תרגיל 2

**Introduction to Artificial Intelligence (236501)**

מגישים: אייל אמדור, בארי זיטלני

318849270, 209351626ת.ז:

ImprovedGreedy

שאלה 1

נפרמל את הבעיה:

* S – מרחב המצבים, כאשר נגדיר כך:  
  כאשר נגדיר מושגי עזר:  
  ונגדיר:
  + – המיקום הנוכחי שלי בלוח, יוגדר להיות tupple באופן הבא:
  + battery – כמה סוללה נשארה לרובוט, ערך גדול/שווה 0.
  + credit – כמה נקודות צבר הרובוט, ערך גדול/שווה 0.
  + package – חבילה שיש ברשות הרובוט, יכול לקבל כערך:
* O – פונקציית מעברים, כאשר נגדיר:  
  כאשר:  
  כלומר בכל תור משחק בדיוק סוכן אחד.
* I – קבוצת המצבים ההתחלתיים, כאשר נגדיר:  
  כאשר,
* G – קבוצת מצבים הסופיים:  
  כלומר או שלאחד הרובוטים נגמרה הסוללה או שקצבת הצעדים הכוללת למשחק נגמרה.

שאלה 2

נגדיר את היוריסטיקה הבאה עבור :

שאלה 3

מימוש בקוד.

RB-Minimax

שאלה 1

בהינתן מינמקס מוגבל במשאבים, שימוש ביורסטיקה קלה לחישוב ייתן עץ פעולות עמוק יותר ולכן ייצג בחירה על מרחב אפשרויות רחב יותר – כלומר ניבוי שחוזה בצורה מדויקת יותר את "עתיד" המשחק. לעומת זאת, היורסטיקה הינה קלה ולכן מתארת פחות נאמנה את המציאות, כלומר מביאה לידי ביטוי פחות פרמטרים או מורכבויות שצעדים מסוימים עלול ליצור.

מנגד, שימוש ביורסטיקה כבדה לחישוב ייתן עץ פעולות עמוק פחות ולכן ייצג בחירה על מרחב אפשרויות מוגבל יותר כלומר ינבא פחות מהשלכות הצעד במשחק ("רואה פחות רחוק"). לעומת זאת, היורסטיקה הכבדה מתארת בצורה מהימנה יותר את המציאות ולכן השערוך היורסטי טובה יותר.

שאלה 2

ייתכן ולדנה אין באג באלגוריתם. כתלות במימושה לאלגוריתם מינימקס, אין התחייבות של האלגוריתם לבחירת הניצחון המהיר ביותר, לכן אם יש לסוכן של דנה אפשרות לנצח בתור הנוכחי והיא לא נבחרת ייתכן וקיימת אפשרות לניצחון בהמשך המשחק והאלגוריתם בוחר בה.

שאלה 3

אלגוריתמי any-time הם אלגוריתמים שיכולים לשפר את ביצועיהם בהינתן משך ריצה ארוך יותר. במקרה הנוכחי, הגבלת זמן במקום הגבלת עומק תתבטא בהרצת מינימקס עד לעומק L=1 ובסיום כל איטרציה של האלגוריתם נגדיל את L ב-1 ונריץ מחדש את RB-Minimax עם L גדול 1.

אלגוריתם any-time נוסף שנלמד בקרוס הינו ID-DFS שמריץ DFS לעומק מוגבל לזמן ריצה מוגבל.

שאלה 4

מימוש בקוד.

שאלה 5

במשחק מרובה משתתפים (שהוא משחק סכום אפס) יש יותר שחקנים שרוצים להרע לי, כלומר יותר צמתי OR בעץ, כתלות במספר השחקנים (NUM\_OF\_PLAYERS) ולכן נעשה התאמה לפסודו-קוד מהתרגול:

Function RB-Minimax-Multiple-Players (State, Agent, depth, turn\_count = 0):

If G(State) OR depth=0 then return h(State,Agent)

ChildrenSucc(State)

if turn\_count % NUM\_OF\_PLAYERS == 0: // the turn is mine, choose max

CurMax

Loop for c in Children:

v RB-Minimax-Multiple-Players(c, Agent, depth-1, turn\_count+1)

CurMaxMax(v, CurMax)

Return CurMax

else: //the turn is not mine, choose min

CurMin

Loop for c in children:

v RB-Minimax-Multiple-Players(c, Agent, depth-1, turn\_count+1)

CurMinMin(c, CurMin)

return CurMin

Alpha-Beta

שאלה *1*

*מימוש בקוד.*

*שאלה 2*

*מבחינת זמן ריצה, גם הסוכן הזה וגם הסוכן הקודם מוגבלים ע"י אותו חסם ולכן הם עתידים להתנהג בצורה דומה. נציין כי במקרה שבו סוף המשחק (מצב שבו לשני השחקנים אין בטרייה) קרוב, סוכן אלפא-בתא יכול להגיע לפיתוח המצבים האלה מהר יותר ולכן גם לסיים את הריצה מהר יותר.*

*מבחינת בחירת המהלכים, סוכן אלפא-בטא מצליח באותו זמן לפתח עץ עמוק יותר ולכן בחירת המהלכים שלו הינה מושכלת יותר. כלומר, בהחלט ייתכן כי יבחר לבצע מהלכים שונים מאשר הסוכן הקודם.*

Expectimax

שאלה 1

כאשר שחקן מקסימום ימצא בצאצאיו צומת בעל ערך 1, הוא יכול לוותר על פיתוח שאר הבנים שלו. זאת מכיוון שמובטח לו שכל ערך אחר בשאר הבנים שלו חסום על ידי כלומר לא יימצא בהם ערך הגבוה מ- 1 (ולכן גם חישוב התוחלת יחזיר ערך בין 1 ל1-) ולכן נעדיף לגזום אותם.   
באופן דומה, בצומת מינימום, אם שחקן מינימום מוצא צומת בעל ערך של 1-, גם הוא יכול להימנע מפיתוח שאר הבנים שלו. שוב מאותה סיבה, אין אפשרות למצוא ערך נמוך יותר בהמשך הפיתוח ולכן נעדיף לגזום את יתר הבנים.

שאלה 2

מימוש בקוד.

משחק עם פקטור סיעוף גדול

שאלה 1

1. *פקטור הסיעוף לא ישתנה. הגדלת לוח המשחק לא משנה את מספר הפעולות האפשרי בכל צעד (צפון, דרום, מזרח, מערב, הורדת חבילה, איסוף חבילה וטעינה). בנוסף, הוספת מחסומים לא משנה גם היא את פקטור הסיעוף שהוא מספר הפעולות המקסימלי (מחסום עלול לשנות את מספר הפעולות עבור צעד ספציפי בדומה לעמידה בצמוד לקיר שלא מאפשרת בתקדמות בכיוון הקיר אך לא משנה את פקטור הסיעוף).*
2. *פקטור הסיעוף יגדל ב-23. במקרה המקסימלי בו שני הרובוטים מחזיקים את החבילות ועומדים על תחנות העגינה שבמקרה זה גם תחנות הורדת החבילה, נוספו 23 פעולות אפשריות: הנחת בלוק בכל משבצת פנויה (25 משבות סה"כ מתוכן רק 2 לא ריקות). לכן פקטור הסיעוף גדל ב23.*

*שאלה 2*

1. *כפי שנלמד בהרצאות, אלגוריצם* GREEDY *הוא לינארי ולכן גם עם השינוי הנ"ל זמן הריצה שלו לא ישתנה בצורה משמעותית. לעומת זאת, אלגוריתמי ה-*minimax *לסוגיהם משתמשים בעצים ולכן זמן הריצה שלהם הוא אקספוננציאלי למקדם הסיעוף, כלומר עם השינוי הנ"ל הריצה שלהם תתארך בצורה משמעותית.*
2. MCTS – *כפי שלמדנו בהרצאה אלגוריתם* MCTS *ממקסם את היחס בין ה-*exploration *ל-*exploitation *ולכן ימנע חיפושים מיותרים בתוך העץ. סה"כ האלגוריתם מאפשר חיפוש יעיל בתוך עץ ההחלטות ולכן ירוץ בזמן סביר.*

שאלה פתוחה – MCTS

שאלה 1

1. *נחשב:*
2. *לפי הגדרת האלגוריתם הוא ייבחר את הזרוע שממקסמת את UCB ולכן ייבחר בזרוע a2*

*שאלה 2*

*כעת נחשב:*

*שאלה 3*

*התשובה הנכונה היא: 2. כן, עבור כל ערך מתקיים כי מתישהו ייבחר, אולם לא עבור .  
זאת מכיוון שכדי ש- ייבחר נדרוש שערך ה-*UCB *שלו יהיה גדול יותר משל ולכן אם נגדיל את מספר הפעמים שנבחר את מספיק נקבל שערך ה-*exploration *שלו יקטן מספיק בעוד שהערך של ייגדל מספיק ולכן נבחר בו לכל . אם אזי שערך ה-*exploration *לא משפיע כלל ולכן תמיד נעדיף את ערך ה-* exploitation *שהוא תמיד יעדיף את (בעל ערך 1 לעומת 0 של ). מתמטית נדרוש:*

