



אומיד דויד, משה קופל ונתן נתניהו

תצלום: אימלבינק / GettyImages

אבולוציה בשחור-לבן

פיתוח אוטומטי של רב-אמן ממוחשב

048

גליצ'א
2008

ממחקרי אוניברסיטת בריאילן



שחמט ללמוד ולשפר את יכולת המשחק שלה באופן אוטומטי. מחקרם, שיוצג השנה בכנס הבינלאומי הגדול ביותר לחישוביות גנטית ואבולוציונית (GECCO '08), מראה כיצד תוכנת שחמט המתחילה בלא ידע כלשהו יכולה לעבור "אבולוציה" בדקות אחדות לתוכנה ברמת רב-אמן. הם עושים זאת באמצעות מתודולוגיה של אלגוריתמים גנטיים (GA), המבוססת על עקרון הברירה הטבעית של דארווין. גישתם מאפשרת לתוכנה לבצע הנדסה-לאחור של ה"מוח" של ישות תבונית אחרת על-ידי בחינת ה"התנהגות" שלה. לפני פירוט גישת הלמידה שלהם (שהאבולוציה שימשה לה כמקור השראה) ויישומה בשחמט, להלן דיון קצר באופן שבו תוכנות שחמט פועלות (או "חושבות").

כל תוכנות השחמט בנויות משני מרכיבים עיקריים: חיפוש והערכה. חיפוש הוא תהליך החקירה של כמה מהלכים קדימה, והערכה היא אומדן החוזק של מצב נבדק ודירוגו. במצב נתון, נשקלים כל המהלכים החוקיים של הצד שתורו לשחק, ולכל מהלך כזה נשקלים כל המהלכים החוקיים של היריב, וכן הלאה. תהליך זה נמשך עד ל"עומק" מסוים. אפשר לייצג את תהליך החיפוש כעץ חיפוש, ששורשו הוא המצב הנוכחי, וכל מהלך הוא ענף בעץ היוצר צומת חדש (ראו איור 1). בנקודה מסוימת יש לחדול מהעמקת החיפוש, שכן מספר המצבים גדל אקספוננציאלית עם ההעמקה. המצבים שאינם מפותחים הלאה נקראים "צומתי עלה".

