**236501 - Introduction to Artificial Intelligence**

**AI HW2 spring 2022**

|  |  |  |
| --- | --- | --- |
| תום סמולין | 313552739 | tomsmolin@campus.technion.ac.il |
| נעם וולף | 326881240 | noamwolf@campus.technion.ac.il |

**חלק א' – סוכן חמדן משופר**

1. נגדיר היוריסטיקה חדשה שתכלול לפחות 3 פרמטרים נוספים מלבד פרמטר הכסף.

באופן כללי, רעיון היוריסטיקה הוא לעשות כמה שיותר נסיעות קצרות ומשתלמות (מונית תל אביבית ולא מונית ספיישל בינעירונית).

ראשית נגדיר מספר ערכים נוספים (כדי לא להעמיס על ההגדרה עצמה בהמשך).

* **עבור נוסע בסריג שמחכה למונית נגדיר את ערך התשלום עבור הסעתו במונית:**
* **עבור כל נוסע שמחכה לאיסוף במצב נתון (ובתנאי שהוא לא מחכה ביעד – ), הערך הוא:**
* **נגדיר חישוב שמוגדר רק עבור מצבים בהם ישנו נוסע על המונית.**

**החישוב יוגדר כך:**

* **נגדיר לכל חישוב :**

שמחזיר אם קיים נוסע שמחכה למונית כך ש .

אחרת, מחזיר .

כעת נעבור להגדרה של היוריסטיקה עצמה:

נשים לב שהפרמטרים הנוספים בהם השתמשנו (3 במספר):

* – מרחק מנהטן בין המונית הריקה לנוסע מסוים שממתין למונית (תחת התניה נוספת שהיעד שלו הוא לא מיקומו הנוכחי)
* – מרחק מנהטן בין מיקומו ההתחלתי של הנוסע ליעדו הסופי (מחושב רק עבור נוסעים שמקיימים (
* - מרחק מנהטן בין מיקום המונית לבין מיקום יעד הנסיעה (של הנוסע במונית).

1. כיוון שהיוריסטיקה מחלקת את המצבים למקרים, נתייחס למוטיבציה לפי מקרים:

* מקרה א' – המונית ריקה וקיים נוסע שעבור הסעתו ניתן לקבל תשלום חיובי ממש.

במקרה זה היוריסטיקה היא כאמור:

המוטיבציה היא להגיע ולהסיע את הנוסע שעבורו מתקבל המסלול הכולל הקצר ביותר (רעיון – לעשות כמה שיותר נסיעות קצרות ולמקסם את מספר הנוסעים) וגם לקבל כמה שיותר כסף.

כיוון שבחרנו לממש סוכן שמעוניין למקסם את היוריסטיקות שלו – המקדם של חיובי והמקדם של "אורך המסלול: מונית->נוסע->יעד" הוא שלילי (רוצים ערך קטן ביותר).

בפרט, פרמטר הכסף עוזר לנו לקבל החלטה להוריד את הנוסע ביעד.

שכן במצב שבו אנחנו עם נוסע ובאחד המצבים העוקבים אנחנו יכולים להוריד אותו – נרצה שהתשלום הנ"ל יילקח בחשבון בהשוואה ליתר המצבים העוקבים (ביתר המצבים הנוסע עדיין יהיה על המונית). לכן גם הכפלנו ערך זה ב12, כדי לתת לו משקל גדול יותר ביחס ליתר הפרמטרים.

* מקרה ב' – המונית ריקה, אבל לא קיים נוסע שעבור הסעתו ניתן לקבל תשלום חיובי ממש.

במקרה זה היוריסטיקה מחזירה 0. שכן, או שיש שני נוסעים שכבר נמצאים ביעד (ולכן אין טעם להגיע אליהם), או שיש נוסע זמין אחד שנמצא ביעד (ולכן אין טעם להגיע אליו) ונוסע נוסף במונית השנייה (לא נגיש עבורנו).

