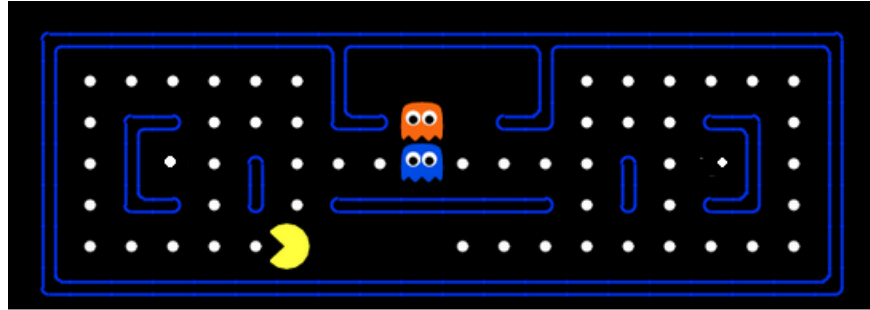


תרגיל בית 3 – תכנון סטוכסטי ללא מודל

מבוא

מטרת התרגיל היא להתנסות בפתרון בעיות סטוכסטיות על דוגמת ה-pacman, בתנאי חוסר מודל



הבעיה

הקלט של הבעיה נשאר זהה לקלט של תרגיל בית 2 – מטריצה $N \times M$ שמייצגת את לוח המשחק, וכניסות בה מייצגות את האובייקטים הנמצאים על הלוח.

בדומה לתרגיל 2, המטרה היא לאסוף כמה שיותר נקודות (כשפירוט הנקודות שמקבלים יבוא בהמשך) בתרגיל קיימת אי-ודאות בתזוזת רוחות הרפאים (לצד גורמי אי-ודאות נוספים), ולכן הפתרון של התרגיל הוא לא תוכנית, אלא מדיניות (policy). זאת אומרת בהינתן מצב העולם, עליכם להוציא את הפעולה שברצונכם לבצע

ההדבלים העיקריים מתרגיל 2 הם:

- הניקוד שאתם מקבלים על גולות לא ידוע מראש, ומשתנה בין סוגי גולות
- הסתברות השולטת על תנועה של רוחות הרפאים לא יודעה מראש
- שינוי ניקוד על איפוס וסיום שלב
- זמנים שבהם צריך לעמוד

תיאור המשימה

עליכם לבנות מחלקה PacmanController שמייצגת את הסוכן שלכם, ושיש בה (בין היתר) constructor ופונקציה choose_next_action.

ה-constructor מקבל את המצב הראשוני, וניתן לעשות בו חישובים מקדימים. ה-constructor אמור לסיים את הריצה תוך 30 שניות.

הפעולה choose_next_action מקבלת את המצב הנוכחי של העולם והפרסים הצבורים עד עכשיו, ומחזירה את הפעולה הבאה שברצונכם לבצע במצב הזה. הקוד של הבדיקה מקבל את הפעולה, מחשב את התוצאה שלה, וקורא ל choose_next_action שוב עד סיום המשחק. לכל בעיה יש זמן מוגדר שיש לסיים את כל הצעדים עד הזמן הזה. הפונקציה אמורה לסיים את הריצה שלה תוך 3 שניות.

על אחריותכם לבדוק את העמידה של הקוד שלכם בזמנים מוגדרים. אם תהיה חריגה, הסימולציה על הקלט נפסקת, ואתם מקבלים 0 נקודות על ההרצה הזאת.

ייצוג הקלט

הקלט ל-constructor הוא tuple שמכיל 2 איברים – מפה, אופק התכנון בצעדים.

כעת יש 3 סוגי גלולות, ולכן ייצוג הקלט נראה ככה:

