



PeRL - Personalized Reinforcement Learning

התאמת יריב ממוחשב לשחקן במשחקי מחשב

חדשנות הפתרון

חדשנות הפתרון מתבטאת בכמה היבטים שונים:

- הפתרון כולל אפיון שחקן בעזרת בינה מלאכותית,** דבר שלא נעשה בעבר ועוזר לשפר באופן משמעותי את ההתאמת היריב ואת חווית המשחק.
- פתרון זה **לא דורש לדעת למי מבין השחקנים יש יתרון** וכמה יתרון בכל רגע במשחק, זאת בניגוד לפתרונות קודמים. **יתרון זה מאפשר ליישם את הפתרון במשחקי מחשב מגוונים ורבים יותר.**
- הפתרון שומר על הרמה של היריב הממוחשב קבועה לאורך המשחק,** דבר המשפר את הדמיון שלו לשחקן אנושי.

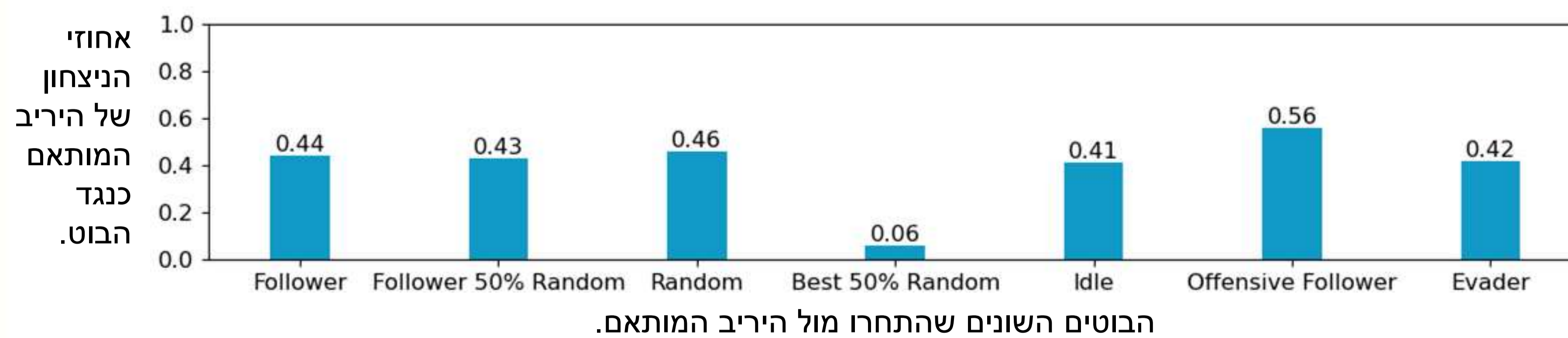
מבט לעתיד

- הפרויקט מציג דרך חדשנית לאפיון צורות משחק של שחקנים שונים באופן יותר כמותי ואובייקטיבי. **האפיון יכול לשמש ככלי עוצמתי להבנת צורות משחק שונות.**
- הפתרון מאפשר **שיפור משמעותי של חווית המשחק אל מול המחשב ביחס לשיטות קיימות** ועשוי לאפשר בעתיד פרסונליזציה משופרת במשחקי מחשב.
- השיטה שפותחה ניתנת ליישום במגוון משחקי מחשב.**

תוצאות ומסקנות

- מהלך הניסוי:** על מנת לבדוק את אפקטיביות הפתרון שלושת שלבי השיטה מומשו במשחק "Pong" ונבדקה הצלחת ההתאמה למספר בוטים שונים (ראו גרף ב').
- תוצאות ככלל:** הגרף מתאר את אחוזי הניצחון של היריב המותאם מול בוטים שונים. ניתן לראות כי בכל המקרים למעט אחד אחוזי הניצחון נעו בין 40% ל-60%, דבר המעיד על הצלחה בהתאמה בשל הקרבה לשוויון.
- התוצאה החריגה:** הבוט היחיד מולו ההתאמה נכשלה הוא בוט שמחצית ממהלכיו אקראיים. לשם השוואה, בוט זה ללא החלק האקראי מנצח 99.4% מהמשחקים מול הבינה המלאכותית ברמה הגבוהה ביותר. דבר זה מעיד שרמת הבוט גבוהה מאוד ביחס לבינה המלאכותית ולכן היא אינה מסוגלת להעריך את מהלכיו בצורה נכונה. מכך ניתן להסיק **שההתאמה תצליח כל עוד הבינה המלאכותית ברמה מספיק גבוהה כדי להעריך נכון את המהלכים שהשחקן מבצע,** דבר המעיד שאימונים נוספים ישפרו את התוצאות.