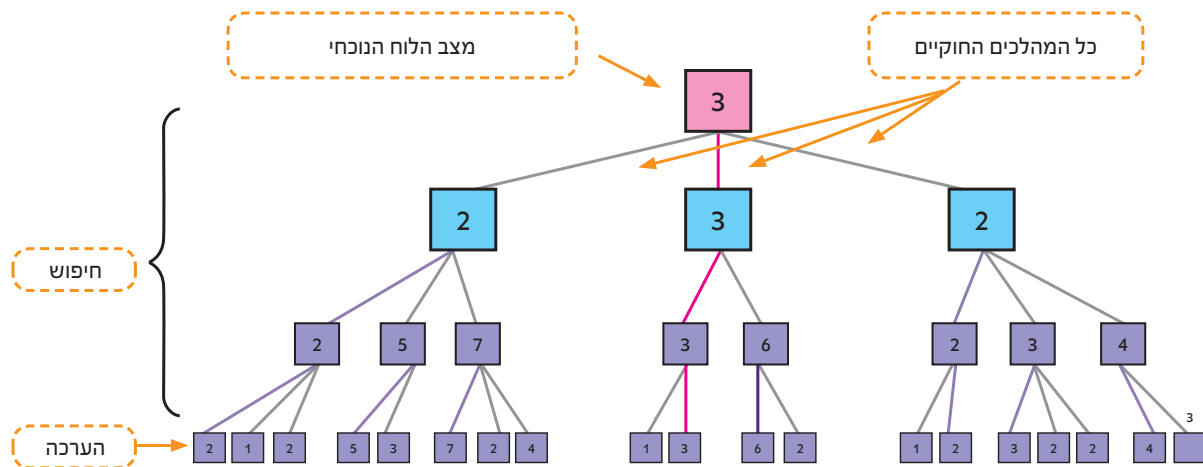


השעה היתה מעט אחרי 3 אחר הצהריים, ב-11 במאי 1997, כאשר אלוף העולם בשחמט, גארי קספרוב, נכנע לאחר 19 מהלכים, בתבוסה המהירה ביותר בכל הקריירה שלו. יריבו היה מכונה ששמה "כחול עמוק", מחשב-על היכול לנתח 200 מיליון מצבים בשנייה. אמצעי התקשורת הציגו את תמונתו של קספרוב המניד ראשו כלא-מאמין והכריזו שהנה, סוף-סוף, "בינה מלאכותית" גברה על בינה אנושית. במשך למעלה מחמישים שנה היה משחק שחמט-המחשב שדה הקרב העיקרי בין המכונה לאדם, ומכאן אחד הנושאים הנחקרים ביותר בתחום הבינה המלאכותית.

מה שרבים לא השכילו להבין בעקבות ניצחונו של "כחול עמוק" היה, שאין כמעט דבר בין אותו מחשב-על לבין בינה זוהי "בסך-הכל" מכונה אדירת כוח לחישוב מצבי שחמט, ואף שהפגינה, לכאורה, התנהגות תבונית (בהביסה את אלוף העולם בשחמט), תהליכי החשיבה שביסוד משחקה לא היו אלא חישובים פשוטים, גם אם מהירים מאוד. נעם חומסקי ניסח זאת מאוחר יותר: "[ניצחונו של 'כחול עמוק'] מעניין בערך כמו העובדה שבולדוזר יכול להרים משקל כבד יותר ממרים משקולות זה או אחר."

בעשור שלאחר התחרות בין "כחול עמוק" לקספרוב, הוסט העניין מיצירת תוכנות שחמט שהן מכוונות חישוב פשוטות ומהירות ליצירת תוכנות שחמט המסוגלות ללמוד לשפר את ביצועיהן באופן אוטומטי. נבחנו שיטות למידה רבות, שיושמו בהצלחה במגוון רחב של בעיות (ובכלל זה משחקים), אך שחמט נותר אגוז קשה לפיצוח עבור כולן. עד היום, אף אחת מהתוכנות המשתתפות במשחקי אליפות העולם השנתית בשחמט-מחשב אינה מפגינה חשיבה תבונית משמעותית במונחים של למידה (על אבי שחמט-המחשב, קלוד שנון, ראו: אמיר קליסמן, "האב הלא-נודע של עידן המידע", "גליליאו" 118).

אומיד דויד (David), משה קופל (Koppel) ונתן נתניהו (Netanyahu), מהמחלקה למדעי המחשב באוניברסיטת בר-אילן, מאמינים שגישתם החדשנית, שהם מכנים "מסתייעת-מנחה" (mentor-assisted), מאפשרת לתוכנת



איור 1 אנטומיה של תוכנת שחמט

של היריב), עמדת כלים (עד כמה עמדתם פעילה), מבנה הרגלים וכו'. לכל פרמטר כזה מוקצה ערך חיובי או שלילי. הניקוד הסופי של המצב מחושב כסכום הערכים של הצד שזהו תורו, פחות סכום הערכים של היריב. כיום, פונקציית ההערכה בכל תוכנות השחמט המובילות נקבעת באופן ידני. כלומר, המתכנתים משקיעים זמן רב (כמה שנים, בדרך-כלל) בכווןן ידני של כל עשרות (ואפילו מאות) ערכי הפרמטרים של פונקציית ההערכה. נוסו גישות רבות לאוטומטיזציה של תהליך זה, כלומר, ליצירת הליך שיאפשר לתוכנה ללמוד באופן אוטומטי את הערכים המיטביים של הפרמטרים. כל השיטות האלה הניבו ביצועים גרועים בהרבה מאלה של תוכנות המכוונות ידנית. אומיד דויד, משה קופל ונתן נתניהו הראו כיצד אפשר להשתמש באלגוריתמים גנטיים לצורך למידה של התוכנה. אלגוריתמים גנטיים, שהראשון לפתחם היה ג'ון הולנד (Holland) ב-1975, הם אלגוריתמים המפתחים על-ידי "אבולוציה" ערכי פרמטרים מיטביים לפתרון בעיות שונות. הם פועלים על-ידי מיפוי הבעיה לזו של קביעת רצף גדול של סיביות (ביטים), רצף המכונה "כרומוזום", ואבולוציה של

