מבוא

נושאו של קורס זה הוא מושג החישוב, סיבוכיותו וענינים בסיסיים הקשורים בו. מושגים ותוצאות אילו נמנים על ההישגים המחשבתיים החשובים של המאה ומהווים את היסוד למדעי המחשב. דיון מעמיק בהם יבהיר את הפער בין מושג החישוב ומושגים כגון מחשבה ואינטואיציה. כמו כן יובהר כי לא כל משימה הניתנת להגדרה ניתנת גם ל"פתרון אוטומטי".

לקורס שני חלקים. החלק הראשון יעסוק במושג החישוב, בעוד שהחלק השני יתמקד בחישובים יעילים. בפרט נסקור בחלק הראשון מודלים פורמליים של חישוב (כלומר, מחשבים) ובעיקר את המודל של מכונת טיורינג. נדגיש מצד אחד את החשיבות של שימוש במודל מדויק (ע"מ להמנע משגיאות הנובעות מהסחפות אחר אינטואיציה בלתי מבוססת), ומצד שני את אי החשיבות של המודל הספציפי שישמש אותנו בדיון, ושל פרטים טכניים משניים הקשורים בו. ינתן ביסוס תאורטי למושגים כגון general purpose machine, ושימוש בסברוטינה. בסיום חלק זה נראה שורה של בעיות מענינות אשר לא ניתן לכתוב תכנית מחשב אשר פותרת אותן.

בחלק השני של הקורס נדון במושג סיבוכיות החישוב ונתמקד בחישובים המוגבלים במשאבים. בפרט נדון בחישוב המוגבל ל"זמן סביר". המושג האינטואיטיבי של "חישוב יעיל" יזוהה עם המושג הפורמלי של "חישוב בזמן פולינומיאלי". האבחנה העיקרית בחלק זה היא בין "בעיות אשר ניתן לפתור אותן בזמן פולינומיאלי" לבין "בעיות אשר ניתן לודא את נכונות הפתרון להן (כשהוא נתון) בזמן פולינומיאלי". אבחנה זו ידועה כאבחנה בין המחלקות **P** ו- **NP**. האם אבחנה זו היא תחבירית או מהותית? (ז.א. האם **P** = **NP**?) לצערינו לא ידועה עד כה תשובה לשאלה, וההשערה המקובלת היא כי מדובר בשוני מהותי (ז.א. **P** ≠ **NP**). נראה שורה של בעיות טבעיות אשר, תחת ההנחה **P** ≠ **NP**, נמצאות ב **NP** – **P**. בעיות "קשות" (במובן שיוגדר בהמשך) אלו, יקראו בעיות **NP**-שלמות. בפרט, הבעיות, "האם בגרף נתון קיים מעגל המילטוני?", "האם פסוק נתון (בתורת הפסוקים) ניתן לסיפוק?", הן **NP**-שלמות.

ספרי לימוד שניתן להעזר בהם לגבי החלק הראשון של הקורס:

(1) Introduction to Automata Theory Languages and Computation / Hopcroft, Ullman

(2) Elements of the Theory of Computation / Lewis, Papadimitriou

ואילו לגבי החלק השני:

(3) Computers and Interactability A Guide to the Theory of NP-Completeness / Garey, Johnson

חוברת זו נכתבה בעיקר עקב אי הסכמתו של המחבר עם הגישה הרעיונית של ספרים אלו (בעיקר הכוונה ל-(1) ו-(2)), גישה הבאה לידי ביטוי בדגשים השונים המוענקים לנושאי הלימוד, ובמשמעות הניתנת (או לא ניתנת) לנושאים אלו, אם במפורש ואם במובלע. לפיכך אין להתפלא על אי התאמות באופן הרצאת החומר בין חוברת זו והמקורות הנ"ל.

מספר הערות לגבי קריאת החומר:

[1] הקורס באוטומטים אינו מהווה קדם לקריאת החומר, כמו שאינו מהווה קדם מהותי לקורס בחישוביות. מה שנרוויח מן הקורס באוטומטים הינו הכרות עם המושגים של א"ב, מילה, שפה ומחלקת שפות. מושגים אלו פשוטים ומוצגים בפרק הרקע הטכני, כך שהרווח הזה זניח, בפרט כשחושבים על הנזק הקונספטואלי הנגרם ע"י הבנה פזיזה של הקורס באוטומטים (אשר עלולה להביא למחשבה כי אוטומט סופי מהווה מודל סביר וכללי של חישוב - ראה הערות סיכום לחלק הראשון). לפיכך מומלץ לקורא "לשכוח" את מה שלמד בקורס ההוא.

[2] חלוקת הפרקים אינה משקפת את החשיבות היחסית שלהם או את הקושי היחסי שלהם, אלא את הארגון הקונספטואלי של החומר בלבד. להלן דעת המחבר על חשיבות הפרקים השונים:

פרקים 1,2: יסוד הקורס כולו, בעל חשיבות עצומה.

פרקים 3,4,5: יציאה ידי חובה.

פרק 6: חשוב בעיקר בגלל הניסוחים השונים של התזה של צ'רץ'.

