

תרגיל בית 2 מבוא לבינה מלאכותית
חיפוש רב סוכנים
בינה מחסנים



הקדמה ואדמיניסטרציה

הנחיות כלליות

- תאריך הגשת התרגיל: **25.6.2023**
- את המטלה יש להגיש בזוגות בלבד – בקשות להגשה ביחידים באישור המתרגל האחראי בלבד (ספיר טובול).
- יש להגיש את המטלה מוקלדת בלבד – פתרון בכתב יד לא ייבדק.
- התשובות צריכות להיות כתובות בשפה העברית או באנגלית.
- אפשר לשלוח שאלות בנוגע לתרגיל דרך הפיאצה.
- המתרגלת האחראית על התרגיל: **אופק גוטליב**.
- בקשות דחיה מוצדקות יש לשלוח למתרגל האחראי בלבד.
- במהלך התרגיל ייתכן שיעלו עדכונים – הודעה תפורסם בהתאם במקרה זה.
- העדכונים מחייבים וזוהי אחריותכם להתעדכן לגביהם עד מועד הגשת התרגיל. עדכונים יופיעו בטופס בצבע צהוב.
- העתקות טופלנה בחומרה.
- ציון המטלה כולל חלק יבש וחלק רטוב. בחלק היבש נבדוק שתשובתכם נכונה, מלאה, קריאה ומסודרת. בחלק הרטוב הקוד שלכם ייבדק בצורה מקיפה באמצעות בדיקות אוטומטיות – אשר ישוו את המימוש שלכם למימוש שלנו. חשוב לעקוב בזירות אחר הוראות התרגיל מאחר שסטיות מהמימוש המבוקש עלולות להוביל לכשל בטסט האוטומטי (גם אם המימוש "נכון ברובו"). במהלך הבדיקה האוטומטית יינתן זמן סביר לכל הרצה – כך שכל עוד תעקבו אחר ההוראות, המימוש שלכם יעמוד בהגבלות הזמנים ויחזיר תוצאות טובות מספיק בשביל הטסט.
- מומלץ להסתכל בקוד בעצמכם. שאלות בסיסיות על פייתון שלא נוגעות לתרגיל כדאי לבדוק באינטרנט לפני שאתם שואלים בפיאצה. מומלץ לקרוא את הקוד הנתון על מנת להבין את אופן פעולתו – במקרה שישנם דברים לא מובנים. (לשם כך יש הערות רבות ואף הסבר מורחב על הסביבה!)
- מומלץ לא לדחות את התרגיל לרגע האחרון מאחר שהמימוש וכתובת הדו"ח עלולים לקחת יותר זמן מהצפוי.
- התייחסות בלשון זכר, נקבה או רבים מתייחסים כלפי כלל המינים.

הוראות הגשה

בתוך קובץ זיפ עם השם: HW2_AI_id1_id2.zip
את הדו"ח היבש בפורמט הבא: id1_id2.pdf
ואת הקובץ: submission.py שבו אתם ממשים את האלגוריתמים

הקדמה

בינה מחסנים רוצים להפוך את המחסן לאוטונומי והם מתלבטים מיהו הרובוט אותו הם רוצים לשכור למשימה. על הרובוטים רובוט R1 ו-R2 להתחרות אחד בשני והמנצח יתקבל לעבודה. בתרגיל זה תממשו ותחקרו אלגוריתמי משחק סכום אפס אדברסריאלים שלמדתם בהרצאות ובתרגולים.

תיאור המשחק

המשחק מתרחש בלוח משבצות בגודל 5X5 כאשר על הלוח: 2 רובוטים, 2 תחנות הטענה, 2 חבילות ו-2 יעדים (אחד לכל חבילה).
לכל רובוט יחידות הטענה (battery), וניקוד (credit).

מטרת כל רובוט לצבור יותר ניקוד מהרובוט האחר עד סוף המשחק, המשחק נגמר כאשר לאחד מהרובוטים נגמרת הסוללה או כאשר נגמר מספר הצעדים המקסימלי לכל רובוט (ערך מוגדר מראש, דוגמה בהמשך).