במקרה זה הפעולות החוקיות הן תנועות בכיוון מסוים או איסוף של נוסע (לא רווחי). כיוון שאנחנו רוצים להישאר זמינים לנוסעים חדשים שעבורם נקבל תשלום, נבצע תנועה אקראית באיזשהו כיוון (כיוון שלפי הפיאצה אנחנו לא יכולים להישאר במקום). בפועל הבחירה בתנועה (ולא באיסוף אם אנחנו נמצאים "על" נוסע) היא כחלק מאקט של "שבירת שוויון". שבירת השוויון באה לידי ביטוי במימוש – תנועה בכיוון מסוים קודמת לאיסוף נוסע (בסדר הפעולות שהוגדר בסביבה), ולכן בקריאה ל על כל הבנים שבמקרה זה יהיו עם אותו ערך – נבחר את הפעולה הראשונה, שהוא תנועה בכיוון צפון (אם אפשרית, אחרת ננסה את הכיוון הבא).

* מקרה ג' – יש נוסע על המונית עבורו אפשר לקבל תשלום.

במקרה זה היוריסטיקה היא כאמור:

כדי לא לסבך את היוריסטיקה יותר מדי עם חישובים מורכבים, גם למצב הזה אנחנו מתייחסים בצורה חמדנית, ובוחרים קודם כל להוריד את הנוסע.

כלומר, ניתן עדיפות למצבים שמקרבים אותנו ליעד באמצעות הפרמטר .

ככל שנהיה יותר קרובים, הערך שהיוריסטיקה מחשבת יהיה יותר גדול. כיוון שאנחנו מנסים למקסם את הערך בין הבנים (המצבים העוקבים), נבחר במצב הקרוב ביותר להורדה.

בנוסף, אנחנו עדיין נותנים משקל לכסף. כלומר, גם אם נעבור בתחנת דלק – נעדיף להמשיך אל היעד ולהוריד את הנוסע כמה שיותר מהר (כי הסוכן אוסף רק נוסעים "רווחיים", והגישה החמדנית שלנו במקרה הזה, רוצה לממש את הרווח קודם כל).

נשים לב שבחרנו להוסיף להיוריסטיקות במקרים א', וג' כדי לשמור על היוריסטיקה אי שלילית.

נשים לב שמרחק המנהטן המקסימלי בגריד הוא (מרחק מנהטן בין שתי פינות נגדיות).

לכן:

ולכן:

לכל :

לדעתנו סוכן חמדן המבוסס על היוריסטיקה שלנו ינצח את הסוכן החמדן הנתון (לפחות במרבית המקרים) כיוון שניתן לומר שהוא "יודע" יותר על העולם. נסביר:

הסוכן החמדן הנתון, מחשב לכל מצב את ההפרש בין מאזני הכסף של המונית שהוא מייצג למונית של היריב. כעת, נשים לב כי בזמן שמגיע תורו של הסוכן הנתון לשחק סכום הכסף של המונית היריבה זהה עבור כל המצבים העוקבים (למצב הנוכחי של הסוכן). המונית היריבה "סטטית" בשלב הזה.

לכן, הדרך היחידה שלו לייצר איזשהו אי-שוויון בין הבנים היא במקרה של הורדת נוסע (או תדלוק. נניח כרגע שהתחלנו לשחק ואין כסף לתדלק).

אבל, כדי להוריד נוסע צריך קודם כל לאסוף אותו. כיוון שאנחנו לא מרוויחים כסף מאיסוף הנוסע, הסיכוי היחיד שהסוכן יבחר לאסוף את הנוסע היא אם תהיה בחירה אקראית שכזו כשהמונית תעמוד "על" הנוסע (כי שוב, הערכים של כל הבנים יהיו זהים. למעט אולי מצב עוקב עבורו מתדלקים).

גם אחרי איסוף הנוסע, זה עניין של מזל בכלל להגיע ליעד של הנוסע (מסיבות דומות – הפרש הכספים לא עוזר לנו להבין איפה היעד, אלא אם כן אנחנו עומדים עליו).