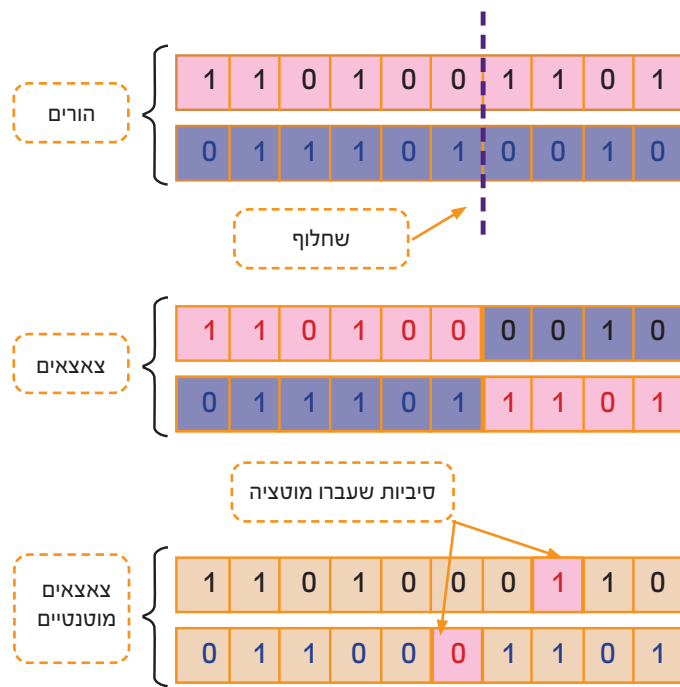
בכל צומת עלה מתבצעת הערכה, והצומת מקבל ערך מספרי המציין את חוזק העמדה. ציונים אלה מחושבים לאחור עד לשורש, על-ידי מתן הציון הטוב ביותר האפשרי לכל צומת (פנימי), כלומר, לכל עמדה. ההנחה היא, כמובן, שכל שחקן נוקט - בתורו - את המהלך הטוב ביותר האפשרי. בסופו של דבר, התוכנה בוחרת את מהלך השורש שלו הניקוד הגבוה ביותר ומבצעת אותו על הלוח. כפי שעולה מן התיאור, הליך החיפוש קשור למרכיב החישובי של התוכנה, ואילו ההערכה מבטאת את הידע וההבנה בשחמט. אלגוריתמים יעילים לחיפוש מצבים בעץ הם נפוצים ביותר (תוכלו למצוא השוואה בין כמה אלגוריתמי חיפוש בספר *Computers, Chess and Cognition*, המופיע ברשימה לקריאה נוספת שבסוף המאמר), והקושי העיקרי הוא ליצור פונקציית הערכה טובה. מכיוון שפונקציית ההערכה משקפת למעשה את הידע והבנת השחמט של התוכנה, היא המרכיב המכריע ביותר. פונקציית הערכה בסיסית צריכה לבדוק עשרות פרמטרים על הלוח, דוגמת חוסר איזון חומרי (למשל, לצד אחד יש פרש עודף), הגנת המלך (אם הוא תחת התקפה



