

הלחנה אלגוריתמית של סולואים
באמצעות אלגוריתמים גנטיים

סמינר בנושא מיוחד במדעי המחשב
20375

15/03/2020

חנן וולט
בהנחיית ד"ר מיה הרמן

תוכן עניינים

1	תוכן עניינים	1
2	מבוא	1
4	חלק I: מהם אלגוריתמים גנטיים?	2
4	2.1 רקע ומוטיבציה	2.1
4	2.1.1 בעיות אופטימיזציה, מרחבי חיפוש ויוריסטיקה	2.1.1
6	2.1.2 אבולוציה כתהליך מיטוב	2.1.2
6	2.1.3 מידול בעיות לתהליכים אבולוציוניים	2.1.3
7	2.2 האלגוריתם הגנטי הבסיסי	2.2
7	2.2.1 תיאור האלגוריתם	2.2.1
8	2.2.2 פסידו-קוד	2.2.2
9	2.3 סקירה מפורטת של שלבי האלגוריתם	2.3
9	2.3.1 אתחול דור ראשוני	2.3.1
9	2.3.2 שחלוף	2.3.2
11	2.3.3 מוטציה	2.3.3
12	2.3.4 שיערוך מידת הכשירות	2.3.4
13	2.3.5 בחירת מועמדים	2.3.5
15	2.3.6 תנאי עצירה	2.3.6
16	3 חלק II: יישום א"ג בהלחנת מנגינות	3
16	3.1 הלחנת סולו כמרחב חיפוש	3.1
17	3.2 ייצוג מנגינה כגנום	3.2
17	3.2.1 ייצוג אבסטרקטי של מנגינה	3.2.1
17	3.2.2 קידוד המנגינה	3.2.2
21	3.3 מימוש אופרטורים גנטיים על כרומוזומים מוסיקליים	3.3
21	3.3.1 אתחול דור ראשון של מנגינות	3.3.1
22	3.3.2 מוטציה של מנגינה	3.3.2
24	3.3.3 שחלוף בין מנגינות	3.3.3
25	3.4 הערכת כשירות המנגינות	3.4
29	4 סיכום	4
30	5 נספח: רקע תאורטי של המוסיקה המערבית	5
34	6 רשימה ביבליוגרפית	6

1 מבוא

הלחנה (קומפוזיציה) היא אומנות העוסקת בחיבור של יצירות מוסיקליות. **אלגוריתם**, בהגדרה חופשית, הוא מתכון המנחה כיצד לבצע תהליך מסוים, באמצעות מעקב אחר רצף סדור של פעולות. **הלחנה אלגוריתמית**, היא טכניקה לחיבור יצירות מוסיקליות בעזרתם של אלגוריתמים במטרה לבצע אוטומציה של התהליך או חלקים ממנו [2]. טכניקה זו היא יישום פרטי של מושג רחב יותר בבינה מלאכותית הנקרא **יצירתיות חישובית** (או יצירתיות מלאכותית), העוסק במלאכת האוטומציה, באופן חלקי או מלא, של תהליך יצירת תוצרי אומנות וניתוח ויצירתיות אומנותית באופן כללי [3].

מסתבר שהשימוש באלגוריתמים כאמצעי עזר למלאכת ההלחנה של יצירות מוסיקליות אינו דבר חדש כלל וכלל! ניתן למצוא דוגמאות לשימושים שכאלו עוד הרבה לפני עידן המחשב, למשל –

- א. במאה השמונה-עשרה עשו שימוש בשיטת 'Musikalisches Würfenspiel' המבוססת על צירוף מקטעים מוזיקליים (שהולחנו מראש) בסדר אקראי שנקבע על סמך תוצאות הטלת קובייה, כך שרצפי הטלות שונים קובעים תמורות חדשות של המקטעים ובהתאם מתקבלת יצירה חדשה [3,7].
- ב. בתחילת המאה העשרים, המלחין היהודי האוסטרי ארנולד שונברג פיתח את 'שיטת שנים-עשר הטונים' – מערכת שיטתית לחיבור יצירות א-טונאליות הנותנת מעמד שווה לכל אחד משנים-עשר הצלילים של הסולם הכרומטי מבלי להיצמד לסולם טונאלי ספציפי [3]. שיטה זו בונה יצירה מוסיקלית ע"י סידור רצף של שנים-עשר צלילי הסולם הכרומטי בסדר כלשהו וחזרה על רצף זה תוך כדי ביצוע וריאציות שונות על כל רצף כאשר הווריאציות המותרות מוגדרות באוסף קבוע מראש. וריאציות לדוגמה הן היפוך סדר הצלילים שברצף ו/או היפוך כיוון המרווחים שבין הצלילים שברצף.

לב היצירתיות שבשיטות אלו אמנם טמון בגורם האנושי שמשתמש בהן (בדוגמה א - הלחנת מקטעי היצירה, ובדוגמה ב - סידור הרצף הכרומטי הראשון ובחירת סדרת הווריאציות העוקבות), אך בהחלט ניתן לראות בהן שיטות של הלחנה אלגוריתמית, שכן הן מוכתבות ע"י סדרת שלבים לביצוע ומובילות לתוצר סופי שמעוצב ומושפע בחלקו לא רק ע"י שיקולי הדעת של המלחין האנושי אלא גם ע"י האלגוריתם עצמו (בדוגמה א - אקראיות הקובייה, ובדוגמה ב - ציות לכללי הווריאציות המותרים).

בעידן המחשב, שילוב אלגוריתמים עם כוח חישובי מאפשר לייצר תכניות אוטונומיות היכולות להלחין יצירות מוסיקליות שלמות באופן אוטומטי ללא התערבות גורם אנושי, ו/או תכניות אינטראקטיביות שבהן הגורם האנושי מעורב בתהליך ונעזר במחשב בכדי להאיץ את תהליך ההלחנה ע"י חיפוש רעיונות חדשים להשראה, כיוון וכיוונון החיפוש של רעיונות ע"פ טעמו האישי, וכן ניתוח ודירוג אוטומטי של טיב הלחן הנוצר ע"י המחשב לצורך סינון פלטי ביניים שיוצאים פחות מוצלחים.

בעשורים האחרונים פותחו שיטות וטכניקות מוגוונות להלחנה אלגוריתמית [3], בין היתר – אלגוריתמים המבוססים על חוקי דקדוק וכללי גזירה, אלגוריתמים אקראיים המשתמשים בתהליכים סטוכסטיים ואלגוריתמי למידת מכונה העושים שימוש ברשתות נוירונים מלאכותיות. בסמינר זה נתמקד במשפחה מסוימת של אלגוריתמים, הנקראים 'אלגוריתמים גנטיים', המבוססים על תהליכים אבולוציוניים וביולוגיים מהטבע.

האלגוריתמים הגנטיים צברו פופולריות רבה בשלושה העשורים האחרונים והוכיחו את עצמם כיעילים במיוחד בתחום ההלחנה האלגוריתמית בכלל ובתחום ההלחנה שנעסוק בו בסמינר זה בפרט – הלחנת מלודיות עבור יצירות קיימות: קומפוזיציה מורכבת מהמון תתי-משימות, בין היתר – הלחנת מהלך הרמוני (רצף אקורדים), הובלת קולות למהלך זה, הלחנת הריתמיקה (קצב), התאמת מלודיה תחת ההרמוניה, ועוד. במשימת ההלחנה של מלודיות עבור יצירות קיימות, המרכיבים הריתמיים וההרמוניים של היצירה כבר קיימים ויש לטפל רק בחיבור המנגינה. משימה זו מופיעה באופן תדיר אצל מוסיקאים ונגנים: במוסיקת הג'אז למשל, קיים אוסף של "סטנדרטים" - אוסף יצירות ג'אז מוכרות, שנוהגים לנגן בהופעות תוך שיבוץ סולואים מאולתרים. אקט זה של אלתור על סטנדרט הוא למעשה מקרה פרטי של הלחנת מנגינה עבור יצירה קיימת. דוגמה נוספת היא הלחנת סולו לשיר חדש: במוסיקה הפופולרית המושמעת ביום יום ברדיו, מבנה טיפוסי של שיר הוא מבנה מהצורה "פתיחה, בית-פזמון, בית-פזמון, סולו, פזמון, סיום". **הסולו** בשיר, הוא חלק שבו תשומת הלב מופנית לאחד מכלי הנגינה (למשל סקסופון או גיטרה חשמלית) שמוטל עליהן לנגן איזושהי מנגינה המתאימה לשיר, מבחינת האווירה הכללית, הרקע ההרמוני, התחושה הקצבית, ה-"צבע" וכד', ועל ידי כך ליצור בה נופך נוסף או אף נקודת שיא, שמרעננים את השיר עם אנרגיות חדשות ומובילים בבטחה לסיום מוצלח שלו.

¹ בגרמנית: 'משחק קוביות מוזיקליות'.

בסמינר זה נעסוק בהלחנה אלגוריתמית של סולואים (מנגינות) עבור יצירות קיימות, תוך שימוש באלגוריתמים גנטיים. הסמינר מאורגן לשני חלקים כדלקמן –

❖ החלק הראשון (פרק 2) עוסק באלגוריתמים גנטיים גרידא, וסוקר אותם מנקודת מבט כללית, מבלי להגביל את הדיון למשימת ההלחנה האלגוריתמית דווקא. הפרק מתחיל ברקע ומוטיבציה שהובילו מלכתחילה לגיבוש משפחה זו של אלגוריתמים, וממשיך בתיאור אלגוריתם גנטי טיפוסי היכול לשמש כמודל ליישומים מעשיים וסקירת חלופות מימוש נפוצות.

❖ החלק השני של הסמינר (פרק 3) מתמקד ביישום מעשי של אלגוריתמים גנטיים עבור המשימה הספציפית של הלחנת מנגינות (וסולואים) ליצירות מוסיקליות קיימות. חלק זה סוקר שימושים מודרניים מהספרות האקדמאית ומדגים כיצד ניתן לממש אלגוריתם גנטי מותאם שיוכל לשמש מוזיקאים ונגנים ככלי עזר במלאכת חיבור מנגינת סולו עבור יצירה קיימת, בהתאם להעדפותיהם האישיות. לחסרי רקע בתאוריה מוסיקלית (תווים, סולמות וכד') מצ"ב נספח קצר בסוף הסמינר עם תקציר כלל המונחים המוסיקליים שנעשה בהם שימוש במסגרת סמינר זה, לרבות מבוא קצר לקריאת תווים.

2 חלק I: מהם אלגוריתמים גנטיים?

2.1 רקע ומוטיבציה

אלגוריתמים גנטיים הם אוסף של שיטות אופטימיזציה מטא-היוריסטיות המבוססות על עקרונות שאומצו בהשראת תהליכים אבולוציוניים מהביולוגיה [5,6]. בכדי להבין לעומק את המוטיבציה העומדת מאחורי אימוץ עקרונות שכאלו, ובכדי להעריך את מלוא הפוטנציאל הגלום בשיטות החיפוש שעושות בהם שימוש, ראשית נסקור את סוגי הבעיות שהן נועדו לפתור – הלוא הן בעיות האופטימיזציה הקשות, ולאחר סקירתן, נמשיך בהצגת הקשר שבין בעיות מסוג זה לתהליכים אבולוציוניים וניווכח בתועלת של אלגוריתמים גנטיים להתמודדות עם בעיות אלו.

2.1.1 בעיות אופטימיזציה, מרחבי חיפוש ויוריסטיקה

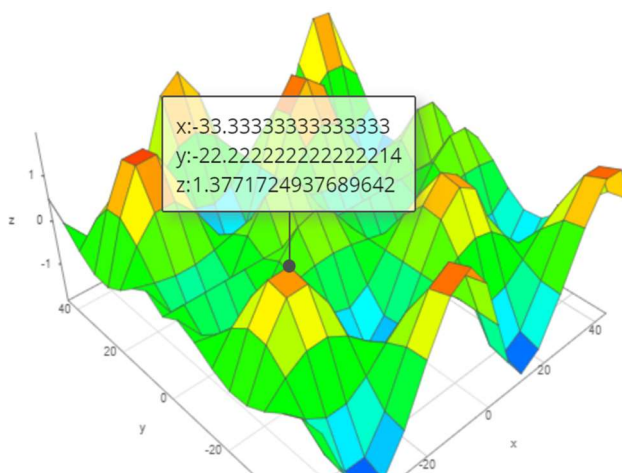
בעיות אופטימיזציה, בהגדרה חופשית, הן בעיות של מציאת קונפיגורציה של משתנים המניבה ערך אופטימאלי (מיטבי) עבור בעיית חיפוש נתונה, בכפוף לאילוצים מסוימים על המשתנים.

ניסוח פורמלי של בעיית אופטימיזציה צריך לכלול הגדרה של שלושה מרכיבים [6]:

1. **פונקציית מטרה** (*objective function*) שיש למטב. למשל – אם המטרה היא למקסם רווחים, פונקציית המטרה תיוצג ע"י פונקציה המגדירה את ערכי הרווח כתלות במשתני תחום ההגדרה הנידון.
2. **תחום הגדרה** (*decision variables*) – זהו תחום ערכי המשתנים שעליהם פועלת (מוגדרת) פונקציית המטרה. הפתרון שאנחנו מחפשים הוא איזושהי תצורה של משתנים אלו.
3. **אילוצים** (*constraints*) שהמשתנים שבתחום ההגדרה צריכים לקיים.

פתרון לבעיית אופטימיזציה הוא איזושהו צירוף (או קונפיגורציה) של משתנים בתחום ההגדרה. **פתרון-קביל** הוא צירוף של משתנים בתחום ההגדרה העומדים באילוצים החלים עליהם. **פתרון אופטימאלי** הוא פתרון קביל אשר מהווה נקודת אופטימום של פונקציית המטרה, דהיינו צירוף ערכים שעבורו פונקציית המטרה מקבלת ערך מיטבי. למשל, כאשר השאיפה היא למקסם את פונקציית המטרה, הפתרון האופטימאלי יהווה צירוף שבו פונקציית המטרה מקבלת את הערך **הגבוה ביותר**; ובאופן אנלוגי, כאשר השאיפה היא דווקא **למזער** את פונקציית המטרה, הפתרון האופטימאלי הוא כזה שמניב את ערך הפונקציה **הנמוך ביותר**.

דוגמה קלאסית לבעיה אופטימיזציה שכזו היא בעיית הסוכן הנוסע [5]: סוכן נדרש לצאת לסיבוב של פגישות במספר תחנות ולחזור בסוף אל נקודת המוצא, תחת האילוץ שאין לעבור באותה תחנה פעמיים (למעט נקודה המוצא). המטרה בבעיה זו היא למצוא מסלול (קונפיגורציה של סדר התחנות בביקור) העובר בין כלל התחנות בעלות מינימאליות, מבחינת הוצאות דלק, תשלומים עבור כבישי אגרה וכד'.



איור 1: הפונקציה $f(x, y) = z = \sin x \cdot \cos y \cdot \sqrt{y}$ **עם נקודת קיצון מקומית רבות שמקשות על מציאת נקודות מקסימום/מינימום גלובאליות לכלל המרחב.**

כאשר פונקציית המטרה היא פונקציה רציפה המוגדרת על משתנה יחיד, נח להמחישה באופן ויזואלי ע"י הצגתה כפונקציה במישור הממשי, כאשר צירי ה- x וה- y מייצגים (בהתאמה) את משתני תחום ההגדרה, ואת ערכי פונקציית המטרה. באופן דומה, פונקציית מטרה רציפה בשני משתנים ניתנת להצגה במרחב, כאשר מישור ה- $x-y$ מייצג את משתני תחום ההגדרה ואילו ציר ה- z את ערך פונקציית המטרה. הצגה שכזו נקרת מרחב הכשירות (*fitness landscape*). עבור מרחבי כשירות פשוטים ניתן לבצע רדוקציה של בעיית האופטימיזציה לבעיית מציאת נקודות מינימום/מקסימום מוחלט במרחב הכשירות.

עם זאת, המצב בד"כ אינו כה פשוט. בעיות אופטימיזציה מעשיות מחיי היום-יום הן בעיות קשות [5,6]. הקושי בא לידי ביטוי באופיין של פונקציות המטרה של בעיות אלו. תכונות טיפוסיות של פונקציות כאלו עשויות לכלול בין היתר – ממד גבוה (תחום ההגדרה מכיל מספר רב של משתנים), פונקציות לא גזירות, אולי אף עם נקודות אי-רציפות (הפונקציה מקוטעת) ומספר רב של נקודות קיצון מקומיות, נקודות פיתול ונקודות אוכף (הפונקציה מאוד תנודתית ומחוספסת ללא כיוון או מגמה ברורים). מטבע הדברים, פונקציות שכאלה נוטות להיות הרבה יותר מורכבות לניתוח. ראה למשל את [איור 1](#) המציג את מרחב הכשירות של פונקציה בשני המשתנים x, y המוגדרת בתת-

מרחב U של $\mathbb{R}^3 : U = \{(x, y, z) | f(x, y) = z\}$, ע"י הפונקציה: $f(x, y) = z = \sin x \cdot \cos y \cdot \sqrt{y}$. באיור מסומנת הנקודה $(x, y) = (-33.33, -22.22)$ שבה מתקבל ערך z שניתן להתרשם בבירור מהאיור שהוא ערך מקסימום מקומי, כלומר, הנקודה $(-33.33, -22.22)$ מהווה אמנם נקודת מקסימום ביחס לסביבה המקומית שלה, אך זו אינה נקודת מקסימום גלובלית ביחס למרחב U כולו. שכן יש נקודות בהן z מקבל ערכים גבוהים יותר. מרחבי כשירות שכאלה מקשים מאוד על החיפוש, שכן בסביבת נקודות הקיצון המקומיות אין פרספקטיבה על המרחב כולו, אז החיפוש עלול לטעות "ולהתביית" על נקודות כאלו בתור נקודות האופטימום לכאורה, על אף שבפועל אינן כאלה.

דרך כללית יותר למידול הבעיה, שמתאימה גם לפונקציות מורכבות שכאלה, היא הגדרת מרחב חיפוש². **מרחב חיפוש** הוא קבוצת כל הפתרונות הקבילים האפשריים, דהיינו כל וקטור במרחב זה הוא פתרון פוטנציאלי לבעיית האופטימיזציה. לעיתים נהוג לכנות את הווקטורים במרחב זה, שהם פתרונות פוטנציאליים לבעיה, גם מצבים, מועמדים או פשוט אינדיבידואלים. כאמור, בכדי לפתור בעיית אופטימיזציה, עלינו למצוא מבין כל הפתרונות הקבילים המועמדים את הפתרון האופטימאלי (או לפחות אחד מהפתרונות האופטימאליים במידה ויש כמה פתרונות המניבים את הערך האופטימאלי). לפיכך, בטרמינולוגיה של מרחב מצבים, עלינו למצוא מצב שבו פונקציית המטרה מקבלת ערך אופטימאלי. כיצד נמצא מצב כזה? כלומר, כיצד ניגש לפתור את בעיית האופטימיזציה? גישה נאיבית למדי, היא להתבונן במרחב החיפוש כעל גרף קשיר, אשר צמתיו הם המצבים (הפתרונות הפוטנציאליים), ולהפעיל על גרף זה סריקות כאלה או אחרות לרוחב או לעומק עד אשר יימצא הפתרון האופטימאלי המבוקש, אולם גישה זו נפסלת על הסף, שכן מרחב החיפוש גדל באופן מעריכי עם הגדלת ממד הבעיה/טווח הערכים של משתני תחום ההגדרה, לכן חיפוש ממצא (exhaustive search), הידוע גם כשיטת brute-force, אינו ישים תחת אף מודל חישובי מעשי לבעיות שכאלו. כיצד אם כן בכל זאת ניתן לבצע חיפוש מחוכם ללא סריקת כלל מרחב החיפוש? גישה פופולרית במיוחד להתמודדות עם בעיות אופטימיזציה קשות שכאלה, היא שימוש בטכניקות היוריסטיות.

