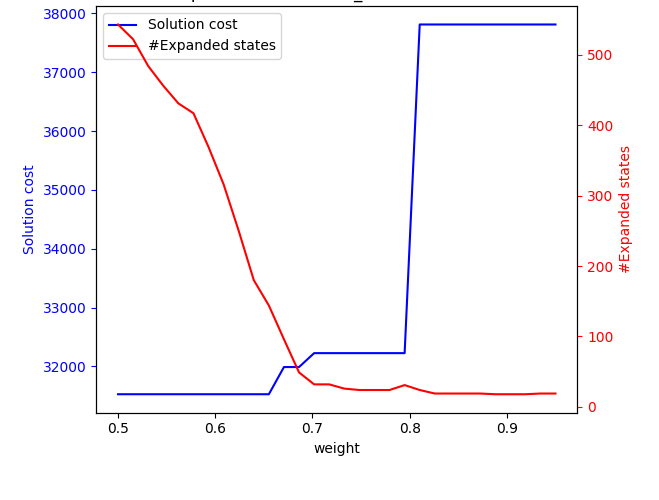
H(s\*)=sum((0,0)=>(-2,0)=>(1,0)=>(3,0))=2+3+2=7

כלומר מצאנו מסלול s\* עבורו *. ולכן זו לא יוריסטיקה קבילה.*

1. רטוב
2. רטוב
3. יבש (4 נק׳): הוכח/הפרך:   
   יוריסטיקה קבילה:

ע"פ הגדרת הפונקציה הערך של מצב s הוא סכום קשתות ה MST של גרף G , כאשר הצמתים בו מוגדרים להיות דירות שעדיין לא בוקרו בהם ממצב S. משקל כל קשת הוא המרחק האווירי בין שני צמתים. על מנת להגיע ממצב s למצב מטרה צריך לעבור בכל הצמתים בגרף G. , כלומר ניתן להסתכל על הבעיה של מסלול בין דירות כמסלול של עץ פורש בגרף G. נתון שהערך היוריסטי של s הוא ערך העץ הפורש המינימאלי של גרף זה , בהכרח כל מסלול אחר בין הדירות יהיה **לפחות**  באותו המחיר. מחיר של כל קשת מוגדר לפי המרחק האווירי בין שתי דירות שהיא מחברת, זאת אומרת ש**מקבילות יוריסטיקת המרחק האווירי**  ומכך שמדובר ב **עץ פורש** נקבל

ע"פ הגדרה קיבלנו ש MDAMSTAirDistHeuristic קבילה.

1. רטוב + יבש (1 נק׳):  
   

הסבר:

עם היוריסטיקה עבור ערכי שוניםASTARהגרפים כמו קודם מתארים מס' צמתים שפותחו ועלות הפתרון בריצת האלג'.

ערכים כדאיים: אנחנו רוצים לשמור על פתרון קביל, לכן בשתי הבעיות נבחר   
 . אם נוכל לוותר על הקבילות, נפתח משמעותית פחות מצבים ונבחר בw=~0.66  
w=~0.7 for moderate  
w=~0.67 for small

## חלק ז' – מימוש והשוואת פונק׳ עלות שונות (21 נק׳ יבש)

1. יבש (2 נק׳): סמן בכל אחד מהתאים כן/לא. האם כל אחת מההיוריסטיקות הנקובות הינה קבילה ביחס לפונק׳ המחיר הנקובה? (אין צורך בנימוק).

|  |  |  |  |
| --- | --- | --- | --- |
|  | MDAMaxAirDistHeuristic | MDASumAirDistHeuristic | MDAMSTAirDistHeuristic |
|  | לא | לא | לא |
|  | כן | לא | כן |

1. רטוב + יבש (0.5 נק' יבש):