ממחקרי אוניברסיטת בריאילן

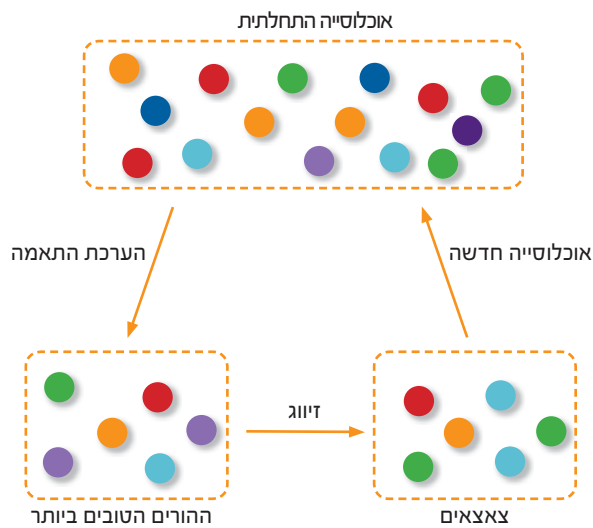


כאשר שני אורגניזמים נבחרים לזיווג, הם יוצרים את צאצאיהם על-ידי החלפת מידע גנטי. הדבר מתבצע באמצעות שחלוף גנטי בין הכרומוזומים שלהם, כלומר, נבחרת נקודה אקראית (אינדקס סיבית) בכרומוזום, כך שכל הסיביות שלפני נקודה זו נלקחות מכרומוזום-הורה אחד, וכל הנקודות שאחרי נקודה זו נלקחות מכרומוזום-ההורה האחר. לדוגמה, שחלוף בסיבית השלישית של הכרומוזום 11111111 עם הכרומוזום 00000000 ייצור את כרומוזום-הצאצא 11100000. לבסוף, הצאצא עובר מוטציה על-ידי "היפוך" כל סיבית (כלומר, הפיכתה מ-0 ל-1 או מ-1 ל-0) בהסתברות נמוכה מסוימת הנקבעת מראש. איור 2 מדגים אלגוריתמים גנטיים, ואיור 3 מדגים זיווג.



איור 3 זיווג באלגוריתמים גנטיים

"אוכלוסייה" גדולה של כרומוזומים במשך כמה דורות, על פי העיקרון של ברירה טבעית. בסופו של דבר, הצאצא הטוב ביותר הוא זה שאמור לתת את הפתרון המיטבי לבעיה. כדי להחיל אלגוריתם גנטי על בעיה, יש אפוא לייצג אותה תחילה ככרומוזום. לדוגמה, אם הבעיה כוללת חמישה פרמטרים שיש למצוא את ערכיהם המיטביים, ואם כל פרמטר מצריך עד 10 סיביות, הבעיה תיוצג על-ידי כרומוזום באורך 50 סיביות. הצעד הבא הוא יצירת אוכלוסייה של כרומוזומים מאותחלים אקראית (כלומר, הסיביות של כל כרומוזום נקבעות כ-0 או 1 באופן אקראי). בשלב זה, יש אוכלוסייה של "אורגניזמים" (כל כרומוזום מייצג אורגניזם אחד). אוכלוסייה זו עוברת אבולוציה במשך כמה דורות, באופן זה: בכל דור, מיוחס לכל אורגניזם ערך המציין כמה הוא "מתאים" (fit). עם סיום דור, כל אורגניזם מזווג באופן יחסי לערך ההתאמה שלו. במילים אחרות, האורגניזם המתאים ביותר יהיה בעל הסיכוי הגבוה ביותר לשרוד ולהעמיד צאצאים.



איור 2 הדגמה של אלגוריתמים גנטיים



אפשר להיווכח שאלגוריתמים גנטיים פועלים בדומה לעקרונות הדארוויניים של הברירה הטבעית. האורגניזם המתאים ביותר מזדווג יותר, והמתאים פחות נעלם עד מהרה. כאשר תהליך זה חוזר ונשנה במשך דורות רבים, מתקבלים אורגניזמים מותאמים היטב, בעלי ציון התאמה גבוה (כלומר, הכרומוזומים שלהם מגיעים לערכים מיטביים לבעיה הנדונה). אנליזה מפורטת של אלגוריתמים גנטיים אפשר למצוא בספרו של דויד גולדברג (ראו: לקריאה נוספת).

במבט ראשון, הכוונון האוטומטי של פונקציית ההערכה של תוכנת שחמט נראה כמשימת אופטימיזציה סטנדרטית, שאפשר לפתרה באמצעות אלגוריתמים גנטיים. את הפרמטרים הרבים של פונקציית ההערכה (קרי, התכונות השונות המזוהות עם עמדה) אפשר לקודד ככרומוזום. יש לאתחל באופן אקראי כרומוזומים רבים כאלה, שכל אחד מהם מייצג מופע אחד של פונקציית ההערכה. לאחר מכן, יש להעביר את האוכלוסייה אבולוציה עד שמתקבלות פונקציית הערכה מכווננות היטב.

ואולם, מכשול משמעותי שמנע עד כה את השימוש באלגוריתמים גנטיים היה כרוך בחישוב פונקציית ההתאמה. בהינתן קבוצת פרמטרים של פונקציית ההערכה (המקודדת ככרומוזום), כיצד מחשבים את ערך ההתאמה? במשך שנים רבות, הגישה המקובלת היתה להניח לאורגניזמים בכל דור (כאשר כל אורגניזם הוא תוכנת שחמט בעלת פונקציית הערכה הולמת) לשחק זה נגד זה סדרה של משחקים ולקחת את הציון של כל אחד מהם כערך ההתאמה שלו.

חסרונה העיקרי של גישה זו הוא משך הזמן הבלתי סביר הנחוץ לאבולוציה של כל דור. לפיכך הוצבו מגבלות חמורות על משך המשחקים הנערכים במשך כל דור ועל גודל האוכלוסייה. עם אוכלוסייה בגודל של 100 ומגבלת זמן מחמירה במיוחד של דקה אחת למשחק לכל צד, ובהנחה שכל שחקן משחק עשרה משחקים לפחות, תארך האבולוציה של כל דור 2,000 דקות. באילוצי זמן אלה,

יידרשו 70 יום להגיע לדור ה-50.

אם כן, קביעת ערכי התאמה באמצעות משחקים רבים איננה מעשית. אפשר, עם זאת, לגייס את המאגר האדיר של מידע קיים לא-מנוצל. אף שפונקציית ההערכה של תוכנות שחמט קיימות הן סוד שמור היטב, תוכנות שחמט נוהגות, מעשה שגרה, לגלות (חלקית) את הציון לכל מצב שבו הן נתקלות במהלך משחק. דויד, קופל ונתניהו ביצעו, באמצעות אלגוריתמים גנטיים, הנדסה-לאחור של פונקציית ההערכה האלה. בגישתם "מסתייעת-המנחה", הם הראו כיצד לבצע הנדסה-לאחור של "התנהגות המנחה" על-ידי אבולוציה יעילה בשלבים כדלקמן:

– צרו רשימה של בעיות אקראיות (במקרה של שחמט – מצבים).

– לכל בעיה, תנו למנחה להעריך את הבעיה ושמרו את התוצאה.

– תנו לכל אורגניזם להעריך את כל הבעיות, וחשבו – על פני כל הבעיות – את הפרש הממוצע בין הערך שנותן האורגניזם ובין הערך שנותן המנחה. ההתאמה תהיה מתכונתית (פרופורציונית) הפוכה להפרש ממוצע זה.

גישה זו היא גישה כללית של "הנדסה-לאחור התנהגותית". במקרה שלנו, כל בעיה מזוהה עם עמדה נתונה על לוח השחמט, והמנחה הוא פונקציית ההערכה של מנוע שחמט עדכני ומשוכלל ביותר. במילים אחרות, יוצרים בכל דור רשימה של מצבי שחמט אקראיים, ומניחים למנוע שחמט חזק להעריך את כולם. לאחר מכן, נותנים לפונקציית ההערכה של כל אחד מהאורגניזמים להעריך את המצבים. ככל שהערכת האורגניזם קרובה יותר להערכת המנחה, ערך ההתאמה של האורגניזם גבוה יותר.

החוקרים בדקו שתי תוכנות שחמט, מאסטרו (Maestro) ופלוקון (Falcon)¹. שתי התוכנות משתמשות באותן שגרות חיפוש, אבל פונקציית ההערכה של מאסטרו מורכבת מקבוצה קטנה הרבה יותר של פרמטרים מוגדרים מראש. הפרמטרים של התוכנה החלשה יותר, מאסטרו, עברו

1. פלוקון היא תוכנת מחשב ברמת רב-אמן שפיתח אומיד דויד, אשר התחרתה בהצלחה באליפויות עולם בשחמט-מחשב.