חיפוש היוריסטי הוא חיפוש המכוון לגילוי מהיר של פתרונות סבירים, תוך השלמה עם העובדה שפתרונות אלו אינם בהכרח האופטימאליים. בבעיות אופטימיזציה קשות היוריסטיקה שכזו יכולה להוות פשרה טובה בהתחשב בנסיבות מרחב החיפוש. מימושים שונים של טכניקות היוריסטיות כוללים בין היתר – שימוש ב-"כללי אצבע", ניחושים מושכלים, אינטואיציה וסימולציות של ניסוי וטעייה. נציג היוריסטיקה לדוגמה במשחק השחמט: ניתן לראות את אוסף כל התצורות האפשריות (כל התצורות התקפות ע"פ כללי המשחק) כמרחב החיפוש ולהגדיר את פונקציית המטרה כפונקציה המעריכה את העמדה של התצורה הנתונה (למשל ע"פ קריטריונים כגון מצב הכלים ההדדי, יכולות התקיפה של כל צד, ומצב אבטחת המלך), ומפיקה ערך חיובי העולה ככל שיתרונו של השחקן הלבן גבוה יותר, ערך אפס כאשר העמדה שקולה, וערך שלילי היורד ככל שיתרונו של השחקן השחור גבוה יותר. מנקודת מבט זו, משחק השחמט הוא בעיית אופטימיזציה, כאשר מטרת השחקן הלבן היא למקסם את פונקציית המטרה, ומטרת השחקן השחור למזער אותה. שחקן מתחיל בשחמט שאינו בקיא בפתיחות השחמט המבוססות, יכול בשלב הפתיחה הקריטי לנקוט באסטרטגיה היוריסטית כדלהלן: "שאף לשליטה במרכז הלוח תוך פיתוח כלים מהיר ושיבוש תכניות היריב". אסטרטגיה זו היא די מופשטת, ומימושים שונים שלה לא בהכרח יובילו לעמדה האופטימאלית, אך יש בה יתרון רב שכן היא אינה דורשת ידע מוקדם של הכרת מאות או אפילו אלפי וריאציות, היא מקצרת את שלב קבלת ההחלטה על בחירת המסע הבא בטרטוריה לא מוכרת של תצורת המשחק, שכן היא מסננת מראש מהלכים שמפרים את העקרונות שלעיל ובאופן כללי מובטח בד"כ למי שמשחק על פי עקרונות אלו שיגיע לעמדה פחות או יותר שקולה לעמדת היריב, או לפחות לעמדה יציבה ומבוססת, גם אם היא לא העמדה הטובה ביותר שיכל להגיע אליה בווריאנט הנתון. נראה כעת את הקשר שבין בעיות אופטימיזציה לתהליכים אבולוציוניים המתרחשים בטבע.

² ידוע גם כמרחב הפתרונות (solution space).

2.1.2 אבולוציה כתהליך מיטוב

אורגניזמים (יצורים חיים) הם אולי העדות הטובה ביותר להמחשת העובדה שאבולוציה טבעית היא תהליך אופטימיזציה מוצלח שעובד הלכה למעשה [5]: ההרכב הגנטי נמצא בשאיפה מתמדת לשיפור. דור אחרי דור המבנה הגנטי מתאים את עצמו לסביבה תחת תנאים (אילוצים) משתנים ומוביל בתהליך של ברירה טבעית הישרדותית להרכב גנטי בעל הגנים החזקים ביותר (פתרון אופטימאלי).

תובנה זו היוותה השראה למדענים במאה העשרים שהחלו לחקור את הרעיון של אימוץ עקרונות ורעיונות מתחומי האבולוציה והגנטיקה לפתרון בעיות מעשיות. מחקרים אלו הביאו לתולדה שאבולוציה היא לא רק תהליך ברירת מחדל שמתרחש בטבע, אלא גם אמצעי אלגוריתמי.

שיטות שונות פותחו בנפרד באופן בלתי-תלוי לפתרון בעיות שונות [5], כאשר המשותף לכולן היה אדפטציה והשראה של רעיונות ועקרונות אבולוציוניים שבטבע. עם הזמן, רעיונות מעניינים שהתגלו בגרסאות מסוימות של האלגוריתם אומצו גם ע"י שיטות אחרות, כך שההבדלים שבין השיטות השונות טושטשו, וכיום מקובל לעגן את כולם במשפחה אחת תחת השם הכולל של אלגוריתמים גנטיים.

נראה כעת כיצד מבצעים פורמליזציה של תיאור בעיית אופטימיזציה במונחים של תהליכים אבולוציוניים.

2.1.3 מידול בעיות לתהליכים אבולוציוניים

בכדי לפתור בעיות אופטימיזציה עם תהליכים אבולוציוניים מלאכותיים, עלינו לבצע תחילה מידול לבעיה ולהמירה לייצוג התואם לאלגוריתמים הגנטיים. בהשראה מתורת הגנטיקה, באלגוריתמים גנטיים מייצגים כל פתרון מועמד ע"י **גנום** (genome), שמתאר את כלל המבנה הפנימי שלו (ובאנלוגיה לגנטיקה, את כלל ההרכב הגנטי של האינדיבידואל). הגנום מורכב מקבוצה לא ריקה של **כרומוזומים** (chromosomes), כל כרומוזום מכיל אוסף של **גנים** (genes), וכל גן מקבל ערכים מתוך קבוצה מוגדרת מראש של **אללים** (allele set). טווחי ערכים נפוצים של אללים הם סיביות, ערכים רציפים, ערכים דיסקרטיים ו/או רשימת סמלים מא"ב נתון.

ייצוג מקובל של כרומוזום כמבנה נתונים הוא ייצוג ע"י וקטור (מערך סדור), כאשר כל רכיב בווקטור (ובאופן שקול, כל תא במערך) מייצג גן בודד, אז בתכנית מחשב כרומוזומים יכולים להיות מיוצגים למשל ע"י מחרוזות של סיביות ו/או מערכים/רשימות של ערכים רציפים/דיסקרטיים.

הגנום המקודד לכרומוזום(ים) נקרא **גנוטיפ**, והפתרון הממשי המתקבל לאחר פענוח הקידוד בחזרה לצורתו המקורית נקרא **פנוטיפ**. ביישומים מסוימים המיפוי בין הגנוטיפ לפנוטיפ הוא מידי וישיר כך שאין מבחינים בין השניים [5].

הערת הבהרה: לעיתים נהוג לייצג את המבנה הפנימי המלא של פתרון מועמד ע"י כרומוזום **מודד**. במצב כזה המושגים של גנום וכרומוזום מתלכדים מבחינת המשמעות שלהם באלגוריתם הגנטי, שכן שניהם מייצגים את כלל המבנה הפנימי של פתרון אינדיבידואל. לפיכך, בהמשך, כאשר נציין כרומוזום, לעיתים הכוונה תהיה בעצם לגנום כולו, ולעיתים לכרומוזום בודד מתוך אוסף הכרומוזומים הכולל של הגנום. לרוב, המשמעות תהיה ברורה מההקשר, ולעיתים הדיון יהיה רלוונטי לשתי המשמעויות גם יחד. במקרים בהם נדרשת הבחנה בין שתי המשמעויות, הדבר יצוין במפורש.

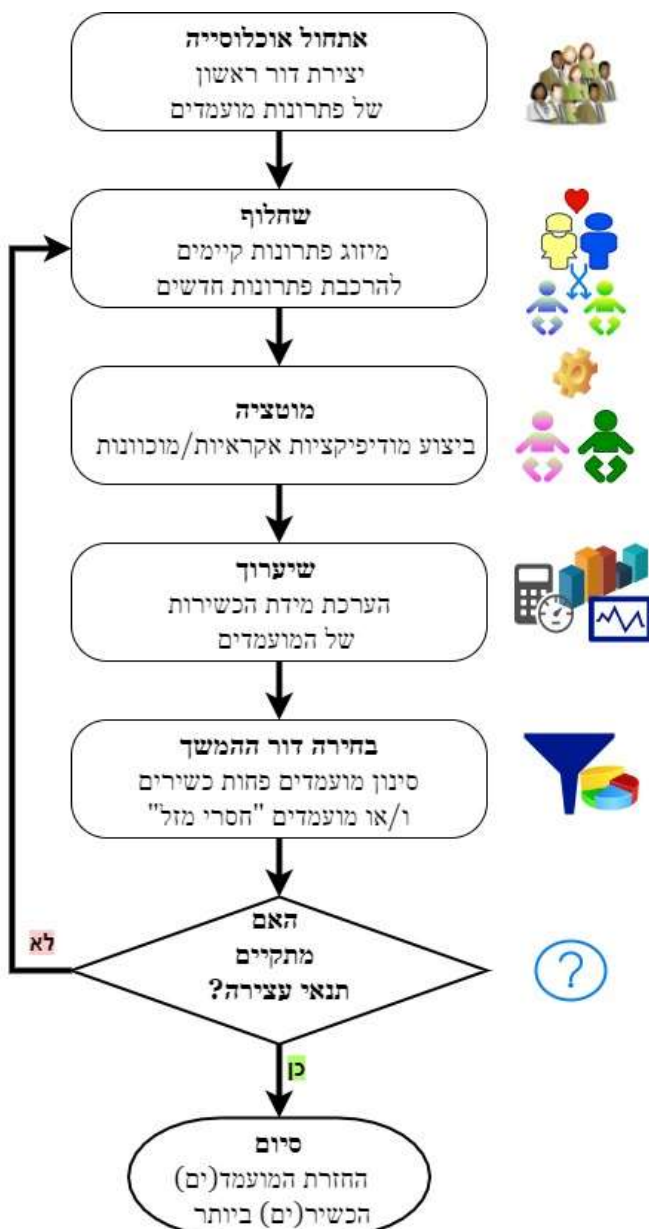
2.2 האלגוריתם הגנטי הבסיסי

2.2.1 תיאור האלגוריתם

בדומה לתהליך האבולוציה הביולוגית שמתרחש באוכלוסיות, ולא אצל פרטים בודדים, גם האלגוריתם הגנטי הקלאסי הבסיסי ששואב ממנו השראה עובד אף הוא על קבוצה (או אוכלוסייה) של פרטים. אוכלוסייה זו היא תת-קבוצה של מרחב החיפוש, ומכאן, כל אינדיבידואל באוכלוסייה זו הוא פתרון מועמד לבדיקה הנידונה. לפיכך כל מועמד עשוי להתגלות בתור הפתרון האופטימאלי שאותו אנחנו מחפשים.

האלגוריתם מתניע את תהליך האבולוציה המלאכותי עם אתחול דור ראשון של פתרונות מועמדים, שניתן לחולל הן באופן אקראי, והן באופן יזום על-סמך פרמטרים נבחרים. על בסיס אוכלוסייה זו של הדור הראשון מתחיל תהליך מחזורי (מעגלי) המדמה תהליכים אבולוציוניים מהביולוגיה (ראה איור 1 משמאל):

1. בשלב ראשון מדמים תהליך של זיווג/רבייה ע"י שימוש באופרטור השחלוף שממזג חלקים משני פתרונות או יותר (ההורים) ובונה מהם פתרונות חדשים (צאצאים).
2. בשלב שני מדמים "רעש" שקיים בטבע כגון מוטציות גנטיות והשפעות סביבתיות שכופות שינוי הסתגלותי (אדפטיבי) על הכרומוזום. שלב זה ממומש ע"י אופרטור מוטציה שאחראי לבצע מודיפיקציות תואמות (שינויים אקראיים ו/או שינויי הסתגלות) בחלק מהמועמדים.
3. בשלב שלישי מבצעים תהליך שיערוך (fitness evaluation) של אוכלוסיית הפתרונות



המעודכנת בכדי לאמוד את טיב האיכות שלהם ע"פ הסיווג המתקבל מפונקציית הכשירות.

4. בשלב רביעי מדמים את הבחירה הטבעית (natural selection) מהתאוריה של דארווין ע"י תהליך סלקציה שבו מבין כל המועמדים באוכלוסייה הנוכחית אילו ימשיכו לדור (מחזור) הבא ואילו יישארו מאחור ו-"יכחדו". קונספט הבחירה הטבעית בא כן לידי ביטוי ע"י מתן העדפה בבחירה למועמדים הכשירים ביותר, בהתאם להערכה שבוצעה בשלב הקודם.

5. לאחר שלב הסינון מסתיים מחזור והאלגוריתם מגיע לנקודת צומת דרכים שבה הוא צריך לקבל החלטה אם להמשיך בחיפוש ע"י התחלת מחזור נוסף בניסיון לשפר עוד יותר את האוכלוסייה הקיימת או שמא הגיעה העת לעצור (למשל אם הפתרונות שנמצאו טובים מספיק, או שמא בשל אילוצי זמן/כוח חישובי יש להסתפק בפתרונות הקיימים).

איור 2 - תרשים זרימה של האלגוריתם הגנטי הקלאסי הבסיסי

2.2.2 פסידו-קוד

להלן פסידו-קוד של האלגוריתם שתואר לעיל –

```

1. אתחל דור ראשון
2. קצף -
   2.1 ne לִפְנֵי
   2.2 מִצֵּיט
   2.3 עִצְרוֹק
   2.4 כְּחִירַת דֹּר הַמֶּשֶׁק
3. קְדוֹק הָאֵם מִתְקִיט תַּנְאִי הַצִּירָה -
   3.1 אֵם כֵּן, הַחֲזִיר כִּפְלֵט אֵת הַמִּוֹצֵא (יֵם) הַכְּשִׁיר (יֵם) בְּיוֹתֵר וְסִיט;
   3.2 אַחֲרַי, חֲזוֹר לִפְנֵי 2.
  
```

כפי שניתן להתרשם מהפסידו-קוד ומהתיאור שקדם לו, האלגוריתם הגנטי הבסיסי מהווה תבנית גנרית, המציינת סדרה של פעולות/שגרות אבסטרקטיות, ללא תיאור קונקרטי של אופן הביצוע שלהן. הסיבה לכך היא שאלגוריתם זה מייצג משפחה של אלגוריתמים, אשר נבדלים ביניהם במימושים השונים של תבנית זו, בכדי להתאימם באופן מיטבי עבור איזושהי בעיית אופטימיזציה מסוימת נתונה (וכן עבור הייצוג באותה בעיה (רציפים/דיסקרטיים). חלקן אף מבצעים את השלבים בסדר אחר או מדלגים על שלב מסוים [5].

בסעיף הבא נסקור במפורט את שלבי האלגוריתם ונדון בגישות שונות למימושים קונקרטיים של השגרות והאופרטורים השונים של האלגוריתם.

2.3 סקירה מפורטת של שלבי האלגוריתם

2.3.1 אתחול דור ראשוני

כאמור, בשלב האתחול יוצרים את האוכלוסייה הראשונית של הפתרונות המועמדים. לאוכלוסייה זו עשויה להיות השפעה מכרעת הן על כיוון החיפוש, הן על קצב התקדמות החיפוש והן על הפתרון הסופי המוחזר בסופו של דבר כפלט, שכן היא מהווה בסיס לפרישת דור ההמשך וע"י כך לוקחת חלק מהותי בעיצוב כל הצאצאים שיווצרו לאחר מכן. לפיכך, חשוב לבצע את האתחול באופן נאות.

אופן האתחול תלוי במידה רבה הן בבעיית האופטימיזציה הנתונה והן באופן הייצוג של הכרמוזומים בתנאי הבעיה, אך באופן כללי ניתן לסווג את המימושים השונים לאחת משתי גישות מרכזיות – אתחול אקראי ואתחול מכוון –

- **אתחול אקראי** מתאים במיוחד לתרחישים בהם מרחב החיפוש מחוספס, עמום ותנודתי, ובאותה מידה גם לבעיות שאנחנו לא בקיאים בתחומן. בגישה זו, המרחב כולו טוב באותה מידה באשר לסיכויי הימצאות הפתרון האופטימאלי, לכן פשוט בוחרים באופן אקראי כמה נקודות (פתרונות) במרחב ומתחילים את החיפוש מנקודות אלה. למשל, אם הנקודות במרחב מיוצגים ע"י רצף סיביות אז מימוש אתחול אקראי לדוגמה יכול להיות פשוט הטלת מטבע הוגן עבור כל רכיב ברצף לקביעה אם לאתחל אותו עם ערך 0 או 1.
- **אתחול מכוון** מבוסס על רקע מוקדם והכרות עמוקה של תחום הבעיה. אתחול זה מתאים לתרחישים בהם למתכנן האלגוריתם יש איזשהו כיוון או השערה על אזורים מסוימים במרחב שבהם יש סבירות גבוהה יותר להימצאות הפתרון האופטימאלי לעומת אזורים אחרים. דרך פופולרית למימוש אתחול מכוון שכזה היא אתחול על-סמך איזשהו "גרעין" (initialization seed) – נקודה במרחב שאנחנו מאמינים שנמצאת לא רחוק מנקודת האופטימום. נקודה כזו יכולה להיות איזושהי נקודת אבן-בוהן (benchmark), דהיינו נקודה שכבר סווגה ע"י פונקציית הכשירות לדירוג שאנחנו מחשיבים כטוב, או לחלופין נקודה המורכבת מרכיבים (דהיינו חלקי פתרונות – גנים) שיש לנו בהם עניין. אוכלוסייה של נקודות גרעין יכולה להיות באופן זה את הדור הראשון. במקרה של נקודת גרעין יחידה, ניתן לבנות ממנה אוכלוסייה גדולה כרצוננו ע"י ביצוע חוזר של פעולות מוטציה ושחלוף (יצירת נקודה חדשה ע"י עדכון נקודה קיימת ולאחר מכן יצירת זוג נקודות חדש על סמך השתיים הקיימות).

מהשוואה שבוצעה בין גישה האתחול האקראית לבין זו המכוונת [7], עולה כי אתחולים אקראיים נוטים להוביל למגוון רחב יותר של פתרונות, אך יחד עם זאת, תהליך ההתכנסות איטי במיוחד ולעיתים אף גרוע מכך – במקרה של "חוסר מזל", בחירת אקראית לא מוצלחת במיוחד של ערכים התחלתיים יכולה ליצור כאוס, שיוביל להתכנסות לערך כשירות יחסית נמוך.

לעומתה, הגישה המכוונת דווקא נוטה להוביל לתהליך התכנסות מהיר לפתרונות המסווגים עם כשירות די גבוהה, אולם כפי שאומר הפתגם הישן: "התפוח לא נופל רחוק מהעץ", בגישה זו הפתרונות שנמצאים אמנם בד"כ טובים אך לא שונים מהותית מהגרעין, ולכן סובלים מחוסר גיוון ומקוריות.

פשרה טובה עשויה להיות מימוש אתחול המשלב את הרציונל שמאחורי שתי הגישות, דהיינו אתחול המבוסס על ידע מוקדם תוך פיזור אקראיות בכדי להרחיב את מנעד החיפוש הראשוני כך שיכלול גם טריטוריה פחות מוכרת.

2.3.2 שחלוף

אופרטור השחלוף (בלעז: crossover), פועל כאמור בהשראת תהליך הזיווג/הרבייה של זוג יצורים חיים בטבע, שבסופו נולד לזוג צאצא שהרכבו הגנטי מכיל צירוף מהגנים שעברו אליו בתורשה מהוריו. באופן דומה, באלגוריתמים גנטיים, אופרטור השחלוף יוצר מועמד חדש (צאצא) שתוכנו הגנטי מתקבל ע"י ערבוב וצירוף של רכיבים גנטיים של שני מועמדים (הורים) או יותר. המוטיבציה העומדת מאחורי זיווג מלאכותי שכזה, היא ניסיון לשמר ולשפר גנים "טובים" המייצגים חלקי פתרונות אופטימאליים, בתקווה ושאיפה שצירוף של גנים טובים יניב בדור ההמשך צאצא שהרכבו הגנטי אף מתעלה על זה של הוריו.

בדומה ליתר האופרטורים הגנטיים, גם אופן מימוש אופרטור השחלוף תלוי במידה רבה באופן ייצוג הכרומוזומים –

- שיטת ה-n-point crossover – עבור ערכים דיסקרטיים וסיביות – עבור כרומוזומים המיוצגים ע"י וקטורים של סיביות או ערכים בדידים, מימוש פופולרי במיוחד להצלבת כרומוזומים הוא שחלוף בשיטת ה-n-point crossover, המבצעת n הצלבות בין זוג כרומוזומים, כאשר n פרמטר שערכו קטן או שווה למספר הגנים בכל כרומוזום. השיטה בוחרת באקראי את המיקומים (אינדקסים) של n נקודות פיצול, ובכל נקודה שכזו שני הווקטורים משחלפים ביניהם את כלל הרכיבים (גנים) הנמצאים בין נקודת פיצול זו לנקודה פיצול הבאה (או כלל יתרת הרכיבים, אם זוהי נקודת הפיצול האחרונה), כאשר הרכיבים נבחרים לסירוגין בכל נקודת פיצול, פעם מהווקטור הראשון, ופעם מהווקטור השני. למשל בהינתן שני וקטורים \bar{x}, \bar{y} באורך k של סדרת סיביות או ערכים דיסקרטיים –

$$\bar{x} = (x_1, x_2, x_3, \dots, x_{k-2}, x_{k-1}, x_k)$$

$$\bar{y} = (y_1, y_2, y_3, \dots, y_{k-2}, y_{k-1}, y_k)$$

עבור $n = 3$ ושלוש נקודות פיצול תואמות s_1, s_3, s_3 , שהבחירה האקראית של מיקומן הניבה למשל את האינדקסים $s_1 = 4, s_2 = i, s_3 = j$ כאשר $4 < i < j \leq k$, נקבל כתוצאה מהשחלוף את זוג הצאצאים "התאומים" הבא (נקודות הפיצול מודגשות ע"י סוגריים מרובעות) –

$$\bar{w} = (x_1, x_2, x_3, [y_4], y_5, \dots, y_{i-1}, [x_i], x_{i+1}, \dots, x_{j-2}, x_{j-1}, [y_j], y_{j+1}, \dots, y_{k-1}, y_k)$$

$$\bar{z} = (y_1, y_2, y_3, [x_4], x_5, \dots, x_{i-1}, [y_i], y_{i+1}, \dots, y_{j-2}, y_{j-1}, [x_j], x_{j+1}, \dots, x_{k-1}, x_k)$$

נציין כי שיטה זו, כאשר היא מופעלת על זוג פתרונות, מניבה תמיד שני פתרונות חדשים. אם ברוח תהליך הרבייה המלאכותית נרצה לנסות להכליל את ההגדרה של שיטה זו באופן שתפעל על מספר כלשהו של פתרונות m , עבור $m \in \mathbb{N}$ ($m \geq 2$) וערך פרמטר n , נקבל בכל הפעלה $m \cdot (m-1)^n$ צאצאים חדשים, שכן עבור כל פתרון המשתתף כאופרנד יש n נקודות פיצול שבכל אחת מהן יש $m-1$ אפשרויות שחלוף (כמספר יתרת הפתרונות המשתתפים בתהליך) לכן כל אופרנד מפיק $(m-1)^n$ צאצאים, ויש m מועמדים המשתתפים, לכן בסה"כ נקבל: $m \cdot (m-1)^n$.