כלומר, אפשר להגיד שהסוכן הנ"ל הוא עם "ידע" מאוד מוגבל על העולם בכל רגע נתון. אמנם המטרה הסופית שלו, היא אכן לקבל המקסימום של הערך אותו הוא בא לחשב בכל מצב - אך כאמור, חישוב ערך זה כערך היוריסטי, לא משרת אותו היטב כדי לנצח.

\*כדי לחדד את עניין התדלוק – גם במקרה שהוא כן הצליח לאסוף ולהוריד נוסע ביעד, הוא עדיין יעדיף לא לתדלק (ייתן ערך היוריסטי יותר קטן, בעוד הסוכן הנ"ל מעוניין למקסם) ולכן לא יבחר בפעולה זו.

הסוכן שלנו לעומת זאת, מגיע עם "ידע" על העולם. כלומר, הוא יודע איפה הנוסעים ממוקמים והוא "יודע" איפה להוריד אותם (מהסיבות שציינו).

ולכן הוא "מחפש" (או בוחר לבצע) באופן אקטיבי את הנסיעות הקצרות ביותר שישתלמו לו עד שיגמר הדלק (אפשר היה לעשות היוריסטיקה יותר מורכבת ויותר מתוחכמת, אבל לא נדרשנו לכך).

להלן התוצאות הסופיות בסדר שבו התבקשנו להציג אותן (ועם תיוג של הסוכנים לטובת הבהירות):

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

**חלק ב' – סוכן Minimax**

1. נניח שלסוכן היריב נגמר הדלק והוא מחזיק כסף כך ש .

לסוכן של החבר (מינימקס) יש והגיע תורו לשחק. נניח למשל שהוא עמד על תחנת דלק. כעת, אלגוריתם המינימקס מכוון אותו לבחור בצעד התדלוק מתוך "הנחה" שניתן למקסם כך את הרווח הסופי של המונית שלו (למשל, אם חסר לו דלק כדי להוריד את הנוסע שנמצא עליו עכשיו, והרווח הצפוי משתלם). כלומר, האלגוריתם מבין שהוא יכול להבטיח תועלת טובה יותר אם יתבצע הצעד של התדלוק (תחת ההנחה שאפשר גם להמשיך לשחק). לכן נקבל שהסוכן של החבר יתדלק ולכן סכום הכסף שלו יהיה (מתדלקים עם כל הכסף).

עכשיו נניח, שזה היה הצעד האחרון להיום. כלומר, בסופו של התור הנ"ל נגמר המשחק.

קיבלנו תרחיש שבו הסוכן של החבר לא מנצח כיוון ש כשהוא היה יכול לבחור צעד אחר (שאינו תדלוק) שהיה מבטיח לו את הניצחון.

שינוי פשוט שניתן להציע כדי למנוע התנהגות כזאת – להתייחס בצורה פרטנית למקרים בהם לסוכן שלנו יש יותר כסף מלסוכן השני ולסוכן השני נגמר הדלק. במקרים כאלה נגדיר לסוכן של החבר לשרוף דלק (ואולי גם לנסות להסיע נוסע נוסף). התועלת שנבטיח תהיה לפחות ובאופן זה נבטיח ניצחון, שזו המטרה הסופית.

\*במקרה שב "לנצח בתור" הכוונה היא לסיים את התור עם יותר כסף, ללא תלות במצב הסופי של המשחק אז בוודאי שיכולים להיות מקרים כאלה במינימקס - שנבצע פעולה מסוימת כדי להבטיח תועלת גבוהה יותר בהמשך. במקרה כזה זה לא בהכרח באג בקוד.

\*\* אפשר גם להתייחס למקרה שבו אין הגבלות על מס' הצעדים ליום. מקרה זה דומה למקרה שתיארנו בפירוט. במקום לגמור את הדלק ולסיים את המשחק – הסוכן יכול לבחור למלא את המיכל מתוך איזושהי הבטחה למקסם את התועלת (ולהמשיך לשחק, אולי עד אינסוף). כלומר שוב נבחר צעד מסוים על פני צעד שיכול היה להבטיח ניצחון.