ממחקרי אוניברסיטת בריאילן



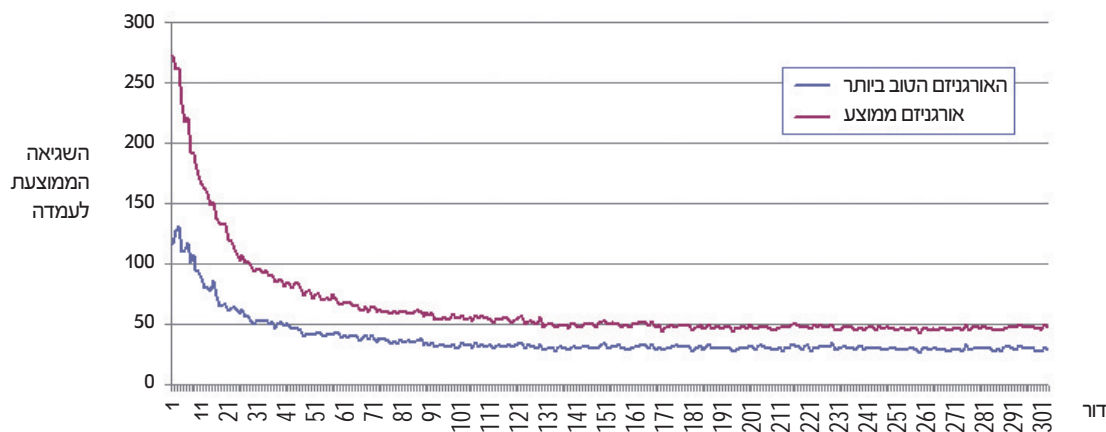
כדי למדוד את חוזק משחקה של התוכנה מאסטרו לאחר שעברה אבולוציה, נערכו 1,800 משחקים. מן התוצאות נראה בבירור שהתוכנה לאחר אבולוציה אינה נופלת מתוכנות שחמט מהחזקות בעולם. ספציפית, מאסטרו לאחר אבולוציה ניצחה את פלקון, המנחה שלה, ב-50% מהמשחקים, מה שמעיד על כך שההנדסה-לאחור של פלקון הצליחה עד כדי כך שהשתיים הפכו למעשה שוות בחוזקן. כמו כן, מאסטרו לאחר אבולוציה ניצחה ב-58% מהמשחקים כנגד קראפטי (Crafty), גרסת המשך של קריי בליץ (Cray Blitz) של רוברט הייאט (Hyatt), פעמיים אלופת העולם בשחמט-מחשב.

התוצאות שתוארו – אשר יוצגו השנה בכנס הבינלאומי על חישוביות גנטית ואבולוציונית – הן הראשונות המדווחות על גישת למידת מכונה שהניבה תוכנה ברמה של רב-אמן. החוקרים בחנו מגוון רחב של יישומים לגישתם, וכעת הם מתמקדים בשיתוף מנחים אנושיים בתהליך. הנדסה-לאחור של מוחו של קספרוב, למשל, תהיה אתגר קשה הרבה יותר, אך הם סבורים שהדבר ניתן לביצוע, לאור הישגי גישתם עד כה.

← ולבסוף, החוקרים מתכוונים להמשיך ולבדוק אם

אבולוציה כדי לחקות את התנהגות המנחה שלה, פלקון. הפרמטרים של מאסטרו אותחלו אקראית, ובכך נמחק למעשה כל ידע השחמט שלה. מאסטרו היתה צריכה ללמוד בכוחות עצמה את הערכים המיטביים של הפרמטרים של פונקציית ההערכה שלה.

החוקרים יצרו אוכלוסייה התחלתית של 1,000 אורגניזמים (שכל אחד מהם היה עותק של מאסטרו עם ערכי אתחול אקראיים), והעבירו אותה אבולוציה של 300 דורות. בכל דור, כל אורגניזם העריך 5,000 מצבים, וציוני ההערכה הושוו לאלה של המנחה, פלקון, לאותם מצבים. בתום כל דור, ניתן לאורגניזמים שציונם היה קרוב ביותר לזה של פלקון סיכוי גבוה יותר להזדווג. לאחר 300 דורות, היה הציון שנתן השורד הטוב ביותר קרוב מאוד לזה של פלקון, כשהפרש הציון הממוצע הוא פחות מ-1/3 ערך של רגלי. (לפי הקונבנציה המקובלת בשחמט, ערכו החומרי של רגלי הוא אחד, או 100 יחידות). במילים אחרות, הגישה מסתייעת-המנחה הצליחה לחולל אבולוציה מתוכנה חסרת כל ידע בשחמט לתוכנה שציונה קרוב מאוד לזה של תוכנת שחמט מן המובילות בעולם. איור 4 מראה כמה מהר מצליח האורגניזם המתפתח לחקות את המנחה.