- שיטת ה-intermediate crossover – עבור ערכים רציפים – עבור כרומוזומים המיוצגים ע"י סדרת ערכים רציפים (ממשיים), מקובל לבצע שחלוף המביא לידי ביטוי את ערכי הגנים (רכיבים) עצמם (להבדיל למשל משיטת ה-n-point crossover שמשחלפת רק את המיקומים שלהם). מימוש פשוט למדי הוא שיטת ה-intermediate crossover (מוכרת גם כ- arithmetic crossover) הקובעת לכל רכיב את הממוצע אריתמטי של ערכי הרכיבים התואמים של הפתרונות המשתתפים בפעולה. שיטה זו מתאימה באופן טבעי לכל מספר אופרנדים. למשל, עבור שלושת הווקטורים $\bar{x}, \bar{y}, \bar{z}$ מעל \mathbb{R}^n הנתונים ע"י –

$$\bar{x} = (x_1, \dots, x_n), \bar{y} = (y_1, \dots, y_n), \bar{z} = (z_1, \dots, z_n)$$

צאצא יחיד \bar{u} שרכיביו נתונים ע"י –

$$\bar{u} = (u_1, u_2, \dots, u_n) = \left(\frac{x_1 + y_1 + z_1}{3}, \frac{x_2 + y_2 + z_2}{3}, \dots, \frac{x_n + y_n + z_n}{3} \right)$$

אם יש עניין בהפקת יותר מצאצא אחד בכל פעולת שחלוף, כפי שמבוצע למשל בשיטת ה-n-point crossover, ניתן להגדיר וריאנט שלה המבצע גם ממוצע גיאומטרי ו/או ממוצע הרמוני, וע"י כך קבלת שניים או שלושה צאצאים בכל פעולה, בהתאמה.

מעבר למימושים על אופן אקט השחלוף, סוגיה נוספת חשובה לא פחות היא השאלה איזה מהפתרונות המועמדים יש לבחור כאופרנדים להשתתפות בפעולת השחלוף. כלומר, מיהם הפתרונות שכדאי לבחור על מנת שישמשו כהורים לצאצאים חדשים? אולי אלו הכשירים ביותר, או שאולי כדאי לבצע איזון ולמזג מועמדים מכל קצוות ה-"פריפריה"? גישה נפוצה במיוחד בשלב השחלוף להתמודדות עם סוגיה מורכבת זו "מבלי לשבור את הראש" היא התחמקות פשוטה ע"י בחירה אקראית של מועמדים תחת התפלגות אחידה לכולם, ולדחות את השיקולים

והדילמות הכרוכים בבירור סלקטיבית לשלב היותר מאוחר של בחירת המועמדים לדור ההמשך לאחר שערוכם ע"י פונקציית הכשירות.

לחלופין, אם יש לנו הכרות טובה עם מרחב הכשירות, או אם אנחנו רוצים לתעל את החיפוש לכיוון או אזור מסוים, נוכל לבחור את ההורים בהתאם – למשל, אם יש אזור במרחב שנראה מבטיח שנרצה להתמקד בו לביצוע חיפוש מקומי לעומק, נוכל לבצע את השחלוף בין מועמדים סמוכים הנמצאים בסביבת אזור זה (למשל ע"י הגדרת פונקציית מרחק על המועמדים ואיתור המועמדים שקרובים דיים לאיזושהי נקודה מרכזית באותה סביבה). לחלופין, אם נרצה דווקא להתפרש לאזורים חדשים ע"י חיפוש רחבי נוכל לבחור במועמדים קוטביים שנמצאים במרחק גדול זה מזה.

2.3.3 מוטציה

אופרטור המוטציה (בלעז: mutation), מבוסס על שינויים אקראיים בגנים של הכרומוזום. בטבע לרוב אלו שינויים מינוריים. באלגוריתמים גנטיים, אופרטור המוטציה מקבל כרומוזום כארגומנט, ופועל לשינוי ועדכון המבנה שלו ע"י ביצוע עדכונים מקומיים בחלק מהגנים שלו. בכדי לשקף את האקראיות מהתהליך שבטבע, נהוג להתנות את ביצוע העדכונים המקומיים השונים ע"י איזושהי פונקציית הסתברות, הניתנת לשליטה והכוונה ע"י פרמטר חיצוני.

באופן כללי, ישנן שלוש תכונות עיקריות שמצופה מאופרטור המוטציה לקיים בכדי שיוכל יתרום לתהליך אבולוציוני אפקטיבי [5] –

- א. **נגישות** – כל נקודה אפשרית במרחב החיפוש צריכה להיות נגישה מכל נקודה אחרת במרחב. באנלוגיה לתורת הגרפים, אם נייצג את מרחב החיפוש כגרף קשיר אשר קבוצת צמתיו היא קבוצת הפתרונות המועמדים, נצפה מאופרטור המוטציה שלא ימחק קשתות שיגרמו לניתוק בין שני רכיבי קשירות, שכן אז כבר לא יוותר מסלול בחזרה שיקשר בין שני אזורי החיפוש.
- ב. **ניטרליות** – העדכונים צריכים להיות נטולי הטיה בכיוון או מגמה מסוימים. תכונה זה רלוונטית בעיקר למרחבים ללא מגבלות ואילוצים. כאשר יש אילוצים, הטיה דווקא יכולה להיות יתרון ע"י הימנעות מלכתחילה משינויים שמובילים למצב המפר את האילוצים.
- ג. **סקלביליות** – האופרטור צריך להיות בעל יכולת התאמה ושליטה על עוצמת השינוי והיקף השינוי, למשל ע"י חשיפת פרמטר אחד שמאפשר לציין את דרגת השינוי ופרמטר שני הסתברותי שמאפשר לציין את ההסתברות המבוקשת לביצוע השינוי.

באנלוגיה לחיפוש במרחב, ניתן לראות את אופרטור המוטציה כצעדי חיפוש במרחב הכשירות. מנקודת מבט זו, תכונת הסקלביליות מאפשרת להתאים את קצב ההליכה ולקבוע עד כמה נרצה להתרחק מהנקודה שאנו נמצאים בה כעת במרחב.

מקובל להבחין בין שתי צורות חיפוש עיקריות במרחבי כשירות – **חיפוש לעומק** – exploitation search (המוכר גם כ- intensification search) המעמיק את החיפוש בסביבה המקומית הנוכחית, תחת ההנחה שזהו אזור מבטיח ושהפתרון האופטימאלי נמצא בקרבת מקום ו**חיפוש לרוחב** – exploration search (המוכר גם כ- diversification search), שמבצע אמנם חיפוש יחסית שטחי, אך שואף לכסות סביבה כמה שיותר רחבה בכדי לנסות לזהות אזורים אטרקטיביים חדשים שבהם יהיה כדאי להתמקד ולבצע חיפוש לעומק. החיפוש לרוחב יכול לשמש גם כאמצעי מניעה לסיכון של "היתקעות" על נקודת קיצון מקומית, ע"י כך שהוא מרשה עצמו להתקדם בצעדים גדולים גם אל מחוץ לסביבה שכרגע יש בה אולי איזשהו קיצון מקומי שלכאורה נראה מבטיח, אולם חיפוש זה כבד יותר בביצועים [5]. בחיפוש לעומק לעומת זאת אמנם הביצועים טובים יותר, אך טווח הראייה קצר יותר, ולכן אם החיפוש יגיע לנקודת קיצון מקומית, ייתכן וייתקע בסביבה שלה.

בהקשר של אופרטור המוטציה, בכדי לבצע חיפוש לעומק יש לכוון את הפרמטרים לשינויים עדינים (דרגת שינוי נמוכה) ובתדירות (הסתברות) נמוכה, ומנגד, בכדי לבצע חיפוש לרוחב יש לכוון את הפרמטרים כך שיבוצעו שינויים משמעותיים יותר ובתדירות גבוהה יותר. אחד היתרונות הבולטים שמייחדים את האלגוריתמים הגנטיים הם שניתן לכוון פרמטרים אלו באופן דינאמי בזמן ריצה בהתאם ל-"מצב בשטח", דהיינו בהתאם לסטטוס החיפוש הנוכחי, האזורים שהתגלו עד כה, וההנחות שלנו באשר לאטרקטיביות שלהם והערכות לגבי הימצאות הפתרונות בהם.

לעתים נהוג לעדכן את הפרמטר ההסתברותי כך שבאיטרציות הראשונות של האלגוריתם יבוצעו שינויים גדולים בתדירות גבוהה, ואט אט, להוריד את תדירות ועוצמת השינוי בהדרגה בין איטרציה לאיטרציה. עדכון שכזה מעודד סריקה רוחבית בתחילת החיפוש כדי לכסות כמה שיותר שטח, ובהמשך לאחר שכבר יש תמונת מצב כללית מצמצמים את טווח החיפוש לאזורים שנראים כיותר אטרקטיביים. ככל שהחיפוש מתקדם, תדירות ורמת השינוי מופחתים, והחיפוש בהדרגה הופך להיות חיפוש לעומק עד אשר הוא מתבסס על המועמדים הטובים ביותר בסביבה המקומית שלו ומביא להתכנסות האלגוריתם.

למשל, אם p הוא פרמטר ההסתברות לביצוע המוטציה ונניח שבהתחלה $p = 0.75$, נוכל להגדיר בין כל זוג איטרציות עוקבות עדכון לינארי $p \leftarrow (p - \lambda)$, כאשר λ הוא פרמטר שקובע עד כמה הדרגתית תהיה ההפחתה של תדירות השינויים.

2.3.4 שיערוך מידת הכשירות

רמת כשירות

בשלב זה אוכלוסיית המועמדים עוברת שלב שיערוך שמסווג ומדרג אותם לרמות כשירות (רמות תאימות). רמת הכשירות שפתרון מועמד מקבל בשלב השערוך בתום כל מחזור של התהליך אבולוציוני בריצת האלגוריתם הגנטי מייצגת עד כמה הפתרון "טוב" או "איכותי": עבור בעיות מקסימום הפתרונות האופטימאליים הם אלו שסווגו עם רמת הכשירות הגבוהה ביותר, ובאופן אנלוגי, עבור בעיות מינימום, הפתרונות האופטימאליים הם אלו בעלי הכשירות הנמוכה ביותר. במידת הצורך ניתן להפוך את היוצרות ע"י מכפלה במינוס אחת: $-M \leq 0 \leq m \Rightarrow -m \leq 0 \leq M$. סיווג זה לרמת כשירות משמש כשלב הכנה לשלב הבא – שלב הבחירה שמבצע תהליך של סלקציה תוך התחשבות ברמת הכשירות של המועמדים.

הערכות אוטומטיות והערכות אינטראקטיביות

ניתן לחלק בצורה גסה את אופן מימוש השערוך לשתי קטגוריות: הערכות אינטראקטיביות והערכות אוטומטיות [4] – הערכות אוטומטיות, כפי ששמן מרמז, מבצעות את השערוך בעצמם באופן אוטומטי ע"י איזושהי **פונקציית כשירות** (fitness function) מוגדרת מראש. בהערכות אינטראקטיביות לעומת זאת, מבוצע מעין "מיקור חוץ" של משימת השערוך אל גורם אנושי, שאחראי לבקר בעצמו באופן יזום את הפתרונות המועמדים ולספק לאלגוריתם פידבק באשר למידת הכשירות שלהם. כלומר, בשיטת הערכה זו האלגוריתם הגנטי נמצא באינטראקציה שוטפת עם המשתמש המפעיל בכדי לקבל ממנו את ערכי הכשירות המעודכנים של המועמדים בין איטרציה לאיטרציה. בהמשך הסמינר נתמקד בשערוך אוטומטי דווקא ע"י מימוש פונקציית כשירות.

התמודדות עם יעדים מנוגדים

אחד האתגרים המרכזיים העומד בפני מימוש פונקציית כשירות עבור שיטות הערכה אוטומטיות הוא כאשר טיב איכות הפתרונות המועמדים נמדד לפי מספר רב של קריטריונים/יעדים, אשר חלקם בעלי אינטרסים מנוגדים זה לזה ויוצרים מצב של שקלול תמורות (tradeoff). דוגמה קלאסית לשני יעדים מנוגדים שכאלה במדעי המחשב הם מקסום רמת הביצועים של מערכת מחד לעומת מזעור העלויות שלה מאידך [5] – מקסום ביצועים כרוך לרוב בשדרוג חומרה (תוספת שרתי עיבוד, ליבות, מעבדים חזקים/מהירים יותר), זמן ארוך יותר לתכנון ובדיקה ע"י תוספת שעות עבודה נוספות וכי, וכל אלה באים לידי ביטוי בתוספת לעלויות, מה שפוגם בציון המדד השני.

דרך אחת פשוטה ופופולרית להתמודד עם סוגיה זו, היא ע"י הערכה משוקללת: בשלב ראשון כל יעד/קריטריון יימדד באופן עצמאי ללא תלות ביתר בקריטריונים, ע"י פונקציית כשירות ייעודית לו. בשלב שני מקצים לכל יעד משקל יחסי ביחס למכלול היעדים, שקובע עד כמה מדד זה דומיננטי/קריטי ביחס ליתר. לבסוף מבצעים סכום ממושקל של ציוני פונקציות הכשירות השונות, בהתאם למשקל היחסי שנקבע לכל אחת.

דרך אלטרנטיבית, מחוכמת יותר, ובהתאם גם מורכבת יותר, מנסה לזהות את אוסף הפתרונות שיביאו ליעילות פארטו (Pareto efficiency), מעין נקודות איזון המטיבות במידת האפשר את כלל היעדים, ע"י הצלבות בניהם וחיפוש אלו הנמצאים על חזית קו התפר שבין הנקודות האופטימאליות [4,5,6]. לשיטה זו יש פוטנציאל רב יותר מהראשונה [6]: לעיתים מתקבל אוסף של פתרונות אופטימאליים בעלי אותה רמת כושר פחות או יותר, ועל מקבלי ההחלטות לבחון

אוסף זה ולהחליט איזה מהפתרונות יש לבחור בסופו של דבר. בבעיות מעשיות, ההבדל בין בחירת פתרון אחד לרעהו יכולה לבוא לידי ביטוי בשינויים בארגון, שינויים בתכנון לוגי, שינויים ברכישות וכיו"ב. בכדי לקבל את ההחלטה הטובה ביותר במצב כזה, ציון הכושר לבדו אינו מספיק (שכן יש כמה פתרונות עם אותו ציון). מה שכן צריך זה פירוט של אופן הרכבת הציון, כלומר כיצד המדדים השונים המרכיבים אותו מתפלגים בחלוקה הפנימית ומה הציונים היחסיים שלהם בכל אחד מן הפתרונות, שבסופו של דבר מובילים לאותו ציון משוקלל. הסכימה המשוקללת של השיטה הראשונה מניבה ציון סופי שלא מאפשר למקבלי ההחלטות לגזור איזה מהמדדים מתקיים בו ובאיזו מידה. השיטה השנייה לעומת זאת, מרכיבה אוסף של נקודות (פתרונות) הפרוש על פני מה שנקרא "חזית פארטו" (pareto front) של גרף האינטגרציה של פונקציות היעד השונות, מה שמאפשר למקבלי ההחלטות להסיק תובנות באשר למידת קיום יעדי המשנה בכל אחת מן הנקודות ולקבל החלטה מושכלת בהתאם.

עם זאת, על אף היתרון הפוטנציאלי של שיטה זו, לא נדון בה במסגרת סמינר זה, שכן רמת המורכבות שלה דורשת התייחסות מעמיקה החורגת מתחום נושא הסמינר.

צוואר בקבוק האלגוריתם

מכל השלבים של האלגוריתם הגנטי, לרוב שלב השערוך הוא היקר ביותר מבחינת הביצועים, שכן פונקציית הכשירות נוטות לכלול חישובים מורכבים, ולהבדיל מאופרטורים אחרים כגון שחלוף ומוטציה המבוצעים רק על חלק מהמועמדים, שלב ההערכה אינו פוסח על אף מועמד, שכן כולם צריכים להיות מסווגים ע"י פונקציות הכשירות לצורך שלב הבחירה. מסיבות אלו, סיבוכיות זמן הריצה של אלגוריתם גנטי נמדדת לרוב כפונקציה של כמות ההערכות הדרושות עד להתכנסות/עצירת האלגוריתם [5].

ניתן לבצע אופטימיזציה מסוימת ע"י הימנעות מהערכה חוזרת למועמד ששרד מדור קודם ולא השתנה בדור נוכחי, ושימוש במטמון לדירוג חלקי פתרונות: אם הייצוג מודולארי דיו ויש חלקי כרומוזום שנמדדים באופן אינדיבידואלי, ייתכן וסביר שאותם חלקי פתרונות (תמורות של סיביות/ערכים רציפים/דיסקרטיים) יחזרו על עצמם לאורך האלגוריתם. שמירת דירוגים עבור חלקי פתרונות שכאלה בטבלת גיבוב (למשל) יכולה לחסוך חישובים נשנים ע"י אחזור הציון המחושב מהטבלה בעלות נמוכה יחסית בהשוואה לשערוך מחדש.

2.3.5 בחירת מועמדים

עקרון הברירה הטבעית

כאמור שלב זה באלגוריתם מדמה את עיקרון 'החזק שורד' ('survival of the fittest') מתאוריית 'הברירה הטבעית' של דארווין [5]. לפי תאוריה זו קיימת תחרות קיומית תמידית בין המועמדים, ואלו "החזקים" יותר, שההרכב הגנטי שלהם כשיר יותר, הם בעלי סיכויים גדולים יותר לשרוד, והם מעבירים את אותם גנים חזקים שסייעו להם לשרוד לצאצאים שלהם.

בכדי לממש עקרון זה, האלגוריתם סורק את אוכלוסיית הפתרונות הנוכחית, ומבצע תהליך של סלקציה על בסיס הדירוג שניתן לכל מועמד בשלב השערוך. מועמדים שנבחרים ממשיכים לדור הבא (דהיינו לאיטרציה הבאה של האלגוריתם), וכל יתר המועמדים נגרעים מהאוכלוסייה ו-"נכחדים".

דטרמיניזם מול אי-דטרמיניזם

ישנן שתי אסטרטגיות לאופן בחירת המועמדים שייסדו את הדור הבא: האחת, דטרמיניסטית, הידועה גם כאסטרטגיה אליטיסטית, והשנייה, אי-דטרמיניסטית (ובהתאמה אי-אליטיסטית). בגישה הדטרמיניסטית תמיד נבחרים מועמדים מהטובים ביותר, ואילו בגישה האי-דטרמיניסטית, עושים שימוש באקראיות שמקנה צ'אנס גם למועמדים פחות טובים להיבחר.

טבעי לתהות מדוע בכלל קיימת הגישה האי-דטרמיניסטית, שהרי אם שלב זה אמור לשקף את הברירה הטבעית, לפיה 'החזק שורד', בכדי להבטיח שלאורך זמן רמת הכשירות של הפתרונות תלך ותעלה בהדרגה, אזי מדוע להסתכן באפשרות שמועמדים פחות טובים ייבחרו וייתפסו את מקומם של פתרונות הטובים מהם, כאשר אנחנו יכולים באופן דטרמיניסטי פשוט לבחור את אלו היותר טובים?

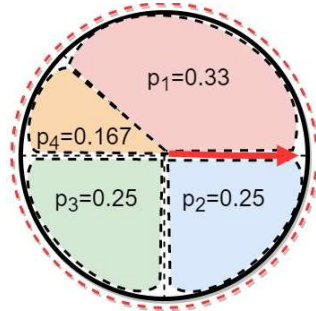
סיבה אחת טובה לכך [5], היא שהשארית פתרונות טובים מאחור ובחירה גם בפתרונות אלטרנטיביים פחות מוצלחים משמשת כאמצעי מניעה נוסף להיתקעות על נקודות קיצון מקומיות: אילו בכל דור ודור ייבחרו אך ורק המועמדים הטובים ביותר ביחס לאוכלוסייה

הנתונה באותה עת, ניסיונות להתרחק מהאופטימום המקומי ע"י שחלופים ומוטציות ייכשלו אם הם לא יספיקו למצוא נקודת שיא חדשה טרום שלב הסלקציה. שילוב מועמדים מהטובים ביותר בבחירה תוך קורץ של אקראיות יכול להוות ניסיון לשימור איזשהו סף של איכות תוך חתירה מתמדת במקביל לבדיקת כיוונים נוספים העשויים להעלות סף זה. נציג להלן כמה דוגמאות למימושים פופולריים של שיטות בחירה בכל אחת משתי הגישות [5].

דוגמאות לשיטות דטרמיניסטיות

בשיטת הבחירה האליטיסטית, נבחרת כאמור קבוצה חלקית של המועמדים הטובים ביותר –

- **בבחירת פסיק** (comma selection) אלו הם μ הצאצאים הטובים ביותר מתוך כלל λ הצאצאים שנוצרו בדור הנוכחי ($\lambda \geq \mu$), כאשר μ הוא פרמטר הנקבע כפונקציה של λ .
- **בבחירת פלוס** (plus selection) לעומת זאת, נבחרים הן μ הצאצאים הטובים ביותר מתוך כלל λ הצאצאים שנוצרו בדור הנוכחי, והן μ ההורים שלהם, כך שהדור הבא יהיה מורכב למעשה מאוכלוסייה המערבת שני דורות.



מבין שתי שיטות אליטיסטיות אלו, בחירת פסיק מתאימה יותר עבור חיפוש לרוחב, שכן בכל איטרציה מחליפים לחלוטין את האוכלוסייה הקיימת עם דור חדש של מועמדים, בעוד שבחירת הפלוס מתאימה יותר לחיפוש לעומק, שכן היא דואגת לשמר חלק יחסי של מועמדים הורים מהדור הקודם וע"י כך דואגת שהחיפוש ימשיך בקרבת מקום סביב האוכלוסייה הנוכחית.

דוגמאות לשיטות אי-דטרמיניסטיות

קצת לשיטות העושות שימוש באקראיות –

- בשיטת "רולטת מזל" (Roulette wheel) מבוצעת בחירת מועמדים תחת התפלגות אחידה, כאשר לכל מועמד מוקצית הסתברות יחסית בהתאם לרמת הכשירות שלו ביחס לכלל.

בשל סיבה זו, שיטה זו ידועה גם בשם **בחירת הכשירות היחסית** (fitness proportional selection). אם למשל לאחר שלב השערוך של אוכלוסייה בת ארבעה מועמדים: c_1, c_2, c_3, c_4 , ערכי פונקציית הכשירות הם כדלהלן: $f(c_1) = 60, f(c_2) = 45, f(c_3) = 45, f(c_4) = 30$, נוכל לבצע נרמול לרמת הכשירות בהתאם לחלקה היחסי מתוך סכום הערכי הכשירות של כלל האוכלוסייה. מתקיים:

$$\sum_{i=1}^4 f(i) = f(c_1) + f(c_2) + f(c_3) + f(c_4) = 60 + 45 + 45 + 30 = 180$$

ומכאן, שהסתברות p_i לבחירת מועמד c_i ($1 \leq i \leq 4$) נתונה ע"י:

$$p_i = \frac{f(c_i)}{\sum_{j=1}^4 f(c_j)} = \frac{f(c_i)}{180}$$

נציב את ערכי פונקציות הכשירות של המועמדים השונים ונקבל (ראה **איור 3**) –

$$P = \{p_1, p_2, p_3, p_4\} = \left\{ \frac{60}{180}, \frac{45}{180}, \frac{45}{180}, \frac{30}{180} \right\} = \left\{ \frac{1}{3}, \frac{1}{4}, \frac{1}{4}, \frac{1}{6} \right\}$$

באופן זה אמנם למועמדים הטובים ניתנים הסיכויים הטובים ביותר לשרוד, בהתאם לעקרון הברירה הטבעית, אך להבדיל מהשיטות הדטרמיניסטיות, בשיטה זו נותנים סיכוי גם למועמדים חלשים יותר להיבחר.

- שיטה אי-דטרמיניסטית נוספת הנותנת משקל כבד יותר לאקראיות היא שיטת **בחירת טורניר** (tournament selection): בשיטה זו, נבחרים תחילה באקראי k מועמדים כלשהם, ללא תלות במד הכשירות שקיבלו, ומתוך המועמדים שנבחרו באקראי מבצעים בחירה של המועמדים הטובים ביותר (למשל ע"י בחירת אליטיסטית או אפילו ע"י שימוש באקראיות נוספת וביצוע בחירה שנייה בשיטת הרולטה. בדומה לשיטת הרולטה, גם שיטה זו מבטיחה לכל אחד המועמדים איזשהו סיכוי הישרדות.

2.3.6 תנאי עצירה

כאמור, מרחבי החיפוש שאנו דנים בהם גדולים מידי מכדי שהאלגוריתם יוכל לסרוק אותם במלואם ולקבוע בחריצות כי הפתרון האופטימאלי נמצא. לפיכך, על האלגוריתם לנתח את תוצאות הביניים המתקבלות בכל איטרציה ולקבל החלטה מושכלת האם כבר הגיעה העת לעצור.

נציג שני תנאי עצירה נפוצים [5]: הראשון הוא הגדרת חסם עליון על מספר המחזורים האבולוציוניים (דהיינו על מספר האיטרציות של הלולאה המרכזית באלגוריתם). עצירה שכזו יכולה לייצג למשל את מיצוי משאבי זמן העיבוד שהוקצו לטובת הריצה. תנאי עצירה חלופי הוא זיהוי סממנים המרמזים על התכנסות קצב השינוי בין איטרציות עוקבות. נח לאתר סממנים שכאלה ע"י אימוץ רעיונות מקריטריון קושי להתכנסות סדרות: נסמן ב- f_i את רמת הכשירות הטובה ביותר המתקבלת באיטרציה ה- i של האלגוריתם, וננסה לחפש איזושהי איטרציה $N \in \mathbb{N}$ כך שלכל שתי איטרציות מאוחרות ממנה, $m, n > N$, מתקיים $|f_m - f_n| < \varepsilon$, כאשר ε הוא איזושהו קבוע חיובי ממשי קטן כרצוננו. אמנם, כמובן שלא נוכל לבדוק שתנאי זה מתקיים לכל $\varepsilon \in \mathbb{R}$ ולכל $m, n \in \mathbb{N}$, אך נוכל לאמוד זאת בצורה גסה ולקבל השערה מושכלת על מגמת התכנסות. למשל אם נקבע את ε מראש לרמת דיוק טובה כרצוננו, למשל: $\varepsilon = 10^{-3} = 0.001$, נוכל להחליט שאם במשך מספר מסוים של איטרציות עוקבות i, j מתקיים $|f_i - f_j| < \varepsilon$, כאשר $|i - j| = 1$, אזי נוכל להסיק שככל הנראה החיפוש כבר לא יתקדם מכאן, ואין כל כך טעם להמשיך באיטרציות נוספות.

ניתן כמובן גם לשלב את שני תנאי העצירה שלעיל בתנאי יחיד עם קשר או אינקלוסיבי (Inclusive or) ולעצור את האלגוריתם אם לפחות אחד משני התנאים מתקיים: ריצת האלגוריתם מיצתה את מספר האיטרציות המקסימאלי שהוקצו לה או שקצב התקדמות שיפור/שינוי הפתרונות מרמז על כך שהוא שואף לאפס ואנחנו נמצאים בסביבת ε קטנה מספיק.

3 חלק II: יישום א"ג בהלחנת מנגינות

בחלק זה נראה כיצד ניתן להשתמש באלגוריתמים גנטיים שהוצגו בחלק הראשון, כאמצעי עזר למשימת ההלחנה של מנגינות וסולואים ליצירות מוסיקליות. חלק זה כולל מושגים ומונחים מתאוריית המוסיקה המערבית (כגון אקורד וסולם) וכן מציג דוגמאות תווים. לחסרי רקע בתחום זה מומלץ לעבור תחילה על [הנספח הנלווה](#) המופיע בסוף הסמינר המציג רקע תאורטי של המוסיקה המערבית.

3.1 הלחנת סולו כמרחב חיפוש

הלחנה היא פרויקט המורכב מהרבה תתי-משימות: קביעת משקל קצבי, חיבור מהלכים הרמוניים, הובלת קולות, קביעת תבנית קצבית לאקורדים, הלחנת מנגינה כנגד המהלכים ההרמוניים ועוד. במסגרת סמינר זה נתמקד אך ורק במשימת ההלחנה של מנגינות (מלודיות) על בסיס איזושהי רקע הרמוני נתון. משימה זו שקולה לחיבור סולו מאולתר על גבי איזושהי סטנדרט ג'אז או שיר פופ/רוק שבו מוקצת סקציה ייעודית לניגון סולו ע"י אחד הנגנים.

היצירה המוסיקלית שעבורה יש לחבר את מנגינת הסולו, מתווה איזושהי מסגרת עבודה לסולן, שצריך להתאים את הסולו שהוא מחבר לקונטקסט של היצירה הנתונה, כך שצבעי הסולו ישתלבו בנוף היצירה. זה יכול לבוא לידי ביטוי בשלל היבטים: בסאונד הכללי והאווירה שמושגת ממנו, התחושה הקצבית, רמת עומק ההבעה, מוטיבים מהיצירה שחוזרים על עצמם בסולו, ועוד. הרבה מהיבטים אלו, באים לידי ביטוי רק בשלב הביצוע, שהרי אותה מנגינה בדיוק יכולה לקבל אינטרפרטציות שונות ע"י נגנים שונים כך שתישמע קצת אחרת ע"י כל אחד מהם. דקויות אלו שמשפיעות על האווירה הכללית של הסולו, כגון דינאמיקה, קישוטים, סאונד וטכניקות ביצוע של כלי נגינה ספציפיים, כל אלו כבודם במקומם מונח הוא, אך אינם נכללים בסמינר זה. סמינר זה מתמקד במשימה של התאמת מנגינת סולו ליצירה תחת הקונטקסט של הרקע ההרמוני שלה, כלומר המיקוד הוא במנגינה עצמה ביחס לרקע הרמוני נתון, ולא סאונד הכללי או אינטרפרטציה שבאים לידי ביטוי רק בשלב הביצוע.

הרקע ההרמוני מגדיר איזושהי סט של סולמות (סדרות של רצפי צלילים), הנגזרים מתאוריית המוסיקה המערבית, כך שבאופן כללי, מובטח לנו שאם נבחר תחת אקורד נתון רק צלילים השייכים לסולם הנגזר ממנו, צלילים אלו יישמעו "סביר" תחת האקורד ללא "זיוף" במובן שהצלילים לא יישמעו צורמים באופן שעלול להתפרש ע"י מאזינים רבים כטעות.

מפרספקטיבה זו של מיקוד משימת הלחנת הסולו לבחירת צלילים מתוך אוסף מוגדר היטב (סולמות) שנגזר מהמהלך ההרמוני של היצירה, נוכל לראות את משימת הלחנת מנגינת הסולו כבעיית אופטימיזציה. מנקודת מבט זו, מרחב החיפוש הוא כלל המנגינות האפשריות שניתן לחבר ליצירה, והאילוצים החלים על האינדיבידואלים במרחב החיפוש הם בין היתר – הצטמצמות רק לצלילים השייכים לאוסף הסולמות שנגזרים מהמהלך ההרמוני ועמידה במשקל הקצבי של היצירה. יתר על כן, כדי למקד את הדיון סביב העיקר, שהוא שימוש באלגוריתמים גנטיים לחיבור מנגינות סולו, נתמקד במודל מאוד מוגבל עם אילוצים נוספים: נצטמצם להלחנת מנגינות חד-קוליות בלבד, עם משקל קצבי קבוע של ארבעה רבעים לתיבה, ונניח שיש חסם תחתון כלשהו על הערך הריטמי (משך שהייה יחסי) הקצר ביותר של כל צליל במנגינה. אילוצים אלו מצמצמים משמעותית את מרחב החיפוש ומאפשרים לדון במודל פשוט יותר, שבו מרחב החיפוש הוא פשוט אוסף של סדרות צלילים עם הערך הריטמי שלהם.

עם זאת, אפילו תחת המודל הכללי המצומצם שבתחום הדיון, מרחב החיפוש גדול מידי מכדי לבצע בו חיפוש ממצה. אם נניח למשל שאורך הסולו הוא 8 תיבות, ושהערך הריטמי הקצר ביותר המותר לכל צליל בתיבה הוא $\frac{1}{16}$ ממנה, אזי אפילו תחת ההנחה המקלה שכל 8 התיבות שייכות לאותו סולם (כאשר סולם דיאטוני סטנדרטי מכיל 7 צלילים שונים), ואפילו אם נניח שמנעד טווח הצלילים של מנגינת הסולו יחסית מצומצם ומוגבל עד לשתי אוקטבות בלבד, ושכל רגע נתון חייבים לנגן צליל ללא הפסקות באמצע, נקבל שבסה"כ, מספר האפשרויות הכולל להרכיב מנגינת סולו ל-8 תיבות נתון ע"י –

$$\#bars \cdot (\#possible_notes \cdot \#octaves)^{\#shortest_rhythmic_value} = 8 \cdot (7 \cdot 2)^{16} = 8 \cdot 14^{16} = 1.74236267 \cdot 10^{19}$$

נחשב ונקבל – $1.74236267 \cdot 10,000,000,000,000,000 = 17,423,626,700,000,000,000$ – מנגינות סולו שונות אפשריות, וזאת עם כל ההנחות המקלות, שבלעדיהן המספר גדול עוד בהרבה. לפיכך, חיפוש ממצה הכולל בדיקה של כל המנגינות האפשריות אינו ישים, ולפיכך שיטות חיפוש מטה-היוריסטיות כמו אלגוריתמים גנטיים יכולות להציל כאן את המצב, כאשר הפתרונות האופטימאליים של מרחב החיפוש הן המנגינות שנשמעות הכי "טוב", ועל כך נרחיב בהמשך בפרק הערכת הכשירות. ראשית כל נראה כיצד בכלל אפשר לייצג ולקודד את הבעיה של הלחנת מנגינת סולו תחת רקע הרמוני נתון, במונחים של אלגוריתמים גנטיים.

3.2 ייצוג מנגינה כגנום

בכדי שנוכל להשתמש באלגוריתמים גנטיים לחיבור מנגינות (ובפרט סולואים), עלינו להגדיר תחילה ייצוג נאות למנגינה במושגים של אלגוריתמים גנטיים, דהיינו כיצד לקודד אותה כגנום (או גנוטיפ). כפי שצוין, ייצוג נאות הוא בעל משמעות קריטית, שכן אופן הייצוג משפיע בין היתר גם על אופן מימוש האופרטורים הגנטיים (שחלוף, מוטציה וכד'), גם על ביצועי האלגוריתם, וגם על הפנוטיפ, דהיינו על התוצר הסופי של המנגינה המפוענחת מתוך הקידוד של הגנוטיפ. כיצד איפה נייצג ונקודד מנגינה?

3.2.1 ייצוג אבסטרקטי של מנגינה

באופן עקרוני, ניתן לאפיין לחן של מנגינה כלשהי ע"י רצף תווים, שכן הם מגדירים באופן חד-משמעי את גבהי הצלילים של הלחן יחד עם הערך הריתמי שלהם (משך השהייה שלהם), לרבות הפסקות (שתיקות) שבין הצלילים [4]. לפיכך, במסגרת המודל הפשוט הנידון שבו אנו מתעלמים מ-"קישוטים מינוריים שמסביב" דוגמת דינאמיקה ותחושת קצב כללית, ייצוג של מנגינה ע"י רצף של גבהי הצלילים והפסקות המרכיבים אותה, יחד עם משך השהייה של כל צליל והפסקה, נראה כייצוג נאות המאפשר לקודד ולפענח כל מנגינה שהיא.

עם זאת, נעיר כי ניתן לטעון (ובצדק) שהתאמה זו בין רצף תווים ללחן של מנגינה אינה התאמה חד-חד ערכית, שכן אופן הרישום בתווים הוא בעצמו אינו חד-ערכי. למשל –

- ניתן להאריך את משך השהייה של ערך ריתמי ע"י סימני קשתות חיבור וניקוד ולקבל ערך ריתמי ארוך יותר, שכבר מוגדר עבורו סימון ייעודי. כתוצאה מכך יוצא שיש כמה דרכים שונות לייצוג אותו משך שהייה כולל – ע"י הערך הריתמי הייעודי, או ע"י חיבור ערכים ריתמיים קצרים יותר שסכומם הוא זה שמיוצג ע"י הערך הייעודי.
- גובה צליל אבסולוטי יכול להיות מיוצג ע"י תווים שונים, למשל גובה הצליל הנמצא באמצע בין דו (C) ורה (D) במרחק של חצי טון מכל אחד מהם באוקטבה כלשהי, יכול להירשם הן כדו-דיאז (C#) והן כרה-במול (Db).

להמחשה, שתי התיבות שלהלן אמנם רשומות באופן שונה, אך נשמעות בדיוק אותו הדבר –

C Bb A C A# A, the 2 quarter-notes are connected to a half note

לטעמי, מבחינת המאזין, מה שחשוב בסופו של דבר הוא הפנוטיפ, התוצר הסופי, דהיינו המוסיקה המתנגנת לו באוזן. התווים הם רק צורת רישום לייצוג היצירה. צורה זו שקופה לחלוטין למאזין, שכן כל דרכי הרישום השונות השקולות מייצגות בסופו של דבר את אותו רצף גלי קול שמהווים את המוסיקה שבה אנחנו חפצים, על כן למרות אי-האחידות בצורות רישום של תווים זוהי מגבלה מלאכותית בלבד. גבהי הצלילים האבסולוטיים בסולם הכרומוטי וכן זמני משכי השהייה שלהם, למשל במדידה לפי שניות, מתלכדים עבור כל צורות הרישום השקולות. לאור האמור לעיל, ניצמד לרעיון הייצוג של מנגינה ע"י רצף סדור של הצלילים וההפסקות שהיא מכילה (לפי הסדר) עם הערך הריתמי שלהם (משך השהייה שלהם).

3.2.2 קידוד המנגינה

הייצוג האבסטרקטי של מנגינה ע"י רצף תווים (דהיינו רצף גבהי הצלילים של המנגינה עם הערך הריתמי שלהם) הוא ייצוג מקובל ונפוץ במחקרים העוסקים בנושא. לעומת זאת, אופן הקידוד הקונקרטי של ייצוג מופשט זה כגנוטיפ יכול להיעשות בשיטות שונות לחלוטין. נציג שתי דוגמאות קידוד המייצגות שיטות פופולריות טיפוסיות.

דוגמא 1: קידוד גובה צליל אבסולוטי, תוך כדי קידוד גובה צליל לחוד וערך ריתמי לחוד

מימוש טבעי לאופן הייצוג שלעיל לרצף תווים הינו הגדרת המבנה האבסטרקטי של הכרומוזום כווקטור של זוגות סדורים מהצורה (pitch, duration), כאשר הרכיב הראשון (pitch) מקודד את הגובה האבסולוטי של הצליל, והרכיב השני (duration) את ערכו הריתמי. דוגמת הקידוד

הראשונה שנציג מהווה מימוש ברוח גישה זו [4], אולם במקום להגדיר את הכרומזום כסדרה של עצמים בעלי שני מאפיינים (שדות) – אחד עבור הצליל והשני עבור הערך הריתמי, בדוגמה זו כל כרומזום מורכב מ-2 וקטורים נפרדים, מאותו אורך, הראשון מכיל את גבהי הצלילים והשני את הערכים הריתמיים שלהם. גובהי הצליל מקודדים לערכים בתחום [0,127], הממופה לגבהים האבסולוטיים ע"פ התאמה הקיימת ב-MIDI³ – למשל, הקליד הלבן השמאלי ביותר (בעל הגובה הנמוך ביותר) במקלדת פסנתר סטנדרטית בעלת 88 קלידים המייצג את הצליל A0 (לה באוקטבה הקודמת לאוקטבה הראשונה) מקודד ע"י הערך 21. הצליל Bb0 המיוצג ע"י הקליד השחור העוקב לו, מקודד ע"י הערך 22, וכך הלאה עד הקליד הימני ביותר במקלדת (זה שמפיק את הצליל הגבוה ביותר) – המייצג את הצליל C8 (דו באוקטבה השמינית), המקודד ע"י הערך 108 (הערכים "החסרים" שבקצוות התחום [0,20] ו-[109,127] מייצגים צלילים כה נמוכים וגבוהים בהתאמה, עד כי כמעט ולא ניתן לזהות אותם ע"י אוזן "ממוצעת", והם כמעט ולא נמצאים בשימוש במוזיקה הפופולרית המערבית).

בכדי לקודד הפסקה במקום גובה צליל, משתמשים בערך השלילי -1, הנמצא מחוץ לתחום הערכים שמשמש את ערכי הצלילים הממשיים.

עבור הערכים הריתמיים, הקידוד יכול להיעשות באופן אבסולוטי, למשל - 0.25 עבור רבע פעימה, או לחלופין באופן רלטיבי ביחס לאיזושהו ערך ריתמי שמוגדר מראש בתור הערך הבסיסי לייחוס. בדוגמה זו למשל, הערך הריתמי של רבע מקודד ע"י 1.0, וכל יתר הערכים מקודדים ביחס אליו: למשל, ערך ריתמי של חצי המהווה שני רבעים, מקודד ע"י הערך 2.0, ושמינית המהווה חצי מרבע מקודדת ע"י הערך 0.5.

להלן ארבע התיבות הראשונות של 'ארץ' של אילנית, יחד עם קידוד שני הווקטורים של הכרומזום התואם (הקידוד בשורה העליונה מציין את גובה הצליל של כל תו וסימוני הפסקות, וזה בשורה התחתונה את הערך הריתמי של כל צליל ו/או הפסקה) –

	Cm	Cm/Bb	Ab	Fm	G	G/F	Cm	Dm7/5-G
pitch:	< 67,63, 63,60,60, 55, -1,	63,63, 62, 60,56, 58, 56	55, 67,67,65,63, 63,62,	63, -1, -1 >				
rhythm:	< 0.5,0.5,0.5,0.5,0.5,1.0, 0.5,	0.5,0.5,0.75,0.25,1.0, 0.5,0.5,	0.5,0.5,0.5,0.5,1.0, 0.5,0.5,	2.0, 1.0, 1.0 >				

דוגמא 2: קידוד גובה צליל רלטיבי, תוך כדי קידוד משולב של גובה הצליל יחד עם ערכו הריתמי

נציג כעת גישה שונה לחלוטין למימוש הקידוד, שבה משתמש Biles [1] במערכת התוכנה שפיתח – GenJam. מערכת זו מחברת סולואי ג'אז מאולתרים בזמן חי בהופעות תוך כדי התאמת תוכן

Chord	Scale	Notes
Cmaj7	Major (avoid 4th)	C D E G A B
C7	Mixolydian (avoid 4th)	C D E G A Bb
Cm7	Minor (avoid 6th)	C D Eb F G Bb
Cm7b5	Locrian (avoid 2nd)	C Eb F Gb Ab Bb
Cdim	W/H Diminished	C D Eb F Gb G# A B
C+	Lydian Augmented	C D E F# G# A B
C7#5	Whole Tone	C D E F# G# Bb
C7#11	Lydian Dominant	C D E F# G A Bb
C7alt	Altered Scale	C Db D# E Gb G# Bb
C7#9	Mix. #2 (avoid 4th)	C Eb E G A Bb
C7b9	Harm Minor V (no 6th)	C Db E F G Bb
CmMaj7	Harmonic Minor	C D Eb F G Ab B
Cm6	Dorian (avoid 7th)	C D Eb F G A
Cm7b9	Melodic Minor II	C Db Eb F G A Bb
Cmaj7#11	Lydian	C D E F# G A B
C7sus	Mixolydian	C D E F G A Bb

טבלה 1: הגדרות מיפוי סולם (אוסף צלילים) ע"פ סוג אקורד
כפי שממומשות במערכת GenJam של Biles [1]

הסולו לאלתורים המבוצעים ע"י נגנים אחרים. שיטת קידוד זו נבדלת משמעותית מזו שהוצגה הדוגמה הקודמת בשני היבטים מרכזיים:

א. קידוד גובה הצליל אינו אבסולוטי, אלא רלטיבי ביחס לסולם נתון.

ב. הקידוד משתמש בשדה משותף לקידוד הן עבור גובה הצליל והן עבור הערך הריתמי התואם לו.

נבהיר מה המוטיבציה שעומדת מאחורי שיטה כזו וכיצד היא עובדת הלכה למעשה. נתחיל בהסבר קידוד גובה צליל יחסי: בתאוריית המוסיקה המערבית קיימים כללים לניתוח יצירה מוסיקלית על סמך המהלך ההרמוני שלה, דהיינו על סמך רצף האקורדים שלה. ע"פ כללי ניתוח כאלה ניתן לשייך כל מקטע ביצירה

³ MIDI (Musical Instrument Digital Interface) – סטנדרט המגדיר פרוטוקולים להתממשקות דיגיטאלית בין כלי נגינה ומחשב. הסטנדרט מגדיר בין היתר מיפוי של 128 צלילים לתחום [0-127] הכוללים את כל 88 הצלילים המופקים מקלידי פסנתר סטנדרטי ו-40 צלילים נוספים.

(אקורד, תיבה, מספר תיבות,...) לאיזשהו סולם (ולעיתים אף מספר סולמות), כך שמובטח כי צלילים הלוקוחים מתוך סולם זה יישמעו "טוב" או "נכון" (ללא "זיופים") על גבי מקטע זה, ביחס להרמוניה המלווה אותם. במימוש של Biles [1], במקום לתת כל יצירה מחדש ע"פ המהלכים והמעברים ההרמוניים שבה, הוגדרו כללים לניתוח הרמוני מקומי, המתאימים לכל סוג של אקורד סולם קבוע, ללא תלות בהרמוניה שבאה לפני ו/או אחרי האקורד נתון (ראה [טבלה 1](#) להמחשה). הסולם מגדיר את תבנית המרווחים שבין כל שני צלילים בסולם, ואלו חוזרים על עצמם באופן מחזורי באוקטבות שונות. לפיכך, לאחר שמקבעים מהיכן מתחילה הספירה, כלומר מיהו הצליל הראשון בסולם ומהו גובהו האבסולוטי, ניתן להתייחס לכל יתר הצלילים באופן רלטיבי ביחס לצליל זה, כלומר הערך 1 יקודד את הצליל הראשון בסולם, 2 את הצליל השני שבסולם וכן הלאה, כאשר בכל סולם הרכב הצלילים משתנה בהתאם לסוג (מבנה) האקורד, ושורש האקורד. במימוש של GenJam [1], הספירה של כל הסולמות מתחילה מהצליל C4 (דו באוקטבה הרביעית), כך שבכדי לקודד סולם הממופה לאקורד שהשורש שלו שונה מ-C (דו), יש לחשב תחילה מהם צלילי הסולם ע"פ תבנית המרווחים המוגדרת לסולם זה החל מהשורש שלו, ולאחר מכן לבצע היסט כך שהאינדקס של מופע ה-C הראשון בסולם יהפוך להיות אינדקס 1. נדגים: עבור סוג אקורד 7 המושרש בתו מסוים, ע"פ [טבלה 1](#) הסולם שמוגדר מהשורש ועד המופע הבא שלו באוקטבה העוקבת הוא רצף המרווחים כדלקמן – שורש, טון, טון, טון וחצי, טון, חצי טון, שורש וחוזר חלילה. עבור סולם שכזה המושרש למשל ב-F (פה) כנגד אקורד F7, מתקבל הסולם (משמאל לימין): $F_1, G_2, A_3, C_4, D_5, E_6, F_7, G_8, A_9, \dots$, ולאחר היסט מתאים כך שאינדקס 1 יחל מ-C נקבל את הסולם המותאם (משמאל לימין): $C_1, D_2, E_3, F_4, G_5, A_6, C_7, D_8, E_9, F_{10}, \dots$. באופן זה הקידוד בעצם אינו אבסולוטי אלא רלטיבי ביחס לאקורד הנוכחי והסולם הנגזר ממנו, למשל מעל אקורד C7, גובה צליל בעל ערך 3 משמעו E (באוקטבה מתאימה) בעוד שמעל אקורד F7 ערך 3 משמעו דווקא Eb (באוקטבה מתאימה). בכדי לציין הפסקות, גם בשיטת קידוד זו, בדומה לשיטת הקידוד הקודמת משתמשים בערך הנמצא מחוץ לתחום הערכים המייצגים צלילים ממשיים – במקרה הזה משתמשים בערך המיוחד 0 כדי לקודד הפסקה. כעת נסביר כיצד מקודדים את הערכים הריתמיים בשילוב יחד עם גבהי הצלילים בשדה משותף באותו הווקטור: מגדירים מראש מהי היחידה הריתמית הקצרה ביותר המותרת לשימוש במנגינה שתיווצר ע"י האלגוריתם $(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \frac{1}{8}, \dots)$, נסמן את המכנה שלה ב-q. בתיבה אחת יש q יחידות ריתמיות באורך $\frac{1}{q}$, על כן ניתן מראש לחלק כל תיבה ל-q יחידות ריתמיות. ניתן ע"י קשת הארכה לחבר שתיים או יותר יחידות של $\frac{1}{q}$ וע"י כך להפיק יחידות ריתמיות ארוכות יותר (למעשה את כל היחידות $\frac{1}{p}$ כך ש-p מחלק את q. למשל: אם $q=12$, דהיינו היחידה הריתמית הקצרה ביותר המותרת היא $\frac{1}{12}$ מאורך של תיבה אחת, אם נחבר שתי יחידות ריתמיות כאלו ע"י קשת הארכה נקבל יחידת ריתמית עם משך שהייה של $\frac{1}{12} + \frac{1}{12} = \frac{2}{12} = \frac{1}{6}$ תיבה, חיבור שלוש יחידות יניב משך שהייה של $\frac{1}{4}$ תיבה וכן הלאה. הרעיון העומד מאחורי "יחידה ריתמית קצרה ביותר" $\frac{1}{q}$ שכזו, הוא שניתן לחסוך ע"י כך את הצורך בקידוד כל יחידה ריתמית בנפרד, ולהניח שכל עוד לא מצוין אחרת ברירת המחדל של משך שהייה של היחידות הריתמיות הן $\frac{1}{q}$ מאורך התיבה. בשביל לציין יחידה ריתמית ארוכה יותר מ- $\frac{1}{q}$ משתמשים בערך מיוחד שזר לתחום ערכי גבהי הצלילים, שמציין שיש לחבר את יחידה זו בקשת ליחידה שקדמה לה, דהיינו להאריך את משך שהייה של התו הקודם ב- $\frac{1}{q}$. במימוש של GenJam [1] מלכתחילה הגדירו מנעד מצומצם של שתי אוקטבות הפורש לכל היותר ארבע-עשר צלילים במסגרת אקורד/סולם נתון, מה שמותיר את הערך 15 כזמין לקידוד ביצוע פעולת הארכת משך שהייה. כעת, ע"פ שיטה זו, מספיק לקודד את גבהי הצלילים בלבד ללא צורך בקידוד ערכים ריתמיים: כברירת המחדל משך שהייה של כל צליל הוא $\frac{1}{q}$, ויחידות ארוכות יותר מתקבלות ע"י שימוש חוזר בערך המיוחד 15, במקום ציון ערך עבור גובה צליל. כיצד נבחר את q? מצד אחד נרצה q יחסית גבוה שכן ככל שנבחר q בעל מספר רב יותר של מחלקים, נאפשר מגוון רחב יותר של יחידות ריתמיות. מצד שני, ב-"יום-יום" משתמשים לרוב בכמה יחידות ריתמיות בסיסיות פשוטות. בחירה ב-q גבוה מידי עלולה להניב יחידות ריתמיות ביזאריות ואי-סימטריות (למשל $\frac{1}{7}$ או $\frac{1}{19}$) שיגרמו למנגינה להישמע כמו כאוס אקראי מוחלט. פשרה טובה היא בחירת q שמייצג פחות או יותר את רמת הדחיסות (טמפו) הרצויה של צלילים בכל פעימה, וכן בחירת q המהווה חזקה שלמה של 2, שכן אז מכני היחידות הריתמיות הבסיסיות של שלם, חצי, רבע, שמינית וכד' כולן מחלקים את

q , ומנגד, יחידות אי-זוגיות ואי-סימטריות ביזאריות אינן מחלקות את q ולכן לא ייווצרו ע"י האלגוריתם. ב-GenJam [1] נבחר הערך $q = 8$, דהיינו היחידה הריתמית הקצרה ביותר היא שמינית תיבה. בחירה זו מתאימה במיוחד למנגינות בסגנון הג'אז, שעבורו יועדה מערכת זו.

נסכם את שיטת הקידוד של הדוגמה השנייה: ערכי גבהי הצלילים וערכים הריתמיים מקודדים יחד בכרומוזום באותו הווקטור: ערכי ביניים (בדוגמה זו [1,14]) מקודדים גבהי צלילים ע"פ המרחק היחסי שלהם מאיזשהו צליל אבסולוטי (בדוגמה זו C4), כאשר המרחק נמדד בצעדים (או דרגות) בסולם הנקבע ע"פ סוג האקורד המתנגן במקביל לאותו צליל. ערך מיוחד (בדוגמה זו 0) מקודד הפסקה, וערך מיוחד אחר (בדוגמה זו 15) מקודד המשך שהיית התו הקודם באורך של היחידה הריתמית הקצרה ביותר הנמצאת בשימוש שנקבעת מראש (בדוגמה זו שמינית). עם סיום תיאור הדוגמה, נציג שוב את ארבע התיבות הראשונות של 'ארץ' של אילנית, הפעם עם הכרומוזום המתאים כשהוא מקודד לפי השיטה המתאימה לדוגמה השנייה – בהתאם, כלל הצלילים הועלו באוקטבה בכדי להתאים למנעד שמתחיל ב-C, היחידות הריתמיות של הרבעים הומרו לייצוג בשמיניות, והסולמות שהותאמו לאקורדים הן סולם דו-מינור הרמוני עבור התיבה של G7, וסולם דו-מינור טבעי דיאטוני עבור יתר התיבות:

Chromosome data: $\langle 12, 10, 10, 8, 8, 5, 15, 0 \quad 10, 10, 9, 8, 5, 15, 6, 5 \quad 5, 12, 12, 11, 10, 15, 10, 9, \quad 10, 15, 15, 15, 0, 15, 0, 15 \rangle$

השוואה בין שיטות הקידוד

גובה אבסולוטי אל מול גובה רלטיבי

קידוד גבהי צלילים ע"פ ערכם הרלטיבי ביחס לאיזשהו גובה אבסולוטי בהקשר של סולם נתון, מבטיח שהמנגינה שתיווצר ע"י האלגוריתם לעולם לא תכיל "טעויות" מבחינת בחירת הצלילים, שכן מלכתחילה טווח הצלילים האפשרי בשיטה זו הוא טווח של צלילים המסווגים כ-"נכונים" ביחס לאקורד שעליהם הם מתנגנים. עם זאת, ההחלטה להצטמצם מראש לסט צלילים קבוע ביחס לאקורד נתון מגבילה משמעותית את טווח החיפוש במרחב ומונעת מראש כניסה לאזורים מסוימים במרחב שהם אמנם חשוכים למראה ואפופים בערפל אך עשויים לעיתים להתגלות כמפתיעים לטובה: האינטרפרטציה של מה נשמע "טוב" במוסיקה היא סובייקטיבית, ולעיתים באומנות נרצה דווקא "לצאת מהמסגרת" של הכללים התאורטיים ולפעול בחופשיות ללא גבולות על מנת לגוון, ובחיבור מנגינה גיוון שכזה יכול להגיע דווקא מהכיוון של צלילים "חריגים" שמתבלטים בנוף ההרמוני ויוצרים עניין.

מהכיוון המשלים, קידוד גבהי צלילים ע"פ ערכם האבסולוטי פורש ספקטרום רחב יותר של אפשרויות בחירה ומאפשר פוטנציאלית לחולל כל מנגינה אפשרית, אך מנעד רחב זה כולל בחובו גם לא מעט "מוקשים" בדמות של צלילים שעשויים להישמע "רע" כאשר משבצים אותם במקום ובזמן שפחות מתאימים.

קידוד משולב של צלילים וקצב אל מול קידוד נפרד

קידוד משולב של ערכי גובה צליל וערכים ריתמיים באמצעות קידוד מובלע של הערך הריתמי ע"י הגדרת ערך ברירת מחדל של יחידה ריתמית קצרה ביותר, מכיל לא מעט יתרונות: מעבר ליתרונות ה-"טכניים" של קומפקטיות בייצוג והימנעות מהצורך לתחזק שני מבני נתונים נפרדים ולדאוג לסנכרון ביניהם, הייצוג המשולב מפשט מאוד את מלאכת החיפוש במובן שהוא מאפשר להתמקד באספקטים הקשורים למנגינה עצמה, כלומר בצלילים המרכיבים אותה, מבלי להיות טרוד במקביל כל זמן החיפוש כיצד לתחזק ולמצוא את הערכים הריתמיים המתאימים: למשל, אורך כרומוזום המייצג תיבה בודדת הוא סטטי (כערך המכנה של היחידה הריתמית הקצרה ביותר) ולכן לא צריכים להתעסק בכמה תווים לכלול בכל תיבה, וכיצד לחלק ביניהם את הפעילות ולדאוג שהחלוקה תהיה תואמת למשקל. כמו כן, במסגרת האופרטורים הגנטיים, ניתן להתייחס לכל גן (תו) כקופסא שחורה, ולבצע עליו שינויים אקראיים בין אם הוא מייצג ערך בעל משמעות ריתמית (למשל 15 בייצוג של דוגמה 2 שלעיל ו/או 0 להפסקה), ובין אם הוא מייצג גובה צליל. כך או כך, יוצא שבלי מאמץ אנחנו משיגים "2 ציפורים במכה אחת", דהיינו תהליך האופטימיזציה של בחירת הצלילים המתאימים מניב באמצעות קידוד זה גם תבניות ריתמיות חדשות, כתוצר לוואי.

לעומת זאת, קידוד נפרד של הערכים הריתמיים הרבה יותר מורכב, שכן בכל פעולת מוטציה ו/או שחלוף, האלגוריתם צריך לקחת בחשבון את ההשלכות של מגוון תרחישים אפשריים: שינוי של צליל ללא שינוי משך השהייה, שינוי משך השהייה תוך שימור גובה הצליל, שינוי שני הרכיבים גם יחד וכד'.

נציין כי מלאכת הרכבת תבניות ריתמיות איכותיות אינה פשוטה כלל וכלל. צליל שנשמע "חריג" תחת איזשהו נוף הרמוני יכול לקבל משמעות שונה לגמרי אם מתזמנים אותו שבריר פעימה לפני או אחרי כצליל מעבר אל צליל אחר המקבל את הדגש, שמביא לאתנחתא מוגברת לאחר רגע של מתח. ישנם אף מחקרים בנושא הלחנה אלגוריתמית באמצעות אלגוריתמים גנטיים המתמקדים ספציפית בנושא הייעודי הזה של חיבור אופטימאלי של תבניות קצביות [3,4].

עם זאת, על אף המורכבות הקיימת בטיפול נפרד בקידודים ערכים ריתמיים, ההפרדה מאפשרת הרבה יותר גמישות: בייצוג המשולב לצורך העניין, מלכתחילה כל הערכים הריתמיים חסומים מלרע ע"י משך השהייה של "היחידה הריתמית הקצרה ביותר" שהוגדרה מראש לטובת הייצוג המשולב, ולכן אין אפשרות "להגביר מהירות" (בהיבט של דחיסת צלילים במקטע) מעל סף מסוים, ויש לא מעט יחידות ריתמיות שפשוט לא ניתן להשתמש בהן מאחר שהמכנה שלהן לא מחלק את המכנה של היחידה הקצרה ביותר: ב-GenJam [1] למשל, בשל הבחירה בשמינית כמשך השהייה הקצר ביותר, אין אפשרות להפיק מנגינה עם טריליות ($\frac{1}{3}$) או שש-עשריות ($\frac{1}{16}$).

לסיכום ההשוואה, יש פה מצב של tradeoff קלאסי בין מצב נטול סיכונים שהתשואה הפוטנציאלית שלו מוגבלת (קידוד גובה יחסי) לבין מצב חשוף לפגיעה שעשוי למצוא את האוצר החבוי (קידוד גובה אבסולוטי), וכן tradeoff בין מודל פשוט ונח המוגבל ביכולות שלו (קידוד משולב) לבין מודל שאינו מוגבל ביכולותיו אך מורכב ומסובך לתפעול (קידוד מפוצל).

3.3 מימוש אופרטורים גנטיים על כרומוזומים מוסיקליים

בפרק זה נציג דוגמאות למימוש האופרטורים הגנטיים של אתחול, שחלוף ומוטציה על כרומוזומים המייצגים מנגינות, בהתאם לאופני הייצוג והקידוד שנידונו לעיל.

3.3.1 אתחול דור ראשון של מנגינות

כאמור בפרק [אתחול דור ראשון](#), ניתן לחלק את שיטות האתחול לשתי מחלקות גסות מרכזיות – אתחולים אקראיים ואתחולים מכוונים. כפי שראינו קיים tradeoff בין הרצון ליצור גיוון וחידוש, הכרוך בסיכון של ביצועים פחותים וקבלת תוצאות ביזאריות (אתחול אקראי), לעומת הרצון להשיג תוצאה טובה במהירות אך עם תוצר שאינו שונה בהרבה מזה המקורי ולכן אינו מחדש הרבה (אתחול מכוון).

איזה שיטת אתחול מתאימה יותר לחיבור מנגינות?

לאור האמור לעיל, בהקשר של חיבור מנגינה עבור סולו או אלתור על בסיס יצירה קיימת, אני מאמין שבד"כ יתאים יותר דווקא אתחול מכוון, שכן סולואים ואלתורים משלבים באופן טבעי מוטיבים מלודיים מתוך מנגינת הבסיס המקורית של היצירה, לכן אתחול על בסיס גרעין המנגינה הקיימת למשל יכול להוות מקור טוב לווריאציות מעניינות שגם יישמעו מוכרות למאזין מצד אחד, וגם יחדשו משהו מצד שני. לעומת זאת, בהקשר של חיבור מנגינה עבור יצירה חדשה, למשל כאשר יש רק תשתית של איזשהו מהלך הרמוני עם מקצב בסיסי מעניין שרוצים לרקוס לה עור וגידים ע"י חיבור לחן על רצף האקורדים, אתחול אקראי כנראה יותר יתאים, שכן המגוון הרחב של אפשרויות שהוא פורש יכול להוות מקור טוב יותר לחיפוש השראה.

כיצד נאתחל איפה דור ראשון של מנגינות?