- 99 – קיר
- 88 – שחקן נאכל ע"י רוח רפאים
- 77 – רעל
- 73 – רעל וגלולה מסוג 3
- 72 – רעל וגלולה מסוג 2
- 71 – רעל וגלולה מסוג 1
- 66 – שחקן (פאקמן)
- 53 - רוח רפאים **אדומה** במשבצת עם גלולה מסוג 3
- 52 - רוח רפאים **אדומה** במשבצת עם גלולה מסוג 2
- 51 – רוח רפאים **אדומה** במשבצת עם גלולה מסוג 1
- 50 – רוח רפאים **אדומה** במשבצת ללא גלולה
- 43 - רוח רפאים **ירוקה** במשבצת עם גלולה מסוג 3
- 42 - רוח רפאים **ירוקה** במשבצת עם גלולה מסוג 2
- 41 – רוח רפאים **ירוקה** במשבצת עם גלולה מסוג 1
- 40 – רוח רפאים **ירוקה** במשבצת ללא גלולה
- 33 - רוח רפאים **צהובה** במשבצת עם גלולה מסוג 3
- 32 - רוח רפאים **צהובה** במשבצת עם גלולה מסוג 2
- 31 – רוח רפאים **צהובה** במשבצת עם גלולה מסוג 1
- 30 – רוח רפאים **צהובה** במשבצת ללא גלולה
- 23 - רוח רפאים **כחולה** במשבצת עם גלולה מסוג 3
- 22 - רוח רפאים **כחולה** במשבצת עם גלולה מסוג 2
- 21 - רוח רפאים **כחולה** במשבצת עם גלולה מסוג 1
- 20 - רוח רפאים **כחולה** במשבצת ללא גלולה
- 13 – משבצת רגילה עם גלולה מסוג 3
- 12 – משבצת רגילה עם גלולה מסוג 2
- 11 – משבצת רגילה עם גלולה מסוג 1
- 10 – משבצת ריקה

אופק התכנון בצעדים – עליכם לבצע את **כל** הצעדים

כמו כן, בתחילת כל תור אתם מקבלים את המפה המעודכנת, והניקוד העכשווי

אופן וחוקיות התזוזה

לרשותכם 5 פעולות של השחקן:

- תזוזה ימינה – מיצוגת ע"י מחרוזת "R"
- תזוזה למטה – מיצוגת ע"י מחרוזת "D"
- תזוזה שמאלה – מיצוגת ע"י מחרוזת "L"
- תזוזה למעלה – מיצוגת ע"י מחרוזת "U"

- איפוס – מיוצג ע"י מחרוזת "reset"

כל פועלת תזוזה גורמת לפאקמן לזוז משבצת אחת בכיוון הרצוי, חוץ מתזוזה למשבצת הקיר. תזוזה כזאת מותרת, אבל לא תשנה את מיקום השחקן. חוץ מזה, אחרי זה גם רוחות הרפאים זזות. תזוזת רוחות הרפאים מתבצעת לפי החוקיות הבאה:

- כל רוח (חוץ מהכחולה) זזה משבצת אחת לאחד מ-4 הכיוונים (למעלה, למטה, ימינה או שמאלה)
- חלק מהרוחות יכולות לזוז גם באלכסון (זאת אומרת יש להן עד 8 אופציות לתזוזה). אין גישה לסוכן סדר תזוזה: קודם כל שחקן, ואז כל הרוחות בסדר אקראי
- כל רוח תזוז לכיוון הכי מקרב אותה לשחקן לפי מרחק מנהטן בהסתברות p , ולכיוון אקראי (הנבחר באופן אחיד מכיוונים מותרים) בהסתברות $1-p$, כשלכל רוח ערך p שונה
- ההסתברות p (לכל רוח הסתברות משלה) לא יודעה לסוכן מראש, ועשויה להשתנות בין קלט לקלט
- אם תזוזת הרוח ל-2 כיוונים מקרבת אותה לשחקן באותה מידה, אז סדר העדיפויות לתזוזה אקראי
- רוח לא יכולה להישאר במקום, אלא אם כן אין לה משבצת שאליה היא יכולה לעבור.
- רוח שדרכה על הרעל – נעלמת מהמשחק, כמו כן גם הרעל נעלם מהמשבצת הזאת. למחברי תרגיל הבית אין הסבר סביר איך רעל יכול להשפיע על רוח רפאים, אבל אלה הכללים.
- רוח לא יכולה לעבור למשבצת שנמצאת בה רוח אחרת או קיר

פעולת איפוס היא פעולה מיוחדת שמחזירה את המפה למצב הראשוני שלה. כמות הצעדים שנשארו והנקודות שנאספו עד עכשיו לא מתאפסים.

חייבים להפעיל את האיפוס במצב שהוגדר בתרגיל 1 ככישלון, שיכול להגרם על ידי כניסה של השחקן למשבצת בה יש רוח רפאים, כניסה של רוח רפאים אל המשבצת של השחקן, או כניסה של השחקן לתא עם רעל.