פרק 7:

סעיפים 7.1 , 7.2 - חשובים מאד.

סעיפים 7.3 , 7.4 - חשובים בעיקר כאנלוגיה לנעשה בפרקים 14 , 13.

סעיף 7.5 - סתם משחק.

פרקים 8,9,10: בעלי חשיבות עצומה.

פרק 11: חשוב.

פרקים 12,13,14: בעלי חשיבות עצומה.

פרקים 15,16,17: חשובים מאד.

פרק 18: פרק מתקדם, מעבר לקורס הרגיל.

פרקים וסעיפים המסומנים ב-* חשובים, אך ניתן לדלג עליהם ללא פגיעה בעיקר.

[3] הערות מתקדמות הן תמיד הערות שקופצות קצת קדימה בחומר, והסטודנט המתקשה יכול להתעלם מהן במידה והן מבלבלות אותו בשלב זה במקום להבהיר. מצד שני יתכן שהקפיצה קדימה יכולה לסייע אפילו למתקשה. לפיכך מומלץ לקורא לנסות לקרוא אותן, ולהתעלם מהן אם אין בהן תועלת לגביו.

אנו מודים לאייל קושלביץ על כתיבת פרק 6.1, וכן לסילביו מיקאלי ולאוניד לוין על שיחות ודיונים הנוגעים לנושא החישוב היעיל. כמו כן אנו מודים ללי-בת נלסון, ארז פטרנק וישי בן-ארויה על הפנית תשומת ליבנו לטעויות בגרסאות קודמות.

רקע טכני ומושגים ידועים

(נניח הכרות עם החומר שכאן, ונחזור למען הסדר הטוב)

I. רלציות, פונקציות וקבוצות בנות מניה

תהיינה S_1 ו-S_2 קבוצות סופיות או אינסופיות.

רלציה: רלציה (יחס) דו-מקומית בין S_1 ל-S_2, היא תת-קבוצה של המכפלה הקרטזית $S_1 \times S_2$. רלציה $R \subseteq S_1 \times S_2$ תקרא מלאה, אם לכל $a \in S_1$ קיים $b \in S_2$ כך ש- $(a,b) \in R$. רלציה שאיננה מלאה תקרא חלקית. לעיתים נסמן aRb במקום $(a,b) \in R$.

פונקציה: פונקציה f מ S_1 ל S_2 $(f : S_1 \to S_2)$ היא רלציה אשר בה לכל $a \in S_1$ יש לכל היותר $b \in S_2$ יחיד המקיים $(a,b) \in f$. במקרה זה נכתוב $b = f(a)$.
הפונקציה f נקראת מלאה, אם לכל $a \in S_1$ קיים $b \in S_2$ כך ש- $(a,b) \in f$. (או במילים אחרות $f(a)$ מוגדר). פונקציה שאיננה מלאה תקרא חלקית (ז.א. קיים $a \in S_1$ כך ש- $f(a)$ אינו מוגדר).

קבוצה אינסופית S נקראת בת-מניה אם קימת העתקה 1-1 ועל, בינה לבין המספרים הטבעיים N (כלומר אם ניתן למנות את אבריה בצורה סדרתית).

משפט: קבוצת הממשיים באינטרוול (0,1) אינה בת-מניה.

הוכחה: נניח בשלילה כי הממשיים בקטע הנ"ל הם קבוצה בת מניה (ז.א. שקיימת העתקה 1-1 ועל בינם לבין הטבעיים), ונראה קיום של מספר ממשי שאינו מותאם לאף מספר טבעי. מספר זה נבנה באופן הבא: נניח קיום התאמה כנ"ל ונסמן ב- r_i את המספר הממשי המותאם למספר הטבעי i. נבנה מספר ממשי r' כך שהספרה ה i-ית ביצוג העשרוני של r', היא 4 (זוהי בחירה שרירותית בלבד) אם הספרה ה i-ית ביצוג העשרוני של r_i שונה מ-4, ואחרת (הספרה ה i-ית של r_i הינה 4) הספרה ה i-ית של r' תהיה 5 (שוב בחירה שרירותית). ראה להלן דוגמא

$r_1 = .47395\cdots$

$r_2 = .47369\cdots$

$r_3 = .25471\cdots$

$r_4 = .37111\cdots$

$r_5 = .98989\cdots$

.

.

$r' = .54544\cdots$

לא יתכן ש- r' מותאם למספר טבעי כלשהו, וזאת משום שאם r' מותאם ל-j אז על פי הבניה הביט ה j-י ביצוג הבינארי של r' שונה מן הביט ה j-י ביצוג הבינארי של r_j. קיבלנו סתירה להנחת קיום ההתאמה הנ"ל.

<u>הערה:</u> שיטת הוכחה מתמטית זו נקראית **שיטת לכסון**, והיא שימושית במספר רב של הוכחות הבאות להוכיח קיום אי-שיוויון בין עוצמות של קבוצות.

II. א"ב, מילים, שפות ומחלקות

<u>הגדרות:</u>

<u>א"ב</u> היא קבוצה סופית, אבריה נקראים אותיות.

<u>מילה</u> (מעל א"ב נתