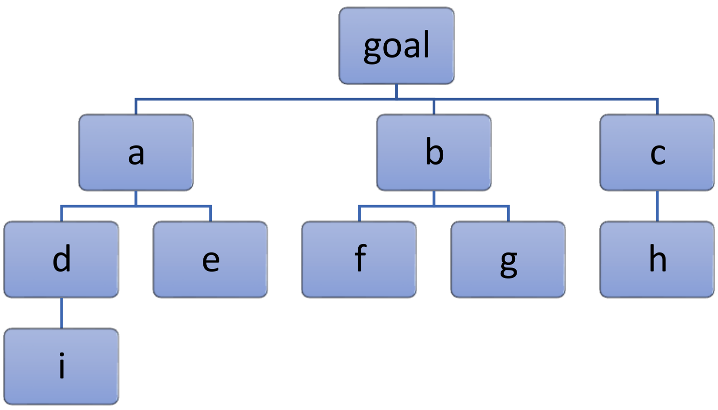
קורס יסודות בינה מלאכותית
מרצה ד"ר עומר לב

אור נחשון אגש
311330500

תרגיל 1

*הערה לגבי הקוד: לא הבנתי אם הגרף מכוון או לא, כי מצד אחד בפורום כתבת שהוא מכוון, ומצד שני כתבת שניתן להניח תקינות קלט ושלא אמור לקרות מקרה שבו x שכן של y אבל y לא שכן של x. ולכן כתבתי את הקוד כך שהגרף מכוון, אבל אם רוצים שהוא יהיה לא מכוון, ניתן להוסיף את שורה 28 שנמצאת בהערה.

1. כאמור כל location מורכב משם (name) ומדינה (2 letter US state code) ועל סמך עובדה זו קבעתי היוריסטיקה שנשענת על השכנות בין המדינות ומרחקן אחת מן השנייה.
מיפיתי את כלל המדינות ועבור כל מדינה בדקתי והגדרתי את המדינות השכנות שלה (מדינה שכנה היא כזו שניתן לעבור אליה בצעד אחד) – ראה נספח 1.
לכל יעד שהוגדר ברשימת ה- goal locations הרצתי אלגוריתם BFS כאשר המדינה שלו היא שורש העץ, השכבה מתחת הן המדינות השכנות, המדינה מתחתיהן הן השכנות של השכנות וכולי. בצורה זו בעצם בניתי עץ המחשב את המרחק של כל מדינה ממדינת היעד ועל סמך מרחק זה נקבע ערך ה-h של כל location (לדוגמא- עבור כל הlocationים במדינה שנמצאה בשכבה השנייה יקבע h=2).
בנוסף, עבור מדינות שאינן נמצאות בעץ (מדינות שלא ניתן להגיע מהן למדינת היעד) קבעתי ערך h להיות 100,000 (מספר גדול מספיק כך שלעולם לא יבחרו).
האם ההיוריסטיקה קבילה? האם היא עקבית?
- ההיוריסטיקה הינה קבילה [admissible] שכן היא תמיד תאמוד ערך שקטן או שווה לערך האמיתי. בגלל שהיא מבוססת על המרחק בפועל בין המדינות הרי שזהו חסם תחתון למספר הצעדים האופטימלי לפתרון, כי במקרה הטוב זה בדיוק יהיה מספר הצעדים לפתרון אך בפועל כנראה שמספר הצעדים יהיה גדול יותר כי בכל מדינה יש הרבה ערים וגם בניהן יש קשרי שכנות. אם הכתובת הנוכחית נמצאת במרחק המינימלי של 3 מדינות מכתובת היעד (כלומר h=3) בהכרח נצטרך לעבור בשלושתם לפחות (או במסלול אחר ארוך יותר) ולכן העלות בפועל תהיה 3 ומעלה.
- ההיוריסטיקה הינה עקבית [consistent], כלומר מתקיים $h(n) \le c(n,n') + h(n')$. ההוכחה:
בשביל הויזואליות- בהינתן כתובת יעד goal, נניח שעץ ה-BFS כשמדינת היעד בקודקוד נראית כך:



ייתכנו מספר אפשרויות:

- n ו-'n נמצאות באותו עומק d.

במקרה זה h(n)=h(n')=d ולכן לא משנה מהו c(n,n') (שהוא בהכרח אי שלילי) יתקיים במקרה זה h(n) ≤ c(n,n') + h(n').

- n נמצאת בשכבה עמוקה יותר מ'n.

לפיכך אם h(n')=d אזי h(n)=d+x כאשר x הינו שלם חיובי. כלומר, d+x ≤ c(n,n') + d.
c(n,n') חייבת להיות לפחות x (או יותר), שכן אם היא קטנה יותר זה אומר שבעת פרישת העץ ב-BFS יכולנו לפתוח את צומת n קודם וזו סתירה.

הסתכלות ויזואלית-

נניח n'=c ו-n=i.במקרה זה לפי ההיוריסטיקה: h(n')=h(c)=1 ו- h(n)=h(i)=3. כמו כן x=2. מכאן c(n,n') יהיה לפחות שתיים: הוא לא יכול להיות 1 כי אז המשמעות היא שניתן לעבור מצומת 'n לצומת n בצעד אחד, אבל אם זה היה המצב אזי בהכרח בעת הרצת ה-BFS n היה בן ישיר של 'n, ולכן יש פה סתירה.

לכן גם במקרה זה יתקיים h(n) ≤ c(n,n') + h(n').

- 'n נמצאת בשכבה עמוקה יותר מn.