1. שני יתרונות עבור :
   1. האלגוריתם מספק מסלול כלשהו לשיפור התועלת (לפחות עבור מס' צעדים ראשוני של המשחק). כלומר, הוא לא מחכה לסיום מלא של החיפוש והוא **מסוגל לספק "תוצאות ביניים" (עבור אלגוריתמי למשל**).
   2. **האלגוריתם (לפחות בסבירות גבוהה) לא "נתקע". כלומר, יש לנו "נקודות יציאה". הסוכן יותר אינטרקטיבי.** שוב, זה מתקשר לגודל העץ שנפרש ולכן במקרים כאלה קיימת המוטיבציה להשתמש באלגוריתם מלכתחילה. למשל על סמך ממצאים באינטרנט, חישוב מלא של משחק שחמט יכול להיערך כשעתיים (=מוטיבציה לעצור באמצע).

**\*\* כלומר, אם לחדד –** יתרון אחד אומר שהאלגוריתם מספק איזשהם פתרונות כל כמה זמן ולכן הוא יכול לשמש אלגוריתמי (מועיל כשאנחנו לחוצים בזמן עבור ריצה מסוימת).

**יתרון נוסף –** אנחנו לא מחויבים לחישובים "ארוכים". ברמת המימוש, אם המשאבים מוגבלים, ניתן לנהל את הריצות בצורה גמישה יותר (כלומר המשאב הוא לא בהכרח זמן. הוא יכול להיות גם ליבת חישוב. למשל (להלן הוראות מדומות לרובוט/סוכן/ישות שמריצה סוכנים) – "חשב איטרציה עד עומק מסוים. עצור. פידבק לא מספיק טוב? חכה רגע. הרץ משהו אחר שלא בהכרח קשור באופן ישיר להחלטה במשחק (אולי זה רובוט שעושה דברים במקביל). בצע איטרציה נוספת עמוקה יותר".

שני יתרונות עבור :

1. האלגוריתם לכשעצמו הוא **שיפור של אלג' ה** . לכן תחת תנאים מסוימים הוא יודע להבטיח תועלת כלשהי עבור כל צומת ובפרט, הוא יודע להגיע לעומקים גדולים יותר ביחס לאלגוריתם המקורי (לכן בפועל ניתן להחליף איתו שימוש במינימקס בלי להשפיע על התוצאה. למעשה **ניתן להשאיל את היתרונות שמנינו עבור ID-מינימקס, רק עם מימוש יעיל יותר**).
2. חוסך זמן ריצה מיותר **על מסלולים שלא ייבחרו** (בענפים שפיתוחם בהכרח לא ישנה את ערך\מסלול\אסטרטגיית המינימקס) ולכן מרכז את הריצה שלו על ענפים שכן יכולים להשפיע על בחירות המינימקס (**ניצול "סלקטיבי" של הזמן**).

ישנם מספר שינויים שניתן לבצע כדי לשפר את השיטות הנ"ל באופן כללי או בהקשר של המשחק שלנו. את השיטה הראשונה, ניתן לשפר באמצעות גיזום כמו שציינו לעיל. את השיטה השנייה ניתן לשפר באמצעות מיון של הבנים כפי שראינו בתרגול. ספציפית למשחק שלנו ניתן לשלב היבטים נוספים, מהירים לבדיקה שיכולים להבטיח את המטרה. למשל כמו בסעיף א', במקרים בהם לסוכן השני נגמר הדלק ולנו יש יותר כסף אז נדע לעצור את זרימת האלגוריתם כפי שהוא מוגדר באופן כללי, כדי לעבור לאסטרטגיה שונה דומה למה שתיארנו (שבה לא בוחרים בתדלוק). אם יודעים מראש את מספר הצעדים המותר ליום, אזי ניתן לקבל עצי מינימקס חסומים בעומקם ובאופן זה להימנע מקבלת החלטות שיפגעו בנו במשחק עצמו (כשהאלגוריתם לוקח צעד שמטרתו לשפר את התועלת בהמשך ולא בטווח הקרוב).