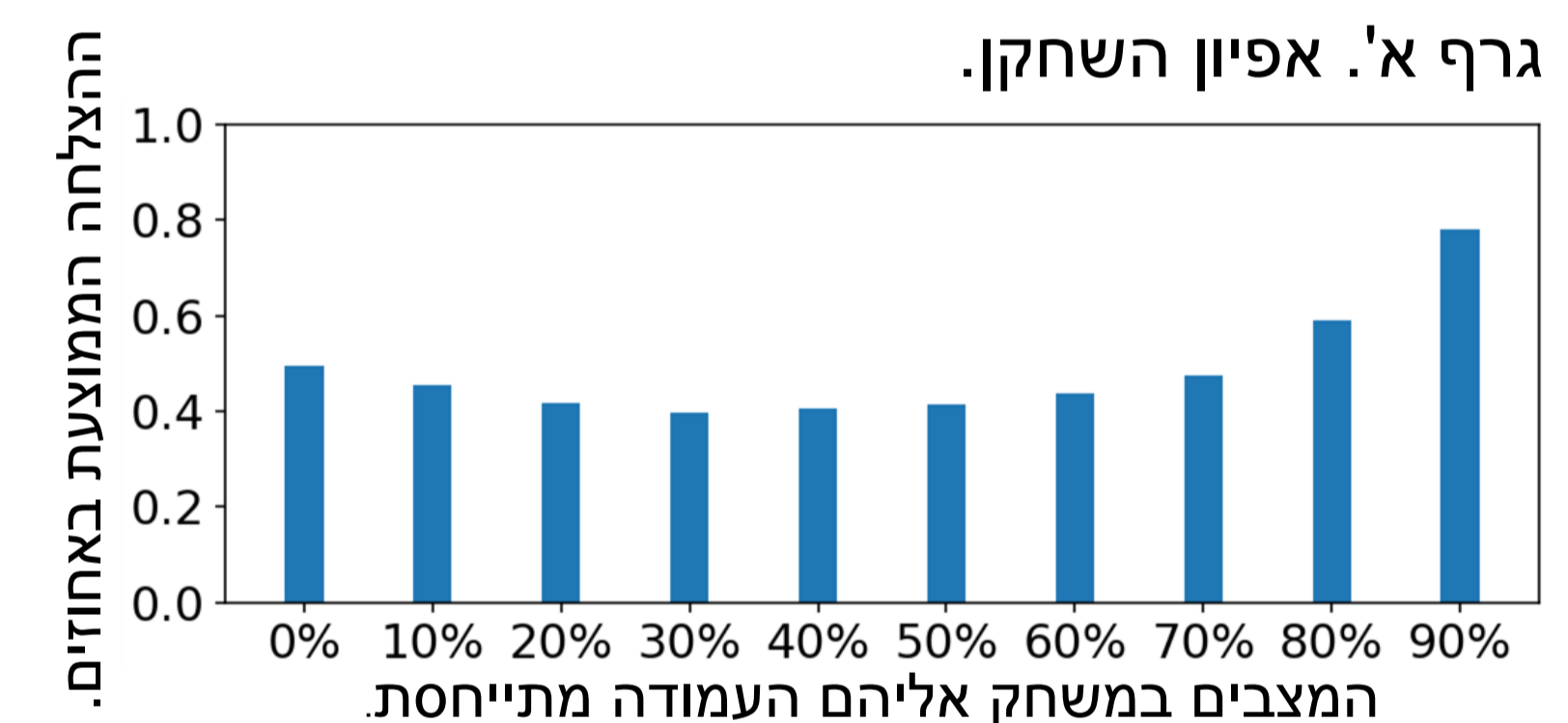
גרף ב' - תוצאות ההתאמה.