עבור אתחול מכוון לטובת חיבור סולו או אלתור על בסיס יצירה קיימת, ניתן להשתמש במנגינה של היצירה הקיימת בתור גרעין, לקודד אותה לכרומוזום ע"פ אחת השיטות המרכזיות שהוצגו [לעיל](#), ולהפעיל על כרומוזום זה סדרת פעולות מוטציה המניבות סדרת כרומוזומים של עותקים "מעודכנים" המהווים וריאציות על מנגינת הגרעין, ועל כרומוזומים אלו ניתן לבצע פעולות שחלוף. באופן זה מתקבל דור ראשון של מנגינות שניתן להחיל עליהן את התהליך האבולוציוני המחזורי.

עבור אתחול אקראי, המימוש משתנה לגמרי בין צורות הקידוד: בקידוד המשולב, אורך הווקטור המייצג את הגנום הוא סטטי, ושווה למכפלת מספר התיבות של היצירה במספר המופעים של

היחידה הריתמית הקצרה ביותר (שנקבעת מראש) בתיבה אחת. לפיכך ניתן פשוט לבצע בחירות אקראיות של ערכים בטווח הקידוד, תוך וויסות רמת הכאוס והדחיסות ע"י פונקציות הסתברות מתאימות לבחירת ערכי ההפסקות והארכת משכי שהייה לעומת צלילים וכן בחירת ערכי צלילים שיחסית סמוכים זה לזה בכדי למנוע קפיצות חריגות. לעומת זאת, עבור קידוד המפריד בין ההפסקות וגבהי הצלילים לבין ערכיהם הריתמיים, האורך של מקטע בכרומוזום המייצג תיבה אינו קבוע, על כן בנוסף לבחירה האקראית של הצלילים עצמם יש גם לבחור בתבנית ריתמית עבורם. כאמור, חיבור תבנית ריתמית טובה הוא נושא מורכב, וקיימים מחקרים לשימוש באלגוריתמים גנטיים ספציפית למלאכה זו של חיבור תבנית ריתמית, אך אנו לא נתעמק בהם כאן, ונתמקד באספקטים הקשורים לצלילים עצמם (גבהי הצלילים) ופחות בערכם הריתמי.

3.3.2 מוטציה של מנגינה

כפי שצוין בפרק על אופרטור המוטציה, אופרטור זה נועד לחקות שינויים אקראיים מהטבע בהרכב הגנטי של הכרומוזום, ומצופה ממימוש כללי לאפשר נגישות, לשמר ניטרליות ולתמוך בסקליליות.

במימוש אלגוריתמים גנטיים עבור חיבור מנגינות, לאור המגוון הרחב של השינויים האפשריים במנגינה, נהוג לפרק ולחלק את המוטציה לאוסף של שגרות, מעין מוטציות-משנה [1,2,7], אשר כל אחת מהן אחראית למימוש איזשהו שינוי. שגרות אלו עשויות לטפל ברמות שונות של מרכיבים במנגינה – שינוי ברמת התו הבודד, שינוי ברמה תיבה שלמה, שינוי של מקטע הכולל אוסף של תיבות וכן הלאה. עם אוסף כזה של מוטציות-משנה, ניתן להגדיר את אופרטור המוטציה ע"י בחירה אקראית של אחת ממוטציות-המשנה בכל הפעלה [2], או לחלופין, לאפשר שינויים ע"י מספר מוטציות-משנה בזה אחר זה בכל הפעלה של האופרטור, ע"י מעבר על אוסף המוטציות השונות והפעלתן באופן מותנה בהתאם לאיזושהי פונקציית הסתברות שמוגדרת מראש עבור אותה מוטציה-משנה [1, 7].

כמו כן, ביישום אלגוריתמים גנטיים לחיבור מנגינות, בעוד שחלק ממוטציות המשנה מקיימות את תכונת הניטרליות המצופה מהן ע"י כך שהן פועלות על ערכי הארגומנטים שלהן כסקלרים מופשטים ללא תלות במשמעות הייצוג, מוטציות-משנה אחרות דווקא מממשות שינויים בהשראת כללים מתורת הקומפוזיציה ואימוץ טכניקות אלתור, מה שהופך אותן למוטות יותר כלפי פתרונות העונים לכללים התאורטיים המוסיקליים שמעצב האלגוריתם מגדיר. הפרה זו של תכונת הניטרליות והטיית כיוון החיפוש במרחב החיפוש המורכב של חיבור המנגינה מבוצעת במטרה שהשינויים אכן יובילו לשיפור רמת הכשירות ובסופו של דבר להתכנסות של האלגוריתם, תחת ההשקפה שהאלגוריתמים הגנטיים הם אמצעי שאמור לתמוך בנו, ולא להגביל אותנו, כך שניתן לראות את כללי העיצוב שלו כעצות והנחיות שאין חובה לאמץ [1].

לטובת תמיכה בסקליליות מרבית, לכל מוטציה-משנה מוקצה פרמטר הסתברות מקומי הקובע את ההסתברות לביצועה בהפעלה מסוימת של האופרטור, ובהתאם לסוג השינוי, לחלק מהמוטציות מוגדר בנוסף גם פרמטר מקומי של רמת/היקף השינוי. הסקליליות באה לידי ביטוי באפשרות לעדכן פרמטרים אלו בין איטרציה לאיטרציה, שבהתאמה מכתבים באופן דינאמי את מגמת החיפוש במרחב: למשל כיוון הפרמטרים לשינויים דרסטיים בהסתברות גבוהה מכתבים חיפוש רוחבי ואילו כיוון הפרמטרים לשינויים מתונים/זניחים בהסתברות נמוכה מכתבים חיפוש לעומק ומובילים להתכנסות מהירה יותר.

נציג כעת דוגמאות אחדות למוטציות מאפליקציות שונות של אלגוריתמים גנטיים לחיבור מנגינות ואלתורים [1,2,7], ונחלק אותן למחלקות בהתאם לארגומנט שעליו פועלת המוטציה (תווים בודדים או תיבה שלמה). נתאר את המוטציות ברמת הפשטה גבוהה. את המימוש הקונקרטי יש להגדיר כמובן בהתאם לשיטת הקידוד שנבחרה עבור האלגוריתם (גובה צליל אבסולוטי/רלטיבי, קידוד מפוצל/משולב). נעיר כי מוטציות מסוימות עשויות להיות קלות מאוד למימוש בשיטת קידוד אחת אך די מורכבות ואף בלתי אפשריות בשיטת קידוד אחרת.

3.3.2.1 דוגמאות למוטציות ברמת התווים (גנים) האינדיבידואליים

מוטציות אלו פועלות על אחד מהתווים שנבחר באקראי –

א. **שינוי גובה צליל** – שינוי גובה הצליל של התו הנבחר לגובה הצליל הנמצא d דרגות מעליו/מתחתיו בסולם, כאשר d הוא פרמטר שניתן לעדכן באופן דינאמי בכל איטרציה במסגרת התמיכה בסקליליות.

- ב. **השתקת צליל** – החלפת הצליל בהשתקה בעלת ערך ריתמי זהה לזה שהיה קודם לצליל.
- ג. **פיצול תו בודד לשני תווים** – התו הנבחר מוחלף בשני תווים שערכם הריתמי הכולל שווה לערך הריתמי של התו המקורי כך שמבחינת הקצב יתר התווים בתיבה לא מושפעים מהשינוי. מבחינת הערך הריתמי של שני התווים המחליפים, אפשר לקבוע ביניהם יחס של 1:1, יחס של 1:3 או יחס של 3:1. למשל, אם לתו המקורי היה ערך ריתמי של חצי תיבה, יחס של 1:1 בין שני התווים המחליפים אותו יניב ערך ריתמי של רבע לכל אחד מהם, יחס של 1:3 יניב ערך ריתמי של שמינית לתו הראשון ורבע מנוקד לתו השני, ויחס של 3:1 יניב חלוקה דומה למקרה הקודם עם חילופי תפקידים. באשר לגובה הצלילים של התווים החדשים, את הראשון נהוג לשמר כמו המקורי, ואת השני ניתן פשוט לשכפל לאותו גובה צליל כמו הראשון, או לחלופין לקבוע אותו כאחד מצלילי האקורד המתנגן במקביל לתו המקורי. הבחירה באחד מצלילי האקורד היא דוגמה לניצול ידע מומחה שמפר את תכונת הניטרליות המצופה מאופרטור מוטציה כללי, ש-Biles טוען שאין חובה לקיימה אם היא לא מקדמת אותנו אל היעד [1].
- ד. **איחוד שני תווים**: פעולה זו פועלת באופן הפוך לזו הקודמת: התו העוקב לתו הנבחר מחובר אליו עם קשת חיבור כך ששניהם מתאחדים לגובה הצליל של התו הנבחר והערך הריתמי החדש שלו הוא סכום הערך הריתמי שלו ושל התו העוקב לו. באופן דומה, ניתן לחבר את התו שנבחר עם קשת חיבור לתו שקודם לו דווקא.
- ה. **הוספת צלילי מעבר** – הוספת צלילי מעבר בין התו הנוכחי לתו העוקב לו, כאשר צלילי המעבר הם סדרה עולה או יורדת של הצלילים בסולם שמפרידים בין התו הנוכחי לזה העוקב לו. בכדי לא להפר את האיזון הקצבי של המשקל, הערך הריתמי שניתן לצלילי המעבר ייגרע מזה של התו הנבחר. בכדי שהדגש על הצליל המקורי יישמר והצלילים החדשים אכן רק ישמשו ל-"החלקת" המעבר בינו לבין הצליל העוקב לו, ניתן לשמר אצל הצליל המקורי חצי מערכו הריתמי המקורי, ואת החצי השני לחלק באופן שווה בין צלילי המעבר.
- זוהי דוגמה למוטציה שלא ניתנת למימוש בשיטת הקידוד המשולבת: כזכור בשיטת הקידוד המשולבת הכרומוזום מורכב מגנים שמייצגים את היחידות הריתמיות הקצרות ביותר. לפיכך "אין מקום" מבחינת קצב להכניס תווים נוספים בין שני תווים עוקבים, שכן כל תו חדש יעלה את הערך הריתמי הכולל של התיבה שבה הוא נמצא, ולהבדיל משיטת הקידוד המפוצלת, בשיטה זו אין אפשרות לקצר את הערך הריתמי של תו נתון (בהנחה שהוא כבר מכיל את הערך הריתמי הקצר ביותר). עם זאת, ניתן להגדיר גרסה מוגבלת של מוטציה זו עבור שיטת הקידוד המשולבת כך שתפעל רק על תווים שמחברים בקשת בקשר חיבור לתו שמלפניהם, דהיינו ערכים שמקודדים עם הערך '15', מה שאומר שמשך השהייה שלהם ארוך כלפחות שתי יחידות ריתמיות קצרות ביותר שניתן לפצל ולנצל לצורך שיכון צלילי המעבר במקום קשתות החיבור (כנגד כל קשת חיבור ניתן להכניס צליל מעבר יחיד).

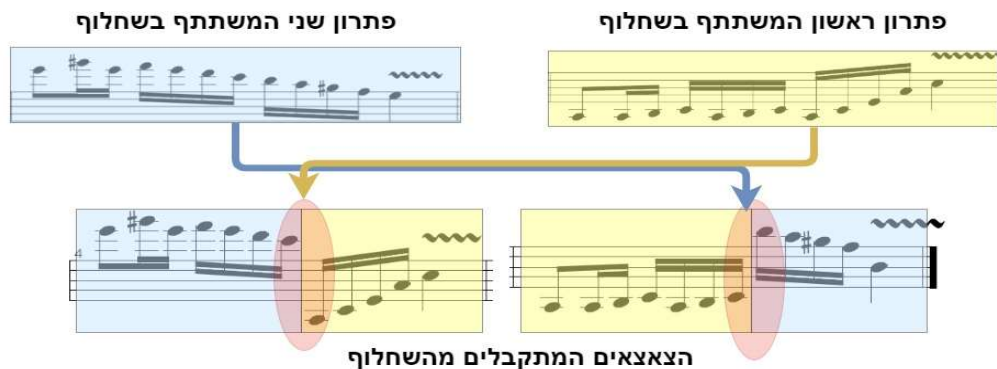
3.3.2.2 מוטציות ברמת תיבה שלמה

- מוטציות אלו פועלות על תיבות שלמות. התיבה שעליה מבוצע האופרטור נבחרת באקראי –
- א. **היפוך סדר התווים בתיבה** – הפיכת סדר התווים בתיבה (האחרון הופך לראשון, אחד לפני אחרון לשני וכן הלאה). מוטציה זו היא דוגמה למוטציה ניטרלית שאינה מודעת לתוכן הייצוג והמשמעות שלו.
- ב. **מיון בסדר עולה/יורד** – סידור מחדש של כל הצלילים בתיבה הנבחרת מנומך לגבוה (מיון עולה) או מגבוה לנמוך (מיון יורד), כאשר המיון מבוצע רק על ערכים המייצגים גבהי צלילים: ערכים שמייצגים הפסקות או קשתות חיבור של משך שהייה שומרים על מקומם המקורי בתיבה (וקטור) שאינו מושפע מהמיון.
- ג. **טרנספוזיציה** – היסט כלל ערכי גבהי הצלילים שבתיבה הנבחרת במספר קבוע של צעדים בסולם כלפי מעלה או כלפי מטה. בדומה למיון, בטרנספוזיציה משנים רק את הערכים שמייצגים גבהי צלילים, שהרי אין משמעות לביצוע שינוי גובה להפסקה או לקשת חיבור, ערכים אלו נשארים כמות שהם גם לאחר השינוי.
- ד. **היפוך בין הפסקות לקשתות חיבור** – הפיכת כל תו הפסקה לתו קשת חיבור שמאריך משך השהייה של התו הקודם לו, ולהפך, דהיינו הפיכת כל תו קשת חיבור להפסקה.
- נציין כי רשימת המוטציות שהובאה כאן היא חלקית ביותר ומציינת רק דוגמאות להמחשה. ניתן להגדיר עוד אינספור מוטציות, שיכולות להיות ניטרליות באופן שתוכן הכרומוזום ומשמעות הייצוג שלו שקוף למוטציה, או לחלופין מוטציות שמעוצבות בהשראת טכניקות אלתור

וקומפוזיציה שמכפיפות את עצמן למסגרת אילוצים הנובעים מכללי מהתאוריה המוסיקלית המערבית.

3.3.3 שחלוף בין מנגינות

האתגרים הקיימים בשחלוף מנגינות: אופרטור [השחלוף](#) מיועד כאמור לייצר פתרונות חדשים ע"י מיזוג חלקי כרומוזומים קיימים לפי סדר המתחלף לסירוגין, בשאיפה כי הפתרונות החדשים (או לפחות חלק מהם) ימוזגו מהחלקים היותר "טובים" של הכרומוזומים היוצרים אותם, או שאקט השילוב עצמו, אפילו אם הוא משלב חלקים "בינוניים" ייצר תולדה של משהו חדש וטוב יותר, וע"י כך הפתרונות החדשים יתעלו על אלו הקיימים. עם זאת, שחלוף עשוי לא רק שלא לייצר פתרונות משופרים, אלא להרע את המצב ולייצר פתרונות גרועים יותר. הדבר בולט עוד יותר במרחב החיפוש המורכב של מנגינות. נניח למשל שיש שתי מקטעי מנגינות "מעולות", הראשונה היא מנגינה שיורדת מאיזשהו תו גבוה, והשנייה היא מנגינה שעולה מאיזשהו תו נמוך. אם מבצעים על שתי מנגינות אלו שחלוף למשל של 1-point-cross-over, עלול להיווצר מצב שנקודת הפיצול הנבחרת היא בלב העלייה/ירידה, מה שגורם לשינוי קיצוני מעלייה לירידה (ולחפך) ולא רק זאת אלא שהמעבר הוא בין מנעד שונים לגמרי, כך שגם יש קפיצה דרסטית פתאומית בגובה הצליל. להלן המחשה למקרה קיצוני שכזה, השטחים באדום מסמנים את המרווח הקיצוני (סדר גודל של שתי אוקטבות) הנוצר כתוצאה משחלוף (נקודת הפיצול בדוגמה זו היא אמצע התיבה/כרומוזום) –



סביר להניח שהמנגינות שנוצרות משחלוף שכזה יפורשו ע"י המאזין ככאוטיות וחסרות סדר וכיוון. כדי להימנע ממצבים כאלה ולהבטיח שהשחלוף יבוצע בצורה חלקה, מימושים אחדים שיפרו את השחלוף הבסיסי כך שיחשב את נקודת התפר האופטימאלית מבחינת המרווחים שייווצרו בין הצלילים שמשני צדי נקודה זו [1,4,7]. באופן זה, אם יוצאים מנקודת הנחה שבתוך כל מקטע קיימת איזושהי "רציפות" מקומית מבחינת היעדר קפיצות חריגות בין צלילים, מספיק שנוודא שרציפות זו נשמרת גם בנקודות התפר שבין המקטעים וע"י כך ביצירה כולה. המחיר הוא כמובן בביצועי זמן הריצה.

כאשר הקידוד מבוצע בשיטת הקידוד הנפרד עבור גבהי הצלילים והיחידות הריתמיות שלהם, מתעוררים בעיות נוספות, שכן להבדיל משיטת הקידוד המשולב, בשיטה זו החלוקות הפנימיות של התיבות אינן רגולריות, כלומר כרומוזומים של תיבות שונות עשויים להיות מורכבים מווקטורים באורכים שונים, ופעולת המיזוג/שחלוף כלל לא מוגדרת היטב עבור כרומוזומים באורכים שונים (למשל בשיטת ה-n-point crossover). ניסיון נאיבי להתמודד עם בעיה זו, יהיה לקבוע את נקודת הפיצול לא לפי הרכיבים של הווקטור המייצג, אלא לפי איזשהו מרחק ריתמי מתחילת התיבה, למשל להחליט שנקודת הפיצול באיזושהי תיבה היא פעימה (רבע) אחת מתחילת התיבה. הצעה זו אמנם מתגברת על הבעיה הראשונה, אך עלולה לייצר בעיה חדשה: ייתכן שיש תו הקודם לנקודת הפיצול באחת מן התיבות שהערך הריתמי שלו גולש וחוצה אל מעבר לנקודת הפיצול, מה שמונע חיתוך "נקי". אותה בעיה עשויה לצוץ אפילו אם מצמצמים את ביצוע השחלופים רק בין תיבות שלמות (שאז מובטח שיש להן אותו משקל כולל), שכן שוב, ייתכן שיש תו באחת מן התיבות שמשך השהייה שלו גולש עד לתיבה העוקבת לו. בכדי להתמודד עם בעיות אלו דרך אחת היא חיפוש נקודות פיצול שבהן ניתן לבצע חיתוך "נקי", דהיינו חיתוך שמתחיל עם רכיבים שמתחילים בדיוק מאותו ערך ריתמי ביחס לראש התיבה. דרך נוספת היא לבצע "כירורגיה" לתכולת הכרומוזום ובמקרה של חלוקה לא נקייה לפצל את היחידה הריתמית שגולשת ליחידות קצרות יותר (למשל פיצול חצי לארבע שמיניות), מה שיאפשר להתייחס לנקודה שקודם הייתה פנימית בתוך איזשהו רכיב מסוים בווקטור כאל רכיב נפרד חדש, וע"י כך ליצור חלוקה נקייה. סוגיה זו מלמדת על דוגמה נוספת של יתרון הפשטות שמביאה שיטת הקידוד המשולב.

3.4 הערכת כשירות המנגינות

סובייקטיביות ההערכה של מוסיקה

כאמור, רמת הכשירות של פתרון מועמד מציינת עד כמה הוא טוב או איכותי. בהקשר היישום הנוכחי שבו הפתרונות המועמדים הם מנגינות, לצד מדדים אובייקטיביים להערכת טיב מנגינה כמו מבניות ושמירה על מרווחים קונסוננטיים, באופן כללי, הקביעה שמנגינה אחת איכותית יותר מאחרת היא דווקא די סובייקטיבית, כפי שהפתגם אומר "על טעם וריח, אין להתווכח", כך גם במוסיקה, קהלים שונים עשויים להחזיק בדעות שונות על טיב איכות של מנגינה בהתאם להעדפות האישיות שלהם, שיכולות לבוא לידי ביטוי למשל, בין היתר, באספקטים הבאים – התחושה הכללית של הרגשות שהמנגינה משדרת (קלילות/כבדות, עליזות/עצבות וכד'), רמת הדחיסות (תדירות הופעה) של הצלילים, רמת הגיוון וכד'.

הערכה אינטראקטיבית לעומת הערכה אוטומטית

לאור הסובייקטיביות של הערכת טיב האיכות של מנגינה, טבעי ביותר לנסות לבצע את הערכת הכשירות באופן אינטראקטיבי, כך שכל משתמש יוכל לדרג את המנגינות שיוצרו בהתאם לראות עיניו (או ליתר דיוק השמיעה באזניו), כך שמנגינות שפחות לטעמו "ייכחדו" עם הזמן וייוותרו לבסוף המנגינות שהכי מוצאות חן בעיניו. כמו כן, שימוש בהערכה אינטראקטיבית מאפשר לאלגוריתם להתייחס לשגרת השערוך כאל "קופסא שחורה", מבלי הצורך להכיר את פרטי המימוש שלה. בפרט, זה חוסך את הצורך להגדיר באופן פורמלי את העדפות המשתמש ואת המדדים לדירוג איכות של מנגינה, במושגים של בעיה אלגוריתמית, מה שהופך את האלגוריתם להרבה יותר פשוט למימוש.