ניקוד נצבר כאשר רובוט מביאה חבילה אל היעד שלה. כאשר רובוט מביאה חבילה ליעד הוא מקבל N יחידות ניקוד כאשר N הוא מרחק מנהטן בין מיקום החבילה המקורי ויעדה של החבילה כפול 2. לאחר שחבילה מגיעה אל היעד שלה היא והיעד שלה נעלמים ובמקומם מופיעים יעד וחבילה חדשים. כלומר, בכל עת על הלוח יש בדיוק שתי חבילות ושני יעדים, אחד לכל חבילה.

רובוטים מתחילים עם טעינה וללא חבילה. כל פעולה של הרובוט עולה לו יחידת סוללה אחת. רובוט יכול לנוע למעלה, למטה, ימינה, שמאלה. בנוסף, רובוט יכול להטעין את הסוללה בתחנת הטענה כאשר הוא נמצא במשבצת של תחנת הטענה, הוא עושה זאת ע"י המרה של כל יחידות הניקוד שלו ליחידות טעינה. רובוט יכול לאסוף ולהוריד חבילה. רובוט אוסף חבילה כאשר הוא עומד באותה משבצת כמוהו ומבצע פעולת אסיפה. רובוט מוריד חבילה כאשר הוא נמצא ביעד ומבצע פעולת הורדה. (אי אפשר להוריד חבילה שנאספה במשבצת שאינה היעד)

סה"כ לרובוט 7 פעולות אפשריות בכל עת:

move north, move south, move east, move west, pick up, drop off, charge

לא יתאפשר צעד למשבצת לא חוקית (משבצת לא חוקית היא משבצת מחוץ לגבולות הלוח או כזו שנמצא בה הרובוט השני). לא יתאפשר הטענה, אסיפה, הורדה במשבצות לא חוקיות. כל סוכן יכול להטעין בכל אחת מהתחנות הטענה.

הסבר על הסביבה והרצת / דיבוג המשחק

בשונה מתרגיל בית 1 הפעם תעבדו עם קבצי py ולא מחברות pynb. מוזמנים לעבוד על התרגיל בכל IDE מתאים שנוח לכם לעבודה בפייתון, אנו ממליצים על pycharm של חברת JetBrains אליה יש לכם מנוי מטעם הטכניון.

הסביבה שאיתה תעבדו ממומשת בקובץ WarehouseEnv.py מוזמנים לעיין בה.

בקובץ Agent.py ממומשים סוכנים מהם הסוכנים שאתם תממשו יורשים, מומלץ להסתכל על הסוכנים הממומשים בה בכדי להבין כיצד הם עובדים עם הסביבה. מומלץ להסתכל על הפונקציות `get_legal_operator` - `apply_operator` בכדי להבין את אופן ההתממשקות שלכם עם הסביבה. שימו לב שבקובץ submission יש סוכן שנקרא hardcoded, מטרתו לעזור לכם להבין את הסביבה.

הרצה

הרצת משחק נעשית ע"י שורת הפקודה הבאה שמריצים מהטרמינל:

```
python main.py greedy random -t 0.5 -s 1234 -c 200 --console_print --screen_print
```

- הארגומנט הראשון (במקרה הזה greedy) מציין את האלגוריתם שלפיו ישחק agent0
- הארגומנט השני (במקרה הזה random) מציין את האלגוריתם שלפיו ישחק agent1
- הגבלת זמן לצעד t- מקבלת ערך שמייצג את מספר השניות המקסימלי לצעד
- גרעין לקבלת ערך רנדומלי s- מקבל ערך שעוזר לחולל סביבה באופן רנדומלי, כאשר מועבר אותו ערך seed תחולל אותה סביבה
- מספר הצעדים המקסימלי עבור סוכן c-
- ערך --console_print דגל אופציונלי, כאשר מועבר מתבצעת הדפסה לקונסולה של המשחק שנראת כך :

```
robot 0 chose move north
[R1][P0][ ][C1][C0]
[D1][ ][ ][ ][ ]
[R0][P1][ ][ ][ ]
[ ][ ][ ][ ][ ]
[ ][ ][ ][D0][ ]
robots: [position:(0, 2) battery: 19 credit: 0 package: [None], position:(0, 0) battery: 20 credit: 0 package: [None]]
packages on street: [position:(1, 0) destination: (3, 4), position:(1, 2) destination: (0, 1), position:(0, 3) destination: (3, 3), position:(2, 3) destination: (0, 1)]
(1, 0) (3, 4) True
(1, 2) (0, 1) True
(0, 3) (3, 3) False
(2, 3) (0, 1) False
charge stations: [position:(4, 0), position:(3, 0)]
```

מודפס מספר הסוכן יכול להיות 0 או 1 והאופרטור בו בחר. לאחר מכן מודפס הלוח לאחר שהסוכן הפעיל את האופרטור. הלוח כולל 25 משבצות כאשר בכל אחת יכולים להיות רובוט, חבילה המחכה לרובוט או תחנת הטענה.

האותיות אשר מופיעות בלוח מסמלות:

R – Robot

C – Charge station

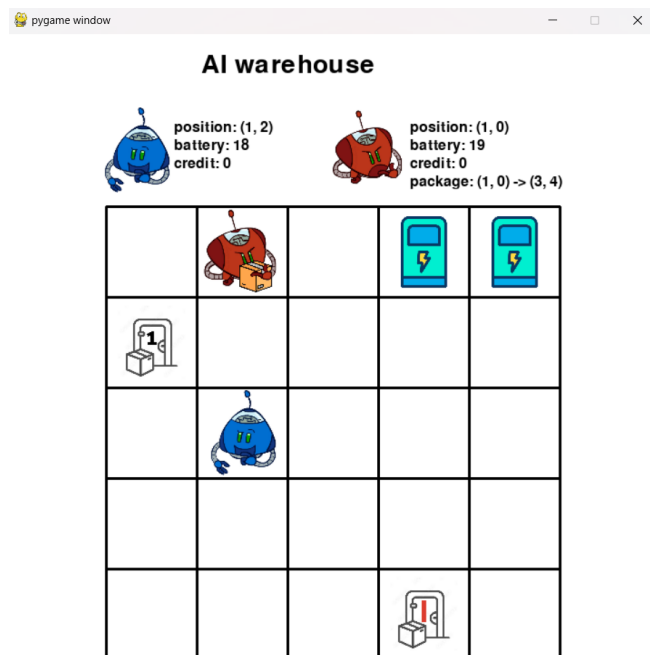
P – Package on street

D – Destination of Package on street

X – Destination of Package taken by Robot

המספר המופיע ליד כל אחת מהאותיות הוא המזהה של האובייקט אליו הוא משתייך – לדוגמא: R1 עבור הרובוט הראשון. עבור חבילה שנאספה ע"י רובוט, המספר ליד X הוא המזהה של הרובוט שאסף אותה. עבור חבילה במרחב, המספר ליד D הוא המזהה של החבילה במרחב. לעיתים מספר אובייקטים מופיעים באותו המיקום בתמונה ואז מודפס רק אחד מהם. במקרה זה, אפשר להשתמש ברשימות המפורשות שמופיעות אחרי הלוח. בנוסף שימו לב! מופיע לכם ברשימות את המקור והיעד של שתי החבילות הבאות שיופיעו על הלוח, מוזמנים ומומלץ להשתמש במידע זה בהמשך התרגיל.

- ערך `--screen_print` דגל אופציונלי, כאשר מועבר יודפס לכם הנפשה של המשחק בחלון pygame שנראה כך:

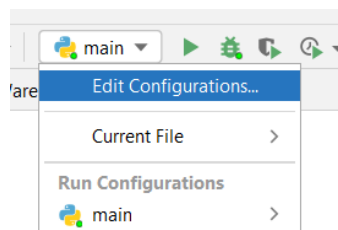


בנוסף בכדי לבדוק את הסוכנים שלכם על מספר רב של משחקים, אתם יכולים להוסיף את הדגל `--tournament` וכך תקבלו הרצה רצופה של n משחקים. מוזמנים לשחק עם המשתנה של מספר המשחקים ולראות האם מקבלים התנהגויות רצויות עבור האלגוריתמים.