**חלק ג' – סוכן Alpha-Beta**

1. מימוש.
2. השינויים שהיינו צריכים לעשות הם כלהלן:

היינו צריכים להכניס לתוך האלגוריתם (ולפיכך למימוש) גם את התורות של יתר המוניות

. כלומר, אם נדמיין את "עץ המשחק" שמתפתח, אז "תור" המונית הראשונה יופיע אחת ל שלבים בעץ (ולא 1 ל2). כמו כן, במימוש של ה במקום לשמור דגל בוליאני שמתעד את הסוכן שזהו תורו (דהיינו הסוכן שמחשב את הצעד הבא שלו – או הסוכן שמולו הוא מתמודד) היינו ככל הנראה מעבירים ברקורסיה את מספרו של הסוכן הבא ש"מסמלץ" (וזוכרים באופן כלשהו איזה סוכן מחשב את הצעד, כלומר הסוכן שנמצא בשורש וקרא לאלג').

בנוסף, אם נרחיב על השינויים ומה שעומד מאחוריהם, אז בכל צומת בעץ נצטרך להחזיק בעל ערכים, המתאר את היוריסטיקות (אם לא מצב סופי) או מצב הכסף (מצב סופי למשל) עבור כל מונית.

נשים לב למספר דברים נוספים – **כשמדובר ביותר משני שחקנים, שיטת המינימקס כפי שהיא פחות מתאימה לפי הספרות** (שכן בשיטה זו מדובר בהתמודדות של שני שחקנים, ובצורה פשטנית ניתן לומר שאם ננסה למזער את "תוצאת" היריב נשפר את הסיכוי שלנו לנצח אבל מוטיבציה זו לא בהכרח באה לידי ביטוי כשיש הרבה שחקנים) **ולכן היינו בוחרים לממש גרסה מעט שונה** של האלגוריתם (באותה הרוח) באופן כזה שיתאים למספר רב של סוכנים. **סביר להניח שבכל שלב היינו מחפשים פשוט למקסם את התוצאה של השחקן "שמסמלץ"** (כלומר אם אנחנו בצומת שמהווה בחירה של שחקן אז שחקן זה יבחר בצומת הבן עם רכיב ה המקסימלי בתוך ה ). לכן הגיזומים לא יגזרו מערכי אלפא ובטא, אלא מסדרה של ערכים שמתארים את המצב המקסימלי ששחקן מסוים יכול להבטיח.

\*נדגיש שאנחנו העלינו את הרעיון הכללי, אך בספרות (ממה שחקרנו) כתוב שאם אין חסמים על סכום הרכיבים בכל , אז גיזום של ענפים שלמים פחות סביר (אם אנחנו מבינים נכון, זה כי אין שום הבטחה שלא יצוץ מצב מאוד שווה עבור שחקן מסוים בהמשך המשחק) ולכן הגישה הזאת לא התפתחה.

אבל ברגע שיש חסם על סכום התוצאות של כל שחקן, ניתן להניח חסמים תחתונים וחסמים עליונים לתוצאות של השחקנים השונים ולבצע את הגיזומים ע"פ ערכים אלו (שכן אם שחקן1 מבטיח לפחות אז שחקן2 יכול לקבל לכל היותר וע"פ זה ניתן לקבל החלטות "מבטיחות" יותר).

\*\*\*\*אם לא נלך על הגישה כפי שהיא מופיעה בספרות, אזי נגדיר עבור הסוכנים היריבים שמטרתם היא למזער את התוצאות של השחקן בשורש. במקרה כזה כן נתחזק את משתני אלפא ובטא (שכן המוטיבציה של כל השחקנים היריבים היא משותפת).

בפועל – אם נשאיר את הגבלת הזמן לכל חישוב צעד של סוכן ונדאג לבצע בזמן, זמן הריצה לא אמור להשתנות יותר מדי.