*כאשר מגדילים רק את אזי האגף הימני מתנהג כך:*

*לכן, כפי שאפשר לראות לכל מתקיים שלאחר מספר גדול מספיק של משיכות המשוואה תתקיים ובפרט כלומר הזרוע של תבחר.*

שאלה פתוחה – אלגוריתם מונטה קרלו למשחקי אינפורמציה חלקית

שאלה 1

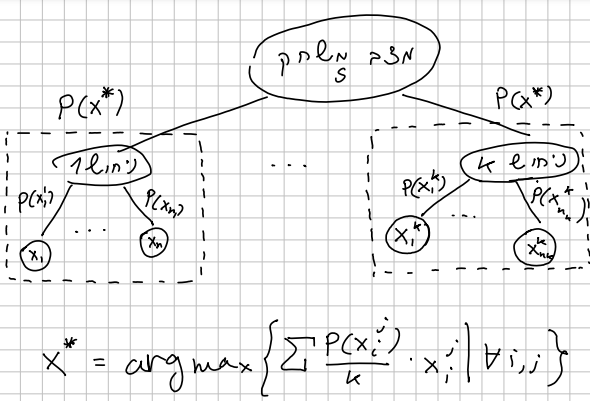
*הבעיה שהלאגוריתם מנסה לפתור היא חוסר מודעות למצב המשחק המלא, בדרך כלל מצב המשחק של היריב לא ידוע ולכן כדי לפתור זאת האלגוריתם ינסה לנחש את k מצבים אפשריים של היריב ובהתאם להם יחליט על דרך פעולה. (למה k ? זאת מטעמי הגבלת משאבים שכן פיתוח כל המצבים האפשריים יהיה יקר וארוך מאוד).*

שאלה 2

*בחירת ערך k גדול מתארת יותר מצבים אפשריים ולכן הסיכוי שלה לנחש נכון את המצב הנוכחי גדול יותר. עם זאת בחירת k גדול עולה יותר משאבי זכרון ותקח יותר זמן להרצת האלגוריתם.*

*עבור k קטן האלגוריתם מנחש פחות מצבים ולכן הסיכוי שלה לנחש נכון את המצב הנוכחי נמוך יותר, אך היא דורשת פחות משאבי זכרון וזמן ריצתה יהיה מהיר יותר.*

שאלה 3

*באלגוריתם* Expectimax *בכל צומת בחירה של השחקן שלי יש* N *אופציות ליריב לשחק ולכל אחת מהן הסתברות מסויימת עם ערך מתאים לה, כאשר האלגוריתם בוחר את האופציה עם ערך התוחלת הגבוה ביותר. נציע מודל שמשתש ב-*Expectimax *באופן הבא:   
מצומת המצב הלא ידוע נבנה* k *בנים* *עבור כל ניחוש של מצב אפשרי. לכל בן כזה נבנה בנים עבור כל פעולה אפשרית של היריב ונחשב את ההסתברות עבור כל פעולה. נחשב עבור כל הפעולות בענף את הפעולה עם התוחלת הגבוהה ביותר (כאשר לכל פעולה עם הסתברות* p(x) *תהיה הסתברות כשמסתכלים על כל הענף), נסמנה ב-*x\*. *כעת, לכל ניחוש אפשרי יש ערך הסתברות של הפעולה הנבחרת, p(x\*) תחת הניחוש הנוכחי, נתייחס לכל בן עם הפעולה כצומת בעץ ה-*Expectimax *ונבחר את הבן עם ההסתברות הגבוהה ביותר.*

*אלגוריתם מונטה קרלו ייבחר את הפעולה עם ערך הגבוה ביותר בנוסחת מונטה קרלו, כאשר במודל המוצע נבחרת הפעולה עם ערך התוחלת הגבוה ביותר שבסבירות גבוהה תקרה. משקילות הנוסחאות (פעולה נבחרת תחת מונטה קרלו אמ"מ היא נבחרת תחת ערך תוחלת גבוה) נקבל שקילות בין האלגוריתמים.*

שאלה 4

רעיון: נריץ אלגוריתם MCTS עם ערך k=1 וכל עוד יש לנו זמן נריץ שוב את האלגוריתם עם k גדול ב1.  
כשיגמר הזמן נחזיר את הפתרון עבור ריצת האלגוריתם ה-k האחרון שסיים את ריצתו.

פסודו קוד:

        depth = 1

        while current\_time < TIME\_LIMITATION:

          op = MCTS(partial\_state, agent, depth, k, TIME\_ LIMITATION)

            depth += 1

        return op

* העברנו את ה- TIME\_ LIMITATION לאלגוריתם-MCTS כדי שנוכל לקטוע אותו באמצע הריצה במידה והזמן חרג.