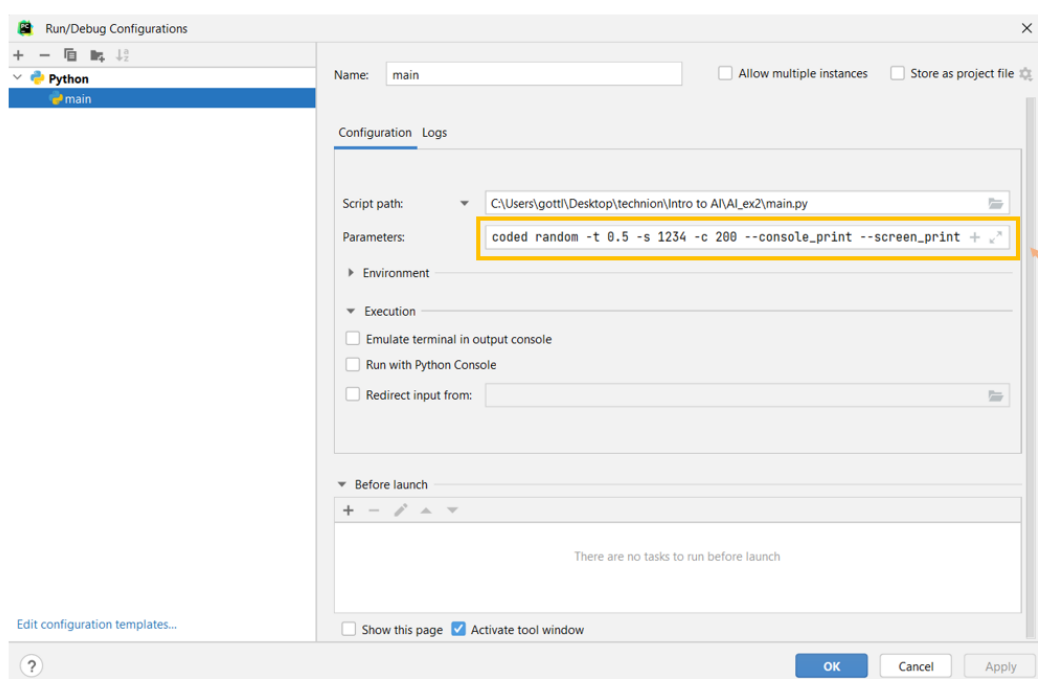
דיבוג

בכדי לדבג את המשחק עליכם לעקוב אחר השלבים הבאים:

1.



2.



מתחילים לכתוב !

חלק א - ImprovedGreedy

1. (יבש: 4 נק') כפי שלמדתם, אנו מגדירים בעייה במרחב בתור רבעייה (S, O, I, G). הגדירו פורמלית את המשחק המתואר לכם ע"פ הנתונים שאתם מקבלים מהסביבה.

2. (יבש: 4 נק') הגדירו היוריסטיקה משלכם להערכת מצבי המשחק. עליכם לתעד אותה בנוסחה מפורשת ועלייה לכלול לפחות שלושה מאפיינים של הסביבה. בחרו שמות ברורים בנוסחה שלכם.

3. (רטוב: 10 נק') ממשו בקובץ submission.py את הפונקציה `smart_heuristic` שבה משתמש הסוכן AgentGreedyImproved

4. (יבש: 2 נק') מהו החיסרון העיקרי של האלגוריתם? (לעומת minimax)

חלק ב - RB-Minimax

1. (יבש: 3 נק') מה היתרונות והחסרונות של שימוש בהיוריסטיקה קלה לחישוב לעומת היוריסטיקה קשה לחישוב בהינתן שהיוריסטיקה הקשה לחישוב יותר מיועדת מהקלה לחישוב? בהינתן שאנו באי-מקסימום מוגבל משאבים.
2. (יבש: 4 נק') חברתכם לקורס דנה מימשה סוכן minimax, היא שמה לב כי לעיתים הסוכן יכול לנצח בצעד אחד אך הוא בוחר בצעד אחר. האם יש לה באג באלגוריתם? אם אין באג הסבירו מה באלגוריתם גורם להתנהגות שכזו. אם יש באג מה הוא יכול להיות?
3. (רטוב: 10 נק') עליכם לממש את המחלקה AgentMinimax בקובץ submission.py. שימו לב! הסוכן מוגבל משאבים, כאשר המשתנה time_limit מגביל את מספר השניות שהסוכן יכול לרוץ לפני שיחזיר תשובה. (הגבלת הזמן עלייה אתם נבדקים הינה שנייה כלומר 1-t).
4. (יבש: 3 נק') נניח שבסביבה היו K שחקנים במקום 2 (תחשבו על משחק כללי לזוג דווקא המשחק שלנו, אך עדיין משחק סכום אפס). אילו שינויים יהיה צריך לעשות במימוש סוכן Minimax?
 - a. בהינתן שכל סוכן רוצה לנצח ולא אכפת לו רק ממכם.
 - b. בהנחה והדבר היחיד שכל סוכן רוצה הוא שלא תנצחו.
 - c. בהנחה שכל סוכן רוצה שהסוכן שאחריו בתור ינצח.

חלק ג - Alpha-Beta

1. (רטוב: 10 נק') ממשו שחקן אלפא - בטא מוגבל משאבים במחלקה AgentAlphaBeta בקובץ submission.py, כך שיתבצע גיזום כפי שנלמד בהרצאות ובתרגולים.
2. (יבש: 3 נק') האם הסוכן שמימשתם בחלק זה יתנהג שונה מהסוכן שמימשתם בחלק ב מבחינת זמן ריצה ובחירת מהלכים? הסבירו.

חלק ד - Expectimax

1. (יבש: 3 נק') בהנחה ואתם משתמשים באלגוריתם Expectimax נגד סוכן שמשחק באופן רנדומלי לחלוטין באיזה הסתברות תשתמשו? ומדעו?

2. (יבש: 4 נק') עבור משחקים הסתברותיים כמו שש בש, בהם יש מגבלת משאבים, משתמשים באלגוריתם RB-Expectimax. הניחו כי ידוע שהפונקציה היוריסטית h באלגוריתם Expectimax-RB מקיימת $\forall s: -1 \leq h(s) \leq 1$. איך ניתן לבצע גיזום לאלגוריתם זה? תארו בצורה מפורטת את תנאי הגזימה, והסבירו את הרעיון מאחוריו.

3. (רטוב: 10 נק') הסוכן של α - β ו-minimax מניח שהסוכן היריב יבחר באופרטור שיוביל לתוצאה האופטימלית בתורו, אולם זה לא תמיד מתרחש. לדוגמה, כאשר אנו מתחרים עם סוכן חמדן, סביר להניח שהוא לא יבחר בפעולה האופטימלית בכל צעד. אפשר להתחשב באפשרות שהיריב יבחר בפעולה שאינה אופטימלית בתורו באמצעות סוכן Expectimax. נבחר את המשקולות של כל אופרטור בצורה שתתחשב בכך שכאשר היריב יכול לפנות אל תחנת דלק הוא יפנה אליה בהסתברות יותר גבוהה. אם הסוכן יכול לפנות למעלה, למטה, ימינה, שמאלה, לאסוף נוסע, להוריד, להטעין ולידו יש עוד תחנת דלק הוא יעדיף בהסתברות של פי 2 את התנועה לכיוון תחנת דלק. ממש שחקן ExpectiMax מוגבל משאבים המתנהג באופן המתואר.

חלק ה - משחק עם פקטור סיעוף גדול

1. (יבש: 6 נק') להלן שינויים אפשריים ששוקלים בינה מחסנים לעשות במשחק בכדי לבחון יכולות נוספות של הרובוטים. עבור כל שינוי ציינו מה ההשפעה שלו על מקדם הסיעוף וחשבו את מקדם הסיעוף החדש המתקבל.

a. הגדלת לוח המשחק להיות 8×8 והוספת מחסומים בסביבה. (מחסומים משמע משבצות שהסוכן לא יכול לעבור בהן)

b. הוספת היכולת של רובוט בכל תור לבחור משבצת על הלוח ולהניח עלייה בלוק, משמע בכל תור יכול הרובוט לנוע למעלה, למטה, ימינה, שמאלה, לאסוף חבילה, להוריד חבילה, להטעין, ולהניח בלוק על הלוח, בלוק יכול להיות מונח על כל משבצת ריקה.