איור 4 השגיאה הממוצעת לעמדה של האורגניזם הטוב ביותר ושל ממוצע האוכלוסייה בכל דור (זמן כולל ל-300 דורות: 442 שניות)

ביום ד' 13.8.08, בשעות 19:00-22:00, יתארח אומיד דויד בפורום "גליליאו" וישמח לענות על שאלותיכם. המעוניינים מוזמנים לפורום מדע וחברה באתר "גליליאו":
www.ifeel.co.il/galileo

לקריאה נוספת:

C. Darwin, *On the Origin of Species by Means of Natural Selection*, John Murray, 1859.
 T.A. Marsland and J. Schaeffer, *Computers, Chess, and Cognition*, Springer-Verlag, 1990.
 J.H. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, 1975.
 D.E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, 1989.
 O. David, M. Koppel, and N.S. Netanyahu, Genetic Algorithms for Mentor-Assisted Evaluation Function Optimization, (to appear) in *Proceedings of the ACM Genetic and Evolutionary Computation Conference*, Atlanta, GA, July 2008.



אפשר להחיל את המתודולוגיה שלהם כדי לכוונן פרמטרים באמצעות אלגוריתם גנטי על מגוון רחב של בעיות שבהן הפלט של פונקציית ההערכה של המנחה זמינה לצורכי אימון. ■

מאנגלית: מיכל גולן

אומיד דויד הוא בעל תואר ראשון ושני במדעי המחשב מאוניברסיטת בר-אילן, וכעת דוקטורנט במחלקה זו. תחום מחקרו העיקרי הוא חישוביות אבולוציונית, ובפרט, שימושיה בתוכנות אינטליגנטיות המתפתחות אוטומטית. הוא המפתח של תוכנת השחמט פלקון, שהשתתפה בשתי אליפויות עולם בשחמט-מחשב, והיה גם חבר הוועדה המארגנת של אליפות העולם ה-12 לשחמט-מחשב, אולימפיאדת המחשבים ה-9 והכנס הבינלאומי ה-4 למחשבים ומשחקים, שהתקיימו באוניברסיטת בר-אילן בשנת 2004.

פרופ' משה קופל מלמד מדעי המחשב באוניברסיטת בר-אילן. בעל תואר דוקטור מאוניברסיטת ניו-יורק על עבודתו בלוגיקה ובתורת המספרים, ושימש כעמית אורח במכון ללימודים מתקדמים בפרינסטון. מחקרו הנוכחי מתמקד ביישומי למידת מכונה, ובייחוד בתחום של סיווג טקסטים. כתב ספרים אחדים בתחום המשפט העברי ומשמש כיועץ בנושאים חוקתיים בכנסת.

פרופ' נתן נתניהו הוא בעל תואר ראשון ושני בהנדסת חשמל מהטכניון, ובעל תואר שני ושלישי במדעי המחשב מאוניברסיטת מרילנד, קולג' פארק. כיום הוא פרופסור חבר במחלקה למדעי המחשב באוניברסיטת בר-אילן, ועמית מחקר במרכז לחקר המוח באוניברסיטת בר-אילן ובמרכז לחקר אוטומציה באוניברסיטת מרילנד, קולג' פארק. תחומי מחקרו העיקריים הם אנליזה ותכנון של אלגוריתמים, גאומטריה חישובית, משחקי מחשב, עיבוד תמונה וחישה מרחוק, זיהוי תבניות ואומדן סטטיסטי חסין. היה שותף לכתיבתם של יותר מ-50 מאמרים שהופיעו בכתבי-עת מדעיים ושימש כעורך-משנה בכתב-העת *Pattern Recognition*.