שיקולים אלו ודומים להם הובילו את Biles [1] להוציא את הגרסה הראשונית של GenJam, מערכת התוכנה שפיתח לחיבור סולואים מאולתרים בג'אז, עם מימוש בשיטת הערכה אינטראקטיבית: המערכת הייתה מנגנת פרזות (משפטים קצרים, דהיינו מנגינות קצרות) והוא היה מדרג אותן בזמן אמת. עם הזמן התגלו שתי מגרעות קריטיות בשיטה זו –

א. צוואר הבקבוק של המגבלה הטמפורלית: להבדיל מיישומי אלגוריתמים גנטיים לעיבוד וסיווג תמונות שבהם שיטות ההערכה האינטראקטיביות דווקא הוכיחו את עצמן [1], בין היתר בשל העובדה שהעין האנושית יכולה לקלוט בזווית הראייה ספקטרום רחב של עצמים (פתרונות) במקביל ולבחור בין רגע את אלו המתאימים ביותר מתוך אוכלוסייה גדולה, ביישומים קוליים לא ניתן להקשיב לכמה מנגינות (פתרונות) במקביל, שכן גלי הקול הפיזיים של מנגינות מתמזגים ולא ניתן להבחין ביניהם. עובדה זו מאלצת את המבקר האנושי להאזין ולדרג כל פתרון מועמד בנפרד, מה שמציב חסם תחתון על משך הזמן של תהליך השערוך (זמן אורך היצירה המוסיקלית, לכל הפחות). אם הרצה טיפוסית של אלגוריתם גנטי מצריכה למשל סדר גודל של כ-100 איטרציות עד להתכנסות לפתרון סביר, ואורך מנגינה טיפוסית הוא כ-30 שניות, אזי הרצה טיפוסית שלמה של האלגוריתם עלולה לקחת כשעה שלמה!

ב. המגבלה האנושית: אפילו אם היה לאדם את כל הזמן שבעולם, Biles מעיד שעם הזמן הדירוג האנושי הופך לפחות ופחות אותנטי [1]: ככך שהזמן חולף, רמת הריכוז של המבקר האנושי יורדת, העייפות עולה והזיכרון של מנגינות קודמות שהושמעו מתפוגג, מה שמקשה עליו להשוות עד כמה מנגינה אחת טובה ביחס לאלו שקדמו לה. מצב זה מוביל לסיווג מאוד כללי והחלוקה לרמות כשירות היא חלוקה למחלקות מאוד גסות כגון "רע", "סביר" ו-"טוב", מבלי יכולת לבדד באופן אותנטי את המנגינות הטובות במיוחד מבין כל אלו הטובות.

מגרעות אלו של ההערכה האינטראקטיבית הובילו לגיבוש פונקציות הערכה אוטומטיות, הן ע"י Biles עצמו בגרסאות מתקדמות יותר של GenJam [1] והן ע"י חוקרים נוספים במערכות ניסיוניות דומות שפיתחו בשנים האחרונות [4,7]. פונקציות ההערכה האוטומטיות מתגברות על צוואר הבקבוק של המגבלה הטמפורלית ע"י ניצול הכוח החישובי של מחשב טיפוס המסוגל לדרג אוכלוסיות גדולות בשבריר שנייה, וכל זאת תוך דבקות במטרה ללא "איבוד ריכוז" שעלול להשפיע על כושר השיפוט, שאליו חשוף המבקר האנושי.

עם זאת, בניית פונקציית הערכה אוטומטית אינה דבר של מה בכך, זוהי משימה מאוד מאתגרת: כיצד ניתן לבצע פורמליזציה של מדדי איכות של מנגינה? כלומר, איך מודל חישובי כלשהו יכול בכלל לנתח ולהעריך עד כמה מנגינה נשמעת טוב מבלי שהוא בכלל מקשיב לה קודם לכן?

יתר על כן, אפילו אם מצליחים להתגבר על אתגר זה ולעצב פונקציה המסוגלת לנתח ולהעריך מנגינות באופן "יבש" רק על סמך התיאור שלהן, כיצד ניתן להתאים פונקציות אוטומטיות אלו באופן שישקפו את ההעדפות האישיות של המשתמש שמריץ את האלגוריתם? בהמשך פרק זה נדון בדיוק בשתי סוגיות אלו.

מימוש פונקציית הערכה אוטומטית

בכדי שפונקציית הערכה תוכל לנתח מנגינה ולאמוד את רמת האיכות שלה על סמך התיאור שלה בלבד, כך שתוכל להחליף את המבקר האנושי ולבצע את מלאכת השערוך באופן אוטומטי, נעשים ניסיונות להטמיע בה את הידע, התובנות והניסיון שיש למבקר האנושי, לצד ההעדפות האישיות שלו. לרוב, הטמעת הידע, התובנות והניסיון מבוצעת ע"י בניית מערכת חוקים הנגזרים מתאוריה מוסיקלית, תורת הקומפוזיציה וכללים היוריסטיים. בכדי להביא לידי ביטוי את ההעדפות המשתמש, ניתן לספק לו את היכולת להגדיר מראש באיזה מדדים הוא מעוניין להשתמש, איזה כללים הוא מעוניין לאכוף, ומה החשיבות של כל מדד/כלל ביחס ליתר, כך שע"י קביעת פרמטרים מתאימים המנגינות שיווצרו יכוונו במידה מסוימת לפי טעמו האישי.

נדגים מימוש לפונקציית הערכה אוטומטית המבוססת על מערכת "כללי נימוס" מוזיקליים היוריסטיים שיכולים לשמש כמדד לטיב המנגינה. בדומה לאופן שבו אופרטור המוטציה הורכב מאוסף של שגרות מוטציות-משנה, גם כאן העבודה המורכבת מחולקת ומפוצלת בין פונקציות משנה להערכה, כאשר כל אחת מהן אחראית לבדוק כלל/מדד מסוים, ולדרג את המנגינה ע"פ כלל זה, ולהחזיר ציון מנורמל לסקלה אחידה שתואמת לכלל הפונקציות. בכדי לשקלל את כלל המדדים גם יחד, לכל פונקציית משנה מוקצה משקל יחסי מתוך המשקל הכולל, בהתאם לחשיבות שלו בהערכת הכשירות, ע"פ הפרמטרים שהמשתמש קובע. לאחר חישוב הציונים היחסיים עבור כלל הכללים שנבחרו לשמש כמדד, פונקציית ההערכה המרכזית משקללת אותם לכדי סכום ממושקל שמהווה את רמת הכשירות של המנגינה כולה. ובאופן פורמלי –

נסמן ב- R את אוסף הכללים המשמשים כמדד לטיב ההערכה, וב- r_i ($1 \leq i \leq k$) את הכללים האינדיווידואליים: $R = \{r_1, r_2, \dots, r_k\}$. כנגד אוסף כללים זה נגדיר קבוצה $F = \{f_1, f_2, \dots, f_k\}$ של פונקציות-משנה אשר מכילה כנגד כלל r_i פונקציית תואמת f_i לכל $1 \leq i \leq k$, אשר מעריכה את הכלל r_i ומחזירה עבורו ציון בטווח $[0,1]$. כמו כן נגדיר אוסף משקלים $W = \{w_1, w_2, \dots, w_k \mid \forall i (1 \leq i \leq k) \rightarrow w_i \in [0,1] \wedge \sum_{i=1}^k w_i = 1\}$ שסכומם 1, המתאים לכל פונקציה f_i משקל $w_i \in [0,1]$, לכל $1 \leq i \leq k$. אזי ערך הכשירות הכולל שיינתן ע"י פונקציית ההערכה האוטומטית עבור איזשהו כרומוזום x מוגדר ע"י הסכום המשוקלל שלהלן –

$$(*) \quad total_fitness(x) = \sum_{i=1}^k w_i \cdot f_i(r_i(x))$$

ניתן לראות את הכללים כחוקים שהמנגינות צריכות לציית להן, ואת פונקציות-המשנה התואמות כשומרות חוק שאחראיות לאכוף את הכללים: מנגינות שיעמדו בהן יקבלו ציון גבוה, ואילו מנגינות אשר יפרו אותם "ייקנסו" ויקבלו ציון נמוך.

דוגמה

נמחיש את אופן השימוש במנגנון הכללים ע"י הצגת כלל לדוגמה. נקרא לכלל זה "סמיכות צלילים שכנים". להלן הכלל – "מנגינה היא טובה אם צלילים עוקבים נמצאים יחסית קרוב זה לזה מבחינת גובהם האבסולוטי". בהתאם להנחת כלל זה (שהינה כמובן הנחה סובייקטיבית שניתן לחלוק עליה), פונקציית ההערכה אחראית לוודא שאין קפיצות חריגות (קפיצות גבוהות מידי) בין תווים סמוכים, דהיינו שהמרווח הגדול ביותר בין כל זוג תווים עוקבים במנגינה אינו גבוה מאיזשהו מרווח מוגדר מראש, למשל סקסטה גדולה (תשעה חצאי טונים). ניתן לחשוף פרמטר זה למשתמש כדי שיקבע הוא בעצמו את המרווח המקסימאלי הרצוי – באופן כללי, ככל שהמרווח קטן יותר אופי המנגינה יקבל צורה סגורה/מופנמת יותר, ואילו מרווחים גדולים יעודדו סאונד יותר פתוח ומוחצן.

נדגים מימוש אפשרי לפונקציית הערכה שאוכפת כלל זה, המתאים לשיטת הקידוד שבה קידוד הצלילים מבוצע ע"פ גובהם האבסולוטי, ושגובהי הצלילים וערכם הריתמי מתוחזקים בווקטורים (מערכים) נפרדים במבנה הנתונים המייצג את הכרומוזום. כזכור, בשיטת קידוד זו המרחק בין כל שני ערכים מציין את המרחק ביניהם ביחידות של חצאי טונים, והערך 1- (מינוס אחת) מציין תו של

הפסקה. הפונקציה שנמש מקבלת כקלט שני פרמטרים: (1) $max_interval$ המציין מהו המרווח המקסימאלי (בחצאי טונים) שנחשב כסביר בין זוג צלילים עוקבים; (2) וקטור גבהי הצלילים (והפסקות). הפונקציה סורקת את הווקטור תוך דילוג על תווי הפסקה, מונה את זוגות המרווחים ה-"חריגים" שהמרחק ביניהם (הפרש בערך מוחלט) גדול ממש מהמרווח ה-"סביר" המקסימאלי כפי שמוגדר בפרמטר הקלט לפונקציה, ולבסוף מחשבת את החלק היחסי של המרווחים ה-"סבירים" בין זוגות תווים שכנים מתוך כלל זוגות המרווחים בין זוגות השכנים. חישוב החלק היחסי מבוצע בכדי לנרמל את הציון לטווח $[0,1]$. באופן דומה יש לנרמל בצורה כזו או אחרת את כלל פונקציות הכשירות לטווח משותף כדי שידרגו לפי אותה הסקלה.

להלן פסידו-קוד של פונקציית הכשירות עבור כלל "סמיכות צלילים שכנים" שתואר לעיל –

Neighbors-Distance-Fitness-Function ($max_interval, pitch_vector$)

```

1.  $total\_num\_of\_intervals \leftarrow 0$ 
2.  $num\_of\_extreme\_intervals \leftarrow 0$ 
3.  $n \leftarrow length[pitch\_vector] - 1$ 
4. for  $i \leftarrow 1$  up to  $n$ 
5.   if  $pitch\_vector[i] = -1$                                 ▷ current note is a silent note
6.      $i \leftarrow i + 1$                                        ▷ skip to next note
7.   else if  $pitch\_vector[i+1] = -1$                         ▷ next note is a silent note
8.      $i \leftarrow i + 2$ 
9.   else do
10.     $total\_num\_of\_intervals \leftarrow total\_num\_of\_intervals + 1$ 
11.    if  $|pitch\_vector[i+1] - pitch\_vector[i]| > max\_interval$ 
12.      then  $num\_of\_extreme\_intervals \leftarrow num\_of\_extreme\_intervals + 1$ 
13.  $normalized\_fitness \leftarrow \frac{(total\_num\_of\_intervals - num\_of\_extreme\_intervals)}{total\_num\_of\_intervals}$ 
14. return  $normalized\_fitness$ 
```

כלל בסיסי זה שהוצג הובא רק כדוגמה להמחשת המנגנון. עבור יישום מעשי, סביר שיידרש אוסף רחב של מדדים נוספים, כגון – רמת דחיסות הצלילים בתיבה, רוחב מנעד (פער בין צליל נמוך ביותר לצליל הגבוה ביותר), שימוש בתבניות/מוטיבים חוזרים, תדירות ההפסקות (חלוקת המנגינה למשפטים), היקף רמת השימוש בסינקופות, חפיפת צלילים לאקורד/סולם שמתנגן אליהן במקביל, קיום עליות וירידות, היקף מגוון סוגי היחידות הריתמיות, ועוד.

לאחר גיבוש רשימת הקריטריונים שרוצים למדוד על-פיהם את טיב המנגינה, חוזרים באופן דומה על תהליך מימוש פונקציית הערכה לכל כלל/קריטריון נוסף כפי שהודגם לעיל עבור כלל סמיכות הצלילים השכנים. לבסוף מקצים לכל פונקציה משקל יחסי כך שסך המשקלים מסתכם ל-1 ומחשבים את רמת הכשירות הכוללת ע"פ משוואה (*).

עם סט מגובש של מגוון קריטריונים אפשריים לאומדן ופונקציות הערכה לדירוג על פיהם, המשתמש המפעיל את האלגוריתם הגנטי יכול לכוון את המנגינות הנוצרות כך שיעוצבו ע"פ טעמו האישי [4], ע"י בחירת המדדים המבוקשים לבדיקה וכוונון הפרמטרים של המשקלים היחסיים של הקריטריונים השונים לצורך דירוגם היחסי ע"י פונקציית ההערכה.

נעיר כי מימוש חלק מהפונקציות אינו טריוויאלי כלל וכלל: למשל, מימוש הקריטריון של חזרה על תבניות/מוטיבים במנגינה, דורש חיפוש תת-מחרוזות דומות שחוזרות פעמיים או יותר בווקטור-כרומוזום. מימוש לא הולם של פונקציה שכזו עלול להוביל בקלות לגידול דרסטי בסיבוכיות זמן הריצה. דוגמה נוספת היא קריטריון קיום מגמה כיוונית של עלייה/ירידה במנגינה. ניסיון אחד למדוד קריטריון כזה במנגינה הוא לסרוק את וקטור הצלילים בכרומוזום ולבדוק בכמה מן הזוגות הרכיב הראשון גבוה יותר מזה העוקב לו, ובכמה להפך, לבסוף למנות מה יש יותר (עליות או ירידות) ובהתאם לכך להסיק האם קיימת במנגינה מגמות עלייה/ירידה [7]. אני סובר שמימוש שכזה אינו מוצלח במיוחד: למשל על מנגינה שהחצי הראשון שלה מהווה עלייה רצופה והחצי השני ירידה רצופה, ברור על

אף שיש כיוונות ברורה בשני חלקים נפרדים, העליות והירידות יקזזו זה את זה לפי שיטה זו, ויניבו ציון זהה כשל פונקציה שעולה ויורדת לסירוגין תו אחר תו (שבאופן מובהק חסרת כיוון מגמתי לעלייה/ירידה).

לסיום פרק זה, נציין כי מודל זה של התמודדות עם ריבוי קריטריונים ע"י מתן משקל יחסי לפונקציות ההערכה שלהם וסכימת כלל הפונקציות לסכום ממושקל אחד בכדי שניתן יהיה להתייחס אליהן כאל פונקציית מטרה אחת, הוא אמנם מודל די נפוץ ביישומי מערכות לחיבור מנגינות ע"י שיטות אבולוציוניות [2,7], אך ישנן גם שיטות אחרות: שיטה אחרת להתמודדות עם ריבוי קריטריונים, היא כפי שצוין [לעיל](#), חתירה ליעילות פארטו ע"י הצלבה בין הקריטריונים השונים. בשיטה זו משתמשים למשל Jaehun et al. בניסיון חתירה במקביל לשיפור שני יעדים מנוגדים למימוש ביצירה: יציבות מחד ומתח מאידך [4]. שיטה זו של יעילות פארטו כאמור חורגת מתחום הדיון של סמינר זה.

4 סיכום

אלגוריתמים גנטיים הם אלגוריתמים שעוצבו בהשראת תהליכים אבולוציוניים מהטבע תחת התבונה שניתן להפעיל תהליכים כאלו באופן מלאכותי כאסטרטגיה לפתרון בעיות אופטימיזציה קשות. אלגוריתמים אלו מבצעים חיפוש היוריסטי איטרטיבי, הצובר אוכלוסייה של פתרונות מועמדים ההולכת ומשתפרת לאורך זמן ע"י הרכבת פתרונות חדשים, השוואתם לאלו הקיימים ושמירת אלו הטובים יותר, תוך שימוש מסוים באקראיות.

ביסוס החיפוש של אלגוריתמים גנטיים על אוכלוסייה של פתרונות מהווה את אחת המעלות הגדולות שלהם, המאפשר בין היתר לבזר את החיפוש בין סוכנים שונים במקביל, שיכולים להתפרש לאזורים שונים במרחב החיפוש, לתקשר ביניהם ולהתאים את אופי ומיקוד החיפוש באופן דינאמי בהתאם לגילויים בשטח.

עם זאת, על אף היתרונות הרבים של האלגוריתמים הגנטיים והספקטרום הרחב של בעיות שהם מסוגלים להתמודד עימם, יש גם אתגר רציני בהפעלתם: הגנריות והגמישות של האלגוריתמים מצריכים כוונן פרמטרים מתאים בכל הפעלה, ולכל בעיה מתאים סט אחר של פרמטרים, כך שבכדי לנצל את מלוא הפוטנציאל של האלגוריתמים ע"י מציאת הפרמטרים האופטימאליים עבור בעיה נתונה, יש להתמודד תחילה עם בעיית אופטימיזציה של מציאת פרמטרים אלו, כלומר ביצוע חיפוש טרם החיפוש, עובדה שמשפיעה משמעותית על סיבוכיות זמן הריצה, שממילא עשויה להיות כבדה לאור הביצועים היקרים של פונקציות כשירות טיפוסיות.

בהקשר של אלגוריתמים ומוסיקה, אלגוריתמים גנטיים צברו פופולריות רבה בשלושת העשורים האחרונים ככלי להלחנה אלגוריתמית, בשלל אספקטים של מלאכת ההלחנה, ביניהם: חיבור מנגינות תחת קונטקסט הרמוני נתון ולהפך (חיבור מהלך הרמוני עבור מנגינה נתונה), חיבור תבניות קצביות, חיבור סולואים מאולתרים, ואף הלחנת יצירות מלאות מאפס.

במשימת ההלחנת מנגינת סולו תחת קונטקסט הרמוני נתון (שזוהי המשימה שבה התמקד סמינר זה), האלגוריתם הגנטי מנסה לאתר את מנגינות הסולו הטובות ביותר שניתן לייצר מתוך אוסף של צלילים וערכים ריתמיים תוך שילוב של אקראיות והיוריסטיקה. הגדרת "טיב" איכות של מנגינות ומוסיקה בכלל היא עניין די סובייקטיבי, כך שמעצב האלגוריתם צריך לדרג בעצמו עד כמה איכותיים הפלטים השונים, בין אם ע"י דירוג אינטראקטיבי ידני, ובין אם ע"י הגדרת פונקציית כשירות אוטומטית, המנסה להעריך את איכות הפלט על-סמך ניתוח "יבש" של המנגינה מתוך תיאור שלה בלבד, תוך שימוש בכללים תאורטיים מהמוסיקה המערבית והעדפות משתמש.

הדירוג האינטראקטיבי גוזל זמן רב וכפוף למגבלות המשאב האנושי. לעומת זאת, דירוג אוטומטי ע"י פונקציית כשירות אמנם מהיר אך לא בהכרח משקף באופן מלא את כל המדדים שמשפיעים על הערכת איכות הפלט. מלאכת הפורמליזציה של מדדים וקריטריונים שגורמים לנו להעריך מנגינה, הינה מלאכה מאתגרת, המורכבת משלל יעדים שונים מנוגדים: הרצון ליצור נקודות של מתח לעומת נקודות יציבות, הרצון לבצע דחיסות צלילים לעומת הרצון להכניס הפסקות תדירות בכדי שתהיה חלוקה ברורה של המנגינה למשפטים וכד'.

אפילו עם פונקציית כשירות טובה שמכסה מגוון רב של קריטריונים, אחד האתגרים הכי גדולים בהלחנה אוטומטית של מנגינות סולו שטרם נמצא לו פתרון שיטתי נאות, הוא יצירת "סיפור" ע"י הסולו, עם פתיח, נקודות שפל, נקודות שיא וסיום: מימושים אחדים של אלגוריתמים גנטיים אמנם מצליחים לתפור בתוך הסולואים שהם מחוללים פרזות (משפטים קצרים בתור המנגינה) שמתלבשות "בול" על בסיס הרקע ההרמוני המתנגן, אך כל שיבוץ שכזה הוא מקומי (לוקאלי), ואינו בהכרח מתחבר לפרזות שלפניו ואחריו, כך שהסולו הנוצר הוא אוסף של משפטים קצרים שכל מהם אולי נשמע מצוין כשלעצמו, אך כמכלול, אין סיפור עלילתי שמחבר בין כל המשפטים האלה.