ניתן (אבל לא חובה) להפעיל את פעולת האיפוס גם בכל מצב אחר. הלוח מתאפס גם ללא הפעלת פעולת איפוס בסיום השלב (ז"א מיד אחרי אכילה של הגלולה האחרונה), וניתן לצבור עוד נקודות אחרי זה.

ניקוד

הסוכן מתחיל עם 0 נקודות, ומטרתו היא לצבור כמה שיותר נקודות. האירועים הבאים משפיעים על כמות הנקודות:

- אכילת גלולה ע"י שחקן: מוכתבת ע"י ההתפלגות של סוג הגלולה
- פעולת איפוס: -5
- סיום השלב (אכילה של כל הגלולות על המפה): +20

לכל סוג גלולה התפלגות משלה (לא ידוע לסוכן שלכם מראש), וכל פעם שהסוכן את הגלולה, נדגמת כמות הנקודות מהתפלגות המתאימה לסוג הגלולה. כמות הנקודות יכולה להיות לא שלמה, ואפילו שלילית.

בגלל החוסר ידע על הניקוד שמקבלים על כל גלולה, כדאי להפעיל שיטות של למידה ע"מ לשערך את הערך הזה.

הערה: להבדיל מתרגיל בית 1, אם הסוכן נהרג באותו התור שהוא אכל את הגלולה האחרונה (ז"א הרוח השיגה את הסוכן אחרי שהוא אכל את הגלולה האחרונה), הוא עדיין מקבל 20 נקודות, והלוח מתאפס.

הסבר קצר על הקוד המצורף

הקוד מורכב מקובץ פייתון ex2.py – קובץ שבו כל העבודה שלכם נעשית, והקובץ היחיד שעליכם לשנות. אמור להכיל את ה-class של PacmanController כפי שמתואר בסעיף "דרישה". הקובץ כולל את חתימות של הפונקציות שעליכם לממש.

כמו כן, מצורפים קבצי עזר mdp.py ו- mdp.py.utils. כולל את המימוש של אלגוריתמים לפתרון mdp ידועים כמו Value iteration ו- Policy iteration, וזאת ע"מ לאפשר לכם להשתמש בפתרונות מתרגיל 2 ללא בעיות תאימות לאחר.

שימו לב: אתם מקבלים כקלט לסוכן את המפה וכמות הצעדים, אבל לא את ההתפלגויות של ניקוד של הגלולות.

אסור לפנות לקוד שלנו ע"מ לברר את ערך הפרמטרים הלא יודעים, וניתן לגלות אותם רק בעקיפין

הבדיקה

הבדיקה תתבצע באופן הבא:

קוד הבדיקה ייצור מחלקה מסוג PacmanController, יקרא לפונק' choose_next_action ע"מ לבחור את הפעולה הבאה עד הגעה לאופק תכנון.

הבדיקה תתבצע ע"י הרצה של הקוד שלכם על מגוון קלטים בגודל שונה. נריץ את הבדיקה כמות פעמים גדולה על כל קלט ע"מ לנטרל את השפעת האקראיות.

איכות הקוד נמדדת על כמות הנקודות הממוצעת לכל קלט וקלט על פני ההרצות (בתנאי של עמידה בזמנים)

הגשה, בדיקת התרגיל וציונים

- הגשה בזוגות או יחידים בלבד
- מספיק שאחד מבני הזוג יגיש. אין צורך ששני בני הזוג יגישו את התרגיל
- את הת.ז. של המגישים יש לרשום במשתנה ids בקובץ ex3.py (בתור strings בתוך רשימה קיימת. אם המגיש יחיד, אז רשימה צריכה להכיל string בודד)
- יש להגיש רק את הקובץ ex3.py. אין להגיש קבצי עזר המצורפים לתרגיל. אין להגיש קובץ בפורמט אחר (לדוגמא zip או rar)
- חלק מהציון ינתן על עמידה בדרישות הבסיס, וחלק ינתן על בסיס תחרות בין המגישים
- התחרות תתבצע על ביצועים של הפתרון על מגוון קלטים. על הסוכן להביא את המערכת לצבירת כמות נקודות גדולה ככל הניתן ללא חריגה מתנאי הזמנים