מס' נשווה את שתי סוגי הבעיות שבדקנו, mda small ו mda moderate, עם monetary cost לעומת distance cost:  
MDA(small\_MDA(5):Distance): total\_cost: MDACost(dist= 31528.659m, money= 49.717NIS, tests-travel= 52112.429m)  
MDA(moderate\_MDA(8):Distance): total\_cost: MDACost(dist= 43034.794m, money= 95.847NIS, tests-travel= 176505.013m)  
MDA(small\_MDA(5):Monetary): total\_cost: MDACost(dist= 31923.809m, money= 42.050NIS, tests-travel= 53317.118m)  
MDA(moderate\_MDA(8):Monetary): total\_cost: MDACost(dist= 54951.037m, money= 77.201NIS, tests-travel= 172922.318m)  
**SMALL**: כשנלקח distance cost, קיבלנו מחיר: NIS49.717  
לעומת בmonetray cost, שקיבלנו: 42  
**MODERATE**: במקרה עם distance cost קיבלנו מחיר: NIS95.847  
עם monetary: 77

1. רטוב
2. יבש (2 נק׳):

היוריסטיקה קבילה עבור testtravel cost:  
ניקח מצב S כלשהו. נניח h\*(s) = 0. אזי אין יותר דירות לבקר ולכן גם h(s) = 0.  
אחרת, h\*(s) > 0, נותרו דירות לביקור. לכל דירה k נסמן מרחק למעבדה הקרובה ביותר k\_closet\_lab ומרחק שהאמבולנס נסע בפועל בפתרון האופטימלי למעבדה k\_real\_lab.  
כיוון שלכל דירה, k\_closet\_lab < k\_real\_lab, אז גם h(s) = sum(k\_closet\_lab) < sum(k\_real\_lab) = h\*(s)  
והדבר נכון לכל S, לכן היוריסטיקה אופטימית תמיד, כלומר קבילה.

1. רטוב + יבש (0.5 נק' יבש):

הפלט שיצא:  
MDA(moderate\_MDA(8):TestsTravelDistance):  
 total\_cost: MDACost(dist= 93226.428m, money= 59.969NIS, tests-travel= 131265.153m)  
ניתן לראות שהמרחק שהבדיקות עברו בפתרון זה: tests-travel= 131265.153m   
קטן מהמרחק שהן עברו בפתרונות הקודמים: tests-travel= 176505.013m

### שילוב בין 2 מדדים

1. יבש (3 נק׳): הוכח/הפרך:

הוכחה:

נניח שקיים פתרון במרחב המקורי, ונסמן את מסלול הפתרון האופטימלי - q= .

מסלול זה הוא מצב מקבל במרחב המסלולים, נסמן את המסלול המשוייך אליו   
ב- .p = כיוון שעלות המצב הכי יקר היא מתקבל שלכל צומת-

לכן עבור העלות שקיבלנו כל צומת יקבל ערך קטן מאינסוף ויוכלו להפתח אם נרוץ מספיק זמן. בסופו של דבר גם יפתח מצב מטרה, ולכן אם קיים פתרון במרחב המקורי האלגוריתם בהכרח יחזיר פתרון.

1. יבש (3 נק׳): הוכח/הפרך:   
   הוכחה:

מהגדרת , יש מסלולים (לפחות אחד, נסמנו p) שכל צמתיו מקיימים את התנאי בגרף   
ש לכל אפסילון. מהגדרת פונ' המחיר, המחיר של מסלולים אלו נמוך משל כאלה שאינם מקיימים תנאי זה, ובפרט לא ייתכן שנפתח מצב מטרה שאינו מקיים תנאי זה (בגלל מינימליות UCS). יתר על כן, מצב מטרה שנפתח הוא המצב המינימלי מכל אלו המקיימים את התנאי (מאותה סיבה) ולכן גם בעל ערך המינימלי.  
סה"כ מצב מטרה שנפתח עם האלגוריתם יהיה המינימלי תחת שתי פונ' העלות ולכן המסלול שיתקבל אופטימלי ע''פ הקריטריון המשולב

1. רטוב + יבש (0.5 נק׳):

הפלט שהתקבל:  
MDA(moderate\_MDA(8):TestsTravelDistance):

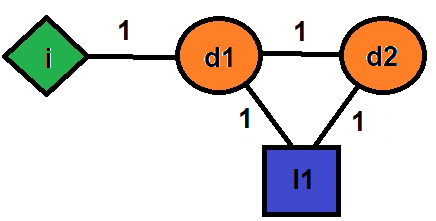
total\_cost: MDACost(dist= 65686.522m, money= 99.486NIS, tests-travel= 132209.981m)

לעומת סעיפים קודמים (מסודר בסדר עולה של dist):

|  |  |  |
| --- | --- | --- |
|  | Test-travel | dist |
| A\* , testtravel cost | 131265.153m | 43034.794m |
| A2 | 132209.981m | 65686.522m |
| A\*, dist cost | 176505.013m | 93226.428m |

ניתן לראות ש 2A אכן מאזן בין המדדים(נמצא בין שניהם), אך מוותר לשם כך על הdist האופטימלי.

חישוב הערך הנל: (65686.522/43034.794) -1 = 0.53 ~ 0.6 , ערך ה-epsilon אכן נשמר!



1. יבש (4 נק׳): הוכח/הפרך:

הפרכה:

דוגמא נגדית-

ניקח גרף כזה : i=(0), d1=(1), d2=(2), l1=(3), עם קשתות המחברות וממושקלות כמו בציור.   
נניח שיש מספיק מטושים כדי לא לדאוג, בכל דירה דייר יחיד. נניח שהאמבולנס בנק התחלה.   
אחרי ריצת ASTAR, נקבל שהמרחק האופטימלי הוא 3 (i->d1->d2->l1) =   
ניקח . (צמתים עם מרחק גדול מ 1.1\*3 לא ניקח)  
נראה ריצה של הUCS על הגרף: [ייצוג צומת- v:(g\_testtravel,g\_dist)]

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
| 5 | 4 | 3 | 2 | 1 | Iter. |
| ---- | D2 | L1 | D1 | i | Next |
| {l1:(4,1),d1:(4,1)} | {d1:(0,3),d2:(0,3)} | {d2:(2,1),l1:(2,1)} | {d1:(1,0)} | {i} | Open |
| {I,d1,l1} | {I,d1,l1} | {I,d1} | {i} | {} | Close |
|  |  |  |  |  |  |

נשים לב ש'נתקענו' באיטרציה החמישית כיוון שאין צמתים העונים על תנאי הdist. לכן אין פתרון

1. יבש (4 נק׳):

הוכחה:

לפי אופן פעולת האלגוריתם, הוא מחפש פתרון לפי פונ' העלות , לכן אם מוחזר פתרון הוא יהיה אופטימלי ביחס לפונ' זו(מאופטימליות UCS). בנוסף, הפתרון מקיים שלכל צומת העלות שלו קטנה מ,(אחרת היה נמחק ולא מפותח),ומאופטימליות נקבל כי הפתרון המוחזר אופטימלי ע"פ הקריטריון המשולב.

1. יבש (1.5 נק׳):

פשוט רץ על מרחב הרבה יותר גדול מ(מרחב המסלולים של ), דבר המייקר את זמן הריצה (למשל במעבר על מס' הצמתים לפיתוח): במקרה גרוע מרחב המסלולים הוא בגודל 2 בחזקת המרחב הרגיל (קבוצת החזקה).

## חלק ח' – מימוש האלג׳ A\* והרצתו (1 נק׳ יבש)

1. רטוב
2. רטוב
3. יבש (1 נק׳):

MDA(small\_MDA(5):Distance) A\* (h=MDA-MST-AirDist, w=0.500) time: 0.60 #dev: 543 |space|: 877 total\_g\_cost: 31528.65909 total\_cost: MDACost(dist= 31528.659m, money= 49.717NIS, tests-travel= 52112.429m) |path|: 8

MDA(small\_MDA(5):Distance) A\*eps (h=MDA-MST-AirDist, w=0.500) time: 38.12 #dev: 492 |space|: 821 total\_g\_cost: 31528.65909 total\_cost: MDACost(dist= 31528.659m, money= 49.717NIS, tests-travel= 52112.429m) |path|: 8

חסכנו בפיתוחים : ניתן לראות כי חסכנו 51 פיתוחים בריצה עם A\*eps. השוני בין האלגוריתמים הוא בבחירת הצומת הבא לפיתוח.

רשימת ה focal מחזיקה צמתים שערך ה f שלהם מינימלי עד כדי כפל ב eps + 1 כלומר ערך ה f שלהם אינו המינימלי, אבל הוא טוב מספיק בשביל לבחון אפשרות שמחיר המסלול שנפתח ממנו יהיה זול. מתוך הרשימה הזאת, בוחרים את הצומת בעלת הערך היוריסטי הנמוך ביותר, כלומר הצומת שמוערך על ידי h כצומת הכי קרוב למטרה.

מכיוון ש f מושפעת מ g ומ- h , גם אם f מקבלת ערך מינימאלי מכל הצמתים בתור ה OPEN, לא מובטח שאכן זה הצומת שיתן את מסלול האופטימלי, למשל יכול להיות מצב שהיוריסטיקה עבור צומת זה הייתה אופטימית ועבור צמת אחר הייתה מחמירה. מכיוון ש**כן**  נותנים משקל להיוריסטיקה אכן מגבילים את f של הצמתים ב OPEN להיות מינימאלית (עד כדי 1+ epx ), וכך גם השתמשנו במידע שהיוריסטיקה נותנת לנו. בנוסף, אפשרנו גמישות כדי לכפר על מקרים שהיא לא מדויקת. העקרון מאחורי זה הוא שנותנים הזדמנות גם לצמתים שלא נראו מבטיחים בהתחלה אבל כן הובילו למסלולים אופטימליים, להיות מפותחים בשלב מוקדם באיטרציות ( מישהו רק היה צריך להאמין בהם)

## חלק ט' – מימוש האלג׳ Anytime A\* והרצתו

1. רטוב
2. רטוב

## חלק י' – שאלה תאורטית (12 נק' יבש)

סעיף (א) – 1 נק' יבש

IDA\* תופס פחות זכרון בזמן הריצה. ליתר דיוק ( O( bd כאשר b הוא branching factor ו d הוא **עומק** העץ. בעוד שבA\* צריכת הזכרון פרופורצינלית למס הצמתים שנוצרו.

סעיף (ב) – 5 נק' יבש

1. (1 נק' יבש)

זמן ריצה

1. (2 נק' יבש)

כי IDA\* לא שומר מבנה נתונים של צמתים שכבר ביקרו בהם , ולכן יכול להיווצר מצב שאלגוריתם מבקר שוב ושוב באותה צומת מבלי לדעת שביקר בו בעבר.

1. (2 נק' יבש)

לא. ראשית מכיוון ש IDDFS לא קביל בעוד ש BFS כן קביל. לעומת A\* וגם IDA\* ששניהם קבילים. בנוסף IDDFS מגדיל את עומק החיפוש כל פעם ב 1 מה שלא מובטח ב IDA\*.

בהמשך השאלה נבחן וריאציה לאלג׳ IDA\* שמטרתה להתמודד עם הבעיה עליה נשאלתם בסעיף ב׳, תוך הקרבה של איכות הפתרון.

סעיף (ג) – 6 נק' יבש

1. (3 נק')

תשובה:

הסבר: כל עוד מוצאים צמתי עם ערך f קטן מ f-limit , ממשיכים לפתח את הצמתים התואמים.אחרת , מגדילים את f-limit ב כל עוד אין פתרון , לכן בסה"כ נקבל כי הערך המקסימלי של f-limit הוא

1. (3 נק')

חסם עליון כנ"ל הוא כיוןון שכל מסלול נבדק בקפיצה של ולכן ערך אפסילון המקסימלי הוא