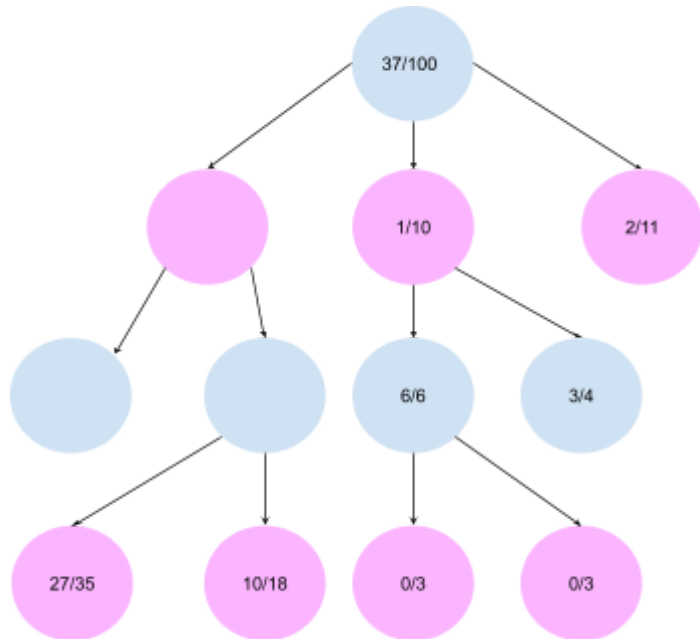
2. (יבש: 6 נק') בהנחה ומימשו את השינוי השני עבור הסביבה (סעיף 1b)

a. האם יש אלגוריתם מהסעיפים הקודמים שנוכל להשתמש בו שזמן הריצה שלו סביר? (סביר משמע לא גדול מהותית מהזמן שלוקח לו להחזיר צעד עבור המשחק בלי השינוי).

b. הציעו אלגוריתם שונה מאלו שמשתמם בסעיפים הקודמים שנלמד בקורס שירון בזמן סביר. הסבירו מדוע בחרתם בו ולמה הוא טוב להתמודדות עם האתגר שנוצר משינוי הסביבה.

חלק ו - יבש - שאלה פתוחה - MCTS

דנה וטל שיחקו משחק, תור של דנה מסומן בכחול ותור של טל מסומן בורוד. להלן עץ המשחק שמתאר את העץ שנוצר בשלב ביניים בהרצת MCTS עם פיתוח צמתים לפי UCB1 על משחק סכום אפס בין שניהם. הניחו כי המשחק בעל אופק מקסימלי של 10 צעדים, נתון $C = \sqrt{2}$



1. (5 נק') חלק מהערכים בצמתים נמחקו, השלימו את החלקים החסרים, אין צורך לנמק.
2. (5 נק') מיהו הצומת הבא שיבחר בשלב ה - selection ? הוסיפו חישובים
3. (5 נק') בהנחה שכל סימולציה מכאן והלאה מסתיימת בניצחון של טל, ונגדיר אב קדום ביותר להיות צומת שקרוב ביותר לשורש (a,b,c) מהו מספר הניצחונות המינימלי הדרוש כדי שצאצא של אב קדום ביותר אחר יבחר בשלב ה - selection ? (לעומת סעיף 2), הוסיפו את החישובים הרלוונטיים
4. (3 נק') כעת רוצים לבצע שינוי כך שנעדיף exploration יותר מ - exploitation. הגישה לנוסחה שמחשבת את ה - UCB1 חסומה לכם, אך הנוסחה משתמשת ב - $N(s)$ אשר אליו יש לכם גישה ואתם יכולים לשנותו. כיצד תשנו אותו בכדי שהנוסחה החדשה שתיווצר תעדיף יותר exploration מ - exploitation לעומת הנוסחה הקודמת.