אבל, אותו החישוב יתבסס על ניתוח של פחות צעדים קדימה במשחק. שכן ככל שנוסיף שחקנים ייקח לנו יותר זמן לחשב את אותה כמות של סבבי תורות.

יתרה מזאת, מהסיבות שציינו לעיל – אם אין חסם על סכום הרכיבים ב אז היכולת לבצע גיזומים מידיים (בספרות ) היא הרבה פחות טריוויאלית (הרי כל שחקן יחפש למקסם עבור עצמו. ואם אין חסם, אין הבטחה שהוא לא יכול למקסם "בהפתעה" בעתיד). לכן גם סביר שנעשה פחות גיזומים.

פחות גיזומים -> זמן ריצה ארוך יותר תחת חישוב של אותה כמות צעדים לסוכן (בפועל אם מגדירים כמות דלק התחלתית קבועה לכל משחק אז כנראה שניתן להגדיר איזשהו חסם גס על כל רכיב ב ולכן גם על סכום הרכיבים. אבל בפרט הגיזום עדיין לא יהיה מיידי כמו זה המתקבל במינימקס ולכן נניח שזמן הריצה יהיה ארוך יותר, עבור אותו מספר סבבים).

1. *בשני החלקים אנחנו משתמשים ב ועבור צמתים ששייכים לקב' המצבים הסופיים מחזירים את הפרש הכספים ועבור מקרים בהם לא הגענו למצב סופי מחזירים את אותה היוריסטיקה שפיתחנו בחלק הראשון.*

*כעת נשים לב, שהאלג' של חלק ג' הוא מעין שיפור של האלג' של חלק ב' (בהינתן שיש לנו היוריסטיקה סבירה, הגיזום מאפשר סריקה יותר יעילה של ה"עץ") שכן הוא מוותר על מסלולים מיותרים בעץ המשחק.* ***לכן אכן, ברוב הגדול של ההרצות אם נחשב את ממוצע זמני הריצה של כל סוכן (עבור אותם מצבים התחלתיים), נקבל שממוצע זמני הריצה של סוכן ה הוא קצר יותר*** *תחת תנאי המשחק הנתונים (אלפא-בטא ימצה מהלכים מהר יותר אם לשניהם יש מספיק זמן להעמיק עד הסוף למשל).*

*יחד עם זאת נציין כי אנחנו קוטעים את ריצת ה אם עברנו אחוז מסוים מהזמן שהוקצב לנו לתור (כדי לעמוד בהגבלת הזמן במימוש). רוצים לומר – הסוכנים רצים פחות או יותר פרקי זמן דומים מבחינה אבוסולוטית. לכן לעיתים, במקרים שבהם לא הגענו למצבים סופיים בהעמקה שלנו – את הזמן שסוכן האלפא-בטא היה אמור לחסוך, הוא מנצל להעמקה שלא מתאפשרת בזמן הקיים עבור סוכן המיני-מקס. לכן יש צעדים שבהם הזמנים כן יכולים להיות שווים ואף ארוכים יותר עבור סוכן האלפא-בטא.*

*מבחינת קבלת ההחלטות, תחת היוריסטיקה שפתחנו (יוריסטיקה שמגדירה התנהגות של מונית לנסיעות קצרות, מה שכינינו מונית "עירונית"/"תל אביבית")* ***הסוכנים מקבלים החלטות זהות במרבית המקרים.*** *לעיתים יש הבדלים כשסוכן האלפא-בטא מספיק להעמיק יותר או דווקא כשהוא בוחר לגזום בנים עם ערכים שווים (לאלפא או לבטא, התנאי הוא קטן/שווה או גדול/שווה ולוותר על מסלולים לאו דווקא מיותרים) ואז הצעדים הנבחרים (אחת לכמה זמן) יכולים להבדל אחד מהשני (אם כי שוב, באופן "גורף" ההחלטות דומות במרבית המקרים).*