הפתרון - PeRL

השיטה מורכבת מ-3 שלבים עיקריים הניתנים ליישום במגוון רחב של משחקי מחשב:

- שלב האימון:** מאמנים בינה מלאכותית עד למצב שהיא מסוגלת להעריך בצורה טובה את איכותו של כל מהלך.
- שלב האפיון:** בוחנים משחקים שהשחקן שחקן בעבר על מנת לאפיון את יכולותיו ואת רמת המשחק שלו. את האפיון של השחקנים ניתן להציג באופן גרפי (גרף א').



גרף א'. אפיון השחקן. ההצלחה הממוצעת באחוזים. במהלך המשחק השחקן נתקל במצבים רבים. לעיתים יש לו יכולת השפעה רבה על התוצאה ולעיתים יכולת השפעה מועטה. כל עמודה מתייחסת ל-10% מהמצבים לפי יכולת ההשפעה של השחקן בהם בסדר עולה. לדוגמה, העמודה הראשונה מתייחסת ל-10% המצבים בהם לשחקן הייתה את יכולת ההשפעה המועטה ביותר.

- שלב ההתאמה:** התייחסות למצב המשחק והאפיון שהתקבל עבור השחקן האנושי. עבור כל מצב במשחק:
 - נבדוק לאיזו עמודה הוא משייך לפי כמות היכולת של היריב המותאם להשפיע על המשחק באותו הרגע.
 - נחפש את המהלך ששימוש הפוטנציאל שלו הוא הקרוב ביותר לגובה העמודה לה זכה השחקן בתהליך האפיון. כך אנחנו בעצם מתבססים על מהלכים שבוצעו על ידי השחקן בביצוע המהלך הנוכחי של היריב המותאם.

מבוא - משחקי מחשב

התאמת משחקי מחשב באופן יותר אישי לכל שחקן יכולה לשפר את חווית המשחק באופן דרסטי, לכן זוהי מטרה נחשקת בעיניהם של יוצרי משחקים רבים.

במשחקים מסוג שחקן נגד שחקן ישנם שחקנים בעלי יכולות משחק שונות. לכן, כאשר משחקים נגד המחשב לכל שחקן יתאים יריב ברמה שונה בהתאם ליכולותיו של השחקן. בנוסף, המחשב אינו משחק כמו שחקן אנושי, דבר הפוגע בחווית המשחק.

מטרת הפרויקט היא ליצור לכל שחקן, יריב ממוחשב מותאם הדומה לשחקן אנושי.

למרות שבעיה זו נשמעת פשוטה, פתרונות קודמים לא השיגו תוצאות מספקות מסיבות שונות, לדוגמה:

- שימוש בהנחות יסוד שלא מתאימות לחלק נרחב ממשחקי המחשב.
- חלקם יצרו יריב שמשנה את רמתו תוך כדי המשחק כדי להוביל לשוויון - דבר שמצד אחד מנע מהיריב להיות יותר מדי טוב ובכך מנע תסכול, אך מצד שני פגע בחווית המשחק.

שיטות וחומרים

על מנת ליצור פתרון חדש ויותר אפקטיבי השתמשנו בשיטות למידת מכונה מסוג:

Reinforcement Learning (RL).

התחום של RL הינו אחד משלושת תחומי הבסיס של למידת מכונה. הוא מתמקד במציאת הדרך הטובה ביותר להשיג מטרה מסוימת בסביבות משתנות. דוגמה לסביבה כזו היא משחק השחמט בו השחקן יכול לבצע מהלכים שונים ומטרתו היא להביס את היריב.

במקרה שלנו, במהלך האימונים הבינה המלאכותית לומדת לשחק במשחק ברמה גבוהה וגם להעריך את טיבו של כל מהלך אפשרי בכל רגע נתון במשחק.

במהלך האימונים הבינה המלאכותית מתאמנת נגד גרסאות קודמות של עצמה (Self Play) כך שאין צורך בהתערבות אנושית במהלך האימונים.