האם בעתיד נצליח לשכלל אף יותר את הפורמליזציה של בעיית ההלחנה והערכת טיב האיכות שלה, ולעצב את החיפוש של האלגוריתמים הגנטיים וטכניקות ההלחנה אלגוריתמית בכלל כך שידמו את החיפוש לזה שמבוצע ע"י בן אנוש, התר אחר סיפור שיבטא וישדר איזושהי אמרה או פרץ רגשות שהוא רוצה להביע?

זוהי שאלה מעניינת, ויהיה זה מעניין לא פחות להקשיב לפלטים שיווצרו ע"י מחשב במידה והתשובה לכך תתגלה כחיובית ביום מן הימים. עד אז, אני סבור שבכדי להפיק את המרב מהאלגוריתמים הגנטיים במלאכת ההלחנת מנגינות סולו, מוטב יהיה שלא להטיל עליהם את המלאכה כולה ולקבל את הפלט שלהם כתוצר מוגמר, אלא להשתמש בהם כאמצעי עזר לחקר וגילוי רעיונות והשראה שישולב בתוך המלאכה הידנית של המוזיקאי, שיוכל לקחת את החלקים הטובים-לוקאליים שחוללו ע"י האלגוריתם ולשבץ אותם בתוך הסיפור שהוא בעצמו בונה, תוך ראיית מכלול היצירה והרגש שהוא מעוניין לשדר, דבר שלפחות בינתיים עוד קשה לבטא ע"י תכנית מחשב.

5 נספח: רקע תאורטי של המוסיקה המערבית

סעיף זה סוקר בקצרה מונחים בסיסיים מהטרמינולוגיה המוסיקלית שמוזכרים במסגרת הסמינר.

1. שנים עשר הצלילים של הסולם הכרומטי

גובה צליל (pitch) מוגדר ע"י תדירות גלי הקול של הצליל – ככל שהתדר של הצליל גבוה יותר, כך הוא גם נשמע גבוה יותר (ולחפץ). בהמשך נקצר ונאמר פשוט 'צליל' כאשר הכוונה היא לגובה צליל. **מרווח** הוא היחס שבין שני (גבהי) צלילים. מקובל למדוד יחס זה ביחידות של טון (או חצאי-טון). **חצי-טון** הוא המרווח הקטן ביותר הנמצא בשימוש במוסיקה הטונאלית המערבית, והוא מבטא יחס של $\sqrt[2]{2}$ בין תדירויות גלי הקול שני צלילים הסמוכים זה לזה (דהיינו צלילים הנמצאים במרחק של חצי-טון זה מזה). בתרבויות אחרות (למשל במוסיקה הודית ובמוסיקה ערבית) משתמשים אף במרווחים קטנים יותר של רבעי-טונים.

מרווח של שנים-עשר חצאי טונים נקרא **אוקטבה**. בהתאם ליחס בין תדירויות גלי קול המוגדר ע"י מרווח של חצי-טון, מרווח האוקטבה מגדיר יחס של $1:2 = \sqrt[2]{2}^{12}$, דהיינו תדירות כל צליל היא בדיוק פי-שתיים מתדירות הצליל הנמוך ממנו באוקטבה. צלילים הנמצאים במרווח של אוקטבה בדיוק זה מזה דומים מאוד אחד לשני, עד כדי כך שבמוסיקה המערבית מסווגים אותם לאותו צליל בסיסי. בהתאם לכך, מבחינים בין סה"כ שנים-עשר צלילים בסיסיים, כך שכאשר מתייחסים לאיזשהו צליל בסיסי, הכוונה היא בעצם למשפחה או מחלקה של צלילים, המכילה איזשהו צליל בסיסי מסוים ואת כל הצלילים הנבדלים ממנו בכפולות שלמות של אוקטבה.

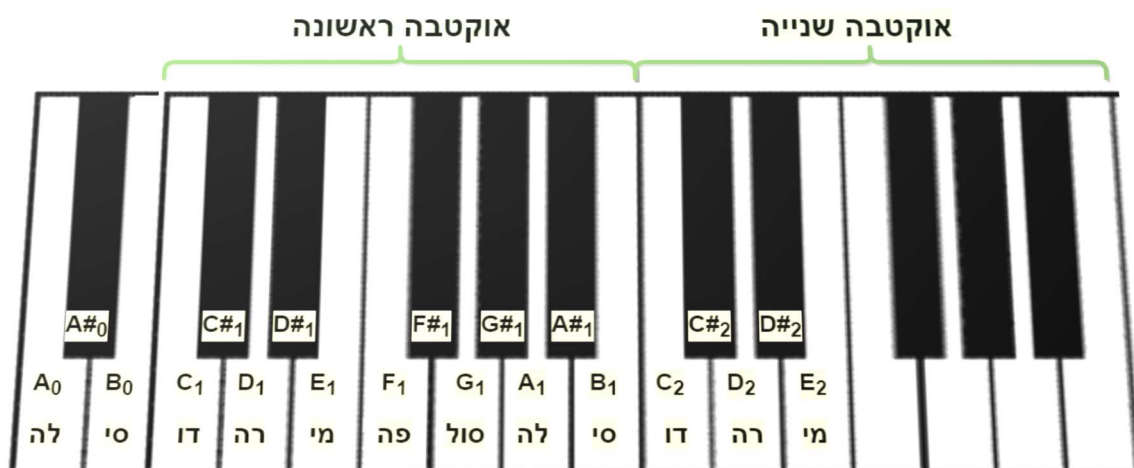
את שנים-עשר הצלילים הבסיסיים נהוג לסמן ע"י שבעת הסמלים לטיניים A, B, C, D, E, F, G (ובהתאמה בעברית: לה, סי, דו, רה, מה, פה וסול) וסימני התק \sharp (דיאז) ו- \flat (במול) המציינים שינוי של חצי-טון כלפי מעלה או מטה, בהתאמה. לפיכך את שנים-עשר הצלילים הבסיסיים נוכל לרשום באופן שקול כך –

- A, A#, B, C, C#, D, D#, E, F, F#, G, G#

או כך –

- A, Bb, B, C, Db, D, Eb, E, F, Gb, G, Ab

(ובהתאמה: לה, לה-דיאז, סי, דו, דו-דיאז, סול-דיאז או לה, סי-במול, סי, דו, דו-רה-במול, לה-במול). בכדי להתייחס לאיזשהו גובה (תדר) קונקרטי של אחד משנים-עשר הצלילים הבסיסיים, מה שמכונה גם גובהו האבסולוטי, נוהגים לחלק את כלל הצלילים לסדרות של שנים-עשר צלילים עוקבים, החל מ-C (דו), ולסדרן על-פי מיקומן היחסי מנמוך לגבוה: בפסנתר סטנדרטי בעל שמונים ושמונה קלידים, סדרת שנים-עשר הצלילים המופקים ע"י שנים-עשר הקלידים התואמים המתחילים מקליד ה-C השמאלי ביותר במקלדת (שהוא גם ה-C הנמוך ביותר במקלדת) נקראת האוקטבה הראשונה. הסדרה העוקבת לה מימין נקראת האוקטבה השנייה, וכן הלאה. באופן זה, ניתן לסמן גובה-צליל קונקרטי ע"י ציון הסמל המייצג את הצליל הבסיסי, סימן התק במידה ונדרש, ואינדקס המציין את מספר האוקטבה הרלוונטי, למשל – C1 מייצג את גובה הצליל C הנמוך ביותר במקלדת הפסנתר, C2 את גובה הצליל C הגבוה ממנו באוקטבה אחת וכן הלאה, להלן המחשה של תוויות שמות/סמלי הצלילים המופקים ע"י 20 הקלידים הראשונים של מקלדת פסנתר סטנדרטית בעלת 88 קלידים –



מנעד הוא היקף של איזשהו טווח של צלילים. לפיכך מנעד של מקלדת פסנתר סטנדרטית הוא 88 חצאי טון.

2. תווים

בכדי לייצג ו/או לתאר יצירה מוסיקלית באופן גראפי על הכתב/דפוס, משתמשים בתווים. **תווים** הם סמלים המתארים גובה צליל או הפסקה (שקט) וכן את משך שהיית הצליל/הפסקה, אשר נקרא גם הערך הריטמי של הצליל/הפסקה. ברישום ניתן לציין גם תכונות הקשורות לצורה המיועדת להפקת הצליל, כמו דרגות דינאמיקה (חוזק/הדגשה) ותחושת מהירות/זרימה הכללית. התווים עצמם רשומים על גבי **חמשה**, מערכת של חמישה קווים אופקיים מקבילים.

2.1. גובה צליל

גובה צליל מיוצג ע"י גובה יחסי תואם בחמשה ביחס למפתח מסוים, עם סימן התק מתאים ע"פ הצורך. להלן דוגמה של רישום הצלילים החל מ-A3 (לה באוקטבה השלישית) ועד ל-G#5 (סול דיאז באוקטבה החמישית) על חמשה לפי מפתח סול⁴:



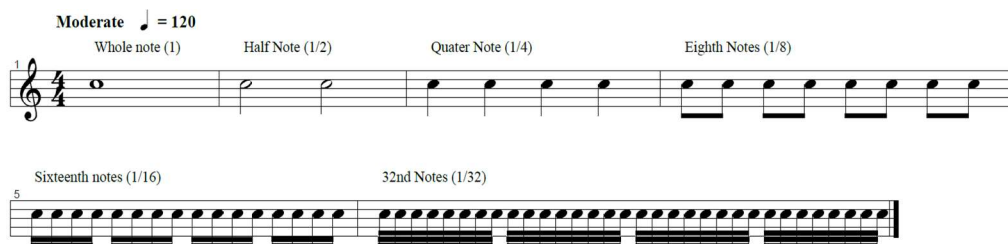
2.2. תיבות, קצב, משקל וערך ריתמי (משך שהייה קצבי)

התווים בחמשה מחולקים ל**תיבות**, לרוב באורך קבוע, המהוות יחידות לוגיות המפרידות בין קבוצות של תווים. ניתן למספר את התיבות ולעטוף אותן בסימני חזרה והוראות בקרה המנחות לחזור עליהן ו/או לעבור למקטע מבוקש (באופן דומה להוראות מחשב כמו goto ו-repeat).

יחידת הזמן הבסיסית במוזיקה היא **פעימה**. קצב הפעימות מתואר בד"כ ע"י ציון ערך BPM^5 , המציין את מספר הפעימות שיש בדקה אחת. למשל, BPM של 120 מציין כי הקצב הוא שתי פעימות בכל שנייה.

לכל יצירה מגדירים **משקל** קצבי ע"י שבר $\frac{p}{q}$: המכנה q מציין את מכנה יחידת הזמן הבסיסית המייצגת פעימה בודדת שהיא $\frac{1}{q}$, והמונה p מציין כמה יחידות זמן בסיסיות כאלו יש בתיבה. למשל, המשקל הנפוץ $\frac{4}{4}$ מציין שיחידת הזמן הבסיסית לפעימה היא רבע (ערך המכנה הוא ארבע) וכן שיש בכל תיבה ארבעה רבעים (ערך המונה הוא ארבע).

בכדי לציין את **משך השהייה** של תו נתון, מה שנקרא **ערכו הריטמי**, מתאימים לתו סימון קצבי המתאים לאורך השהייה שלו ביחס לתיבה שלמה. למשל, במשקל של $\frac{4}{4}$, תו עם ערך ריתמי של **שלוש** מתמשך לאורך תיבה אחת שלמה (ארבעה פעימות), **חצי** מתמשך לאורך חצי תיבה (שתי פעימות), **רבע** למשך רבע תיבה (פעימה אחת בלבד), וכן הלאה. להלן המחשה של סימונים של כמה ממשיכי הצליל הבסיסיים במשקל $\frac{4}{4}$ –



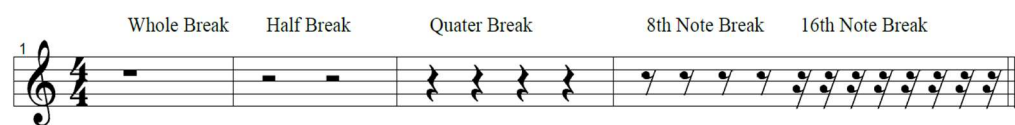
⁴ מפתח סול מציין שמיקום התו G4 (סול באוקטבה רביעית) על פני החמשה הוא השורה השנייה מלמטה בחמשה.

⁵ BPM = Beats Per Minute

מבלי להיכנס לפרטי אופן הסימון, נציין כי המכנה של יחידת זמן משך שהיית צליל אינו מוגבל דווקא לכפולות שלמות של שתיים: ישנן דרכי רישום לכל מספר רציונלי. כמו כן, תיבה לא חייבת להיות הומוגנית ביחידות אורך, ניתן לשלב בה יחידות אורך מעורבות כל עוד משך השהייה הכולל של כלל התווים יחד בתיבה תואם למשקל הקצבי המוגדר ואינו קצר ממנו או חורג ממנו.

כמו כן, ניתן להאריך את משך שהיית צליל/תו אל מעבר ליחידת הזמן הבסיסית שלו ע"י רישום מופע נוסף שלו העוקב למופע המבוקש וקישור שני המופעים בקשת חיבור. הארכה שכזו קובעת את משך השהייה הכולל של התו לסכום משכי הזמנים של התווים המחוברים בקשת. לחלופין, ניתן לציין יחידת זמן מנוקדת, שמשמעה היא משך של פי 1.5 מיחידת הזמן המקורית ללא הניקוד. ע"י שימוש בשתי הטכניקות לעיל ניתן להרכיב אינסוף ערכים ריתמיים שונים, ולייצר **סינקופות**, הדגשה של איזשהו תו ע"י ניגונו שלא על הפעימה עצמה, למשל ע"י קיצור משך השהייה של התו שקודם לתו הנוכחי, וניצול הזמן שהתפנה כדי להקדים טיפה את הנגינה של התו הנוכחי, פעולה שמעוררת אפקט של הפתעה ועניין.

עבור הפסקות (שהיות של שקט) בין צלילים יש סימונים קצב ייעודיים, להלן הסימונים של הפסקות ביחידות זמן הבסיסיות מעל משקל 4/4:



3. הרמוניה ומלודיה

באופן כללי, **מלודיה** (מנגינה) מוגדרת ע"י רצף של צלילים המנוגנים אחד אחרי השני (לחוד), ו**הרמוניה** מוגדרת ע"י רצף של צלילים המנוגנים יחדיו (במקביל). ההרמוניה מלווה את המלודיה.

4. סוגי מרווחים

כאמור, מרווחים במוסיקה המערבית מרווחים נמדדים ביחידות של טון (וחצאי טון). היחס הנקבע ע"י מרווח מסוים משרה איזושהי תחושה כאשר המרווח מנוגן באופן הרמוני (דהיינו כאשר שני הצלילים היוצרים את המרווח מנוגנים יחדיו במקביל). תחושה זה יכולה להיות מסווגת כנעימה (קונסוננטית), צורמת (דיסוננטית) או זכה (צלולה, טבעית). בהתאם, המרווחים מסווגים למחלקות מתאימות. להלן רשימה של המרווחים הבסיסיים הכלולים באוקטבה אחת –

מרווח	מרחק בטונים	מחלקה
סקונדה קטנה	חצי טון	מרווחים דיסוננטיים
סקונדה גדולה	טון	מרווחים דיסוננטיים
טרצה קטנה	טון וחצי	מרווחים קונסוננטיים
טרצה גדולה	שני טונים	מרווחים קונסוננטיים
קוורטה	שניים וחצי טונים	מרווחים זכים
טריטון	שלושה טונים	מרווחים דיסוננטיים
קווינטה	שלושה וחצי טונים	מרווחים זכים
סקסטה קטנה	ארבעה טונים	מרווחים קונסוננטיים
סקסטה גדולה	ארבעה וחצי טונים	מרווחים קונסוננטיים
ספטימה קטנה	חמישה טונים	מרווחים דיסוננטיים
ספטימה גדולה	חמישה וחצי טונים	מרווחים דיסוננטיים
אוקטבה	שישה טונים (שנים-עשר חצאי טונים)	מרווחים זכים

הערה: המרווחים נמדדים הן במרחק ביחידות של גובה הצליל האבסולוטי (חצאי טונים) והן בסימול התווי ע"י המרחק בין התווים בסולם הכרומטי, למשל המרחק האבסולוטי בין דו-דיאז לבין מי-במול הוא טון, אולם מאחר שבסולם הכרומטי של שנים-עשר הצלילים, יש בין דו למי את התו רה, המרווח הזה של טון לא יסווג כסקונדה גדולה, אלא כטרצה קטנה מוקטנת.

5. סולמות, טונאליות ודרגות הרמוניות

סולם הוא סדרה מתוזרית של מרווחים, הקובעת איזשהו יחס סדר בין צלילים, שמרחקם זה מזה מוכתב ע"י סדרת המרווחים של הסולם. במוסיקה הטונאלית המערבית, בכל סולם שכזה יש צליל אחד שהוא "השליט" בסולם, הוא נשמע הכי טוב וטבעי בקונטקסט רצף הצלילים שבסולם: הצלילים בסולם נמצאים באינטראקציה הדדית תמידי של יצירת מתח ופתורו ע"י רגיעה, והרגיעה או אתנחתא הגדולה ביותר שכל הרצפים בסולם שואפים אליה היא כאשר מתנגן הצליל ה-"שליט" של

הסולם. צליל זה מכונה ה-"טוניקה". צליל הטוניקה משתמש גם כשורש הסולם, דהיינו כצליל שממנו מתחילים את סדר רצף המרווחים, שקובע בהתאמה את הרכב הצלילים בסולם. לאוסף הצלילים בסולם מיחסים דרגות, ע"פ מיקומם היחסי ברצף הצלילים שבסולם ביחס לטוניקה: תו הטוניקה עצמו הוא דרגה ראשונה, התו שרחוק ממנו במרחק הנקבע ע"י המרווח הראשון בסדרת המרווחים של הסולם הוא דרגה שנייה וכל הלאה. על בסיס דרגות אלו בונים **אקורדים**, רצפי שלושה תווים שונים או יותר המנוגנים יחדיו. כל סולם מכתיב אקורדים הנגזרים מהדרגות שלו כדלקמן: לכל דרגה (תו) בסולם מוסיפים את התווים שנמצאים ממנו במרחק של 3, 5 ו-7, במובנים של דרגות. לאקורדים הנוצרים באופן זה קוראים **דרגות הרמוניות**. להלן דוגמה של הדרגות ההרמוניות הנגזרות מהסולם דו-מז'ור, הנקבע ע"י סדרת המרווחים: טון-טון, חצי טון, טון, טון, חצי טון, טון, שמושרש בתו דו (כלומר, סדרת המרווחים קובעת סולם מז'ור כלשהו, וכאשר משרישים סדרה זו בדו בתור צלילי הטוניקה, מתקבל הסולם דו-מז'ור) –

I - Cmaj7	II - D minor 7	III - E Minor 7	IV - F maj7	V - G7	VI - A minor	VII - B minor7 b5

לפיכך, כל סולם מכתיב אוסף צלילים ואקורדים. ישנם סולמות שונים בעלי תתי-קבוצות משותפות של צלילים, לכן בהינתן אקורד מסוים, לא ניתן לבצע את המיפוי בכיוון ההפוך ולשייכו באופן חד-חד ערכי לסולם ספציפי, כי אם למספר סולמות שונים.

6 רשימה ביבליוגרפית

- [1] Biles, J.A., *Straight-Ahead Jazz with GenJam: A Quick Demonstration*, in *Musical Metacreation Workshop: AAAI Ninth Conference on Artificial Intelligence and Interactive Digital Entertainment*, (Northeastern University, US, 2013), AAAI.
- [2] Dreier, J. *Algorithmic Music Composition*, Topics in Computer Music, 2015. Retrieved June 08, 2019, from RWTH AACHEN University: <http://hpac.rwth-aachen.de/teaching/sem-mus-15/reports/Dreier.pdf>
- [3] Fernández, J.D., and Francisco, V. *AI methods in algorithmic composition: A comprehensive survey*. *Journal of Artificial Intelligence Research*, 48 (2013), 513-582.
- [4] Jaehun, J., Yusun, K. and Chang, W.A. *A multi-objective evolutionary approach to automatic melody generation*, *Expert Systems with Applications*, 90 (2017), 50-61.
- [5] Kramer, O. *Genetic Algorithm Essentials*. Springer, Cham, 2017.
- [6] Maier, H.R., Razavi, S., Kapelan, Z., Matott, L.S., Kasprzyk, J. and Tolson, B.A. *Introductory overview: Optimization using evolutionary algorithms and other metaheuristics*, *Environmental Modelling & Software*, 114 (2019), 195-213.
- [7] Pavlov, S., Olsson, C., Svensson, C., Anderling, V., Wikner, J. and Andreasson, O. *Generation of music through genetic algorithms*, 2014. Retrieved July 27, 2019, from University of Gothenburg: <http://hdl.handle.net/2077/37004>