לפיכך אם h(n)=d אזי h(n')=d+x כאשר x הינו שלם חיובי. כלומר, d ≤ c(n,n') + d+x.
c(n,n') אינו שלילי ולכן בכל מקרה יתקיים h(n) ≤ c(n,n') + h(n').

- אחת מהנקודות לפחות לא נמצאת כלל בעץ.

אמרנו שבמקרה ונקודה לא נמצאת בעץ ערך ההיוריסטיקה שנקבע לה הוא 100,000. אם אחת מהן לא נמצאת בעץ אזי שמרחק העלות בין n ל-'n הוא אינסוף ולכן תמיד יתקיים h(n) ≤ inf.

הראיתי שבכל המקרים מתקיים h(n) ≤ c(n,n') + h(n') ולכן ניתן לומר שההיוריסטיקה שבחרתי היא עקבית.

- ההיוריסטיקה עקבית וקבילה. למדנו בכיתה שבשימוש בעץ חיפוש עם היוריסטיקה קבילה הפתרון המתקבל הוא אופטימלי, ומכאן שהפתרון שהמסלול שיימצא בהכרח יהיה הקצר ביותר ☺

2. נניח עץ בעל n קודקודים כאשר לכל קודקוד יש בן יחיד ושקודקוד יעד שלנו נמצא בעומק n.

<u>אלגוריתם DFS</u>

יבצע חיפוש לעומק. הוא יתחיל סריקה בראש העץ, וימשיך לסרוק את קודקוד הבן שלו, ואז את הבן של הבן וכו' עד שיגיע לשכבה ה-n וימצא שם את הפתרון וייעצר. למעשה האלגוריתם יעבור קודקוד קודקוד, סה"כ n קודקודים ומכאן זמן הריצה $1 + 1 + 1 + \cdots + 1 = O(\mathrm{n})$

<u>אלגוריתם IDS</u>

כפי שלמדנו מדובר בחזרה אלטרנטיבית של אלגוריתם DLS (DFS המוגבל לסריקה עד עומק מסויים), אשר בכל פעם מגדילים את העומק שבו מחפשים את הפתרון. מכיוון שהאלגוריתם לא יודע באיזה

אור נחשון אגש
311330500

עומק נמצא הפתרון הוא יתחיל בלחפש עבור עומק 1, אם לא יימצא שם יבוצע DLS עד לעומק 2, אם לא ימצא שם ימשיך לעומק 3 וכו'. לפי הדוגמא שלנו הפתרון יימצא רק כאשר האלגוריתם יגיע לשכבה ה-n, כלומר עוד קודם לכן לפני שיגיע לשכבה הוא יבצע את הצעדים הבאים $1 + 2 + 3 + 4 + \cdots + n = \frac{n*(n+1)}{2}$ ולכן זמן הריצה יהיה $O(n^2)$.

אור נחשון אגש
311330500

קורס יסודות בינה מלאכותית
מרצה ד"ר עומר לב

בס"ד

נספחים:

נספח 1: יחסי השכנות בין המדינות


```
{'OH':['WV','IN','MI','PA','KY'],
'OK':['NM','CO','KS','TX','AR','MO'],
'OR':['WA','ID','CA','NV'],
'IA':['SD','NE','MO','IL','WI','MN'],
'ID':['WA','OR','NV','UT','MT','WY'],
'IL':['KY','IN','MO','IA','WI'],
'IN':['OH','MI','KY','IL'],
'AL':['FL','GA','TN','MS'],
'AK':[],
'AZ':['CA','NM','NV','UT','CO'],
'AR':['OK','TX','LA','MS','TN','MO'],
'GA':['NC','SC','FL','AL','TN'],
'DE':['PA','MD','NJ'],
'SD':['WY','MT','ND','MS','LA','NE'],
'ND':['MT','SD','MS'],
'HI':[],
'WY':['UT','CO','ID','NE','SD','MT'],
'WA':['OR','ID'],
'WI':['IL','IA','MN','MI'],
'VA':['MD','WV','KY','NC','TN','DC'],
'WV':['OH','PA','MD','VA','KY'],
'VT':['NH','NY','MA'],
'TN':['VA','NC','GA','AL','MS','AR','KY','MO'],
'TX':['NM','OK','AR','LA'],
'UT':['AZ','NV','CO','ID','WY','NM'],
'LA':['TX','AR','MA'],
'MT':['ID','WY','SD','ND'],
'MO':['NE','KS','OK','TN','KY','AR','IL','IA'],
'ME':['NH'],
'MN':['ND','SD','IA','WI'],
'MS':['LA','AL','TN','AR'],
'MI':['OH','IN','WI'],
'MA':['NH','VA','CT','NY','RI'],
'MD':['PA','DE','VA','WV','DC'],
'NV':['OR','CA','AZ','ID','UT'],
'NE':['WY','SD','KS','MO','CO','IA'],
'NJ':['NY','PA','DE'],
'NH':['ME','VT','MA'],
'NY':['VT','MA','CT','NJ','PA'],
'NM':['AZ','CO','TX','OK','UT'],
'FL':['GA','AL'],
'PA':['OH','NY','WV','MD','DE','NJ'],
'CO':['NM','UT','WY','KS','OK','NE','AZ'],
'CT':['MA','NY','RI'],
'CA':['OR','AZ','NV'],
'KS':['CO','NE','MO','OK'],
'KY':['OH','WV','VA','TN','IN','IL','MO'],
'SC':['NC','GA'],
'NC':['VA','TN','SC','GA'],
'RI':['MA','CT'],
'VI':[],
'PR':[],
'MP':[],
'GU':[],
```

```
'AS':[],
'DC':['MD','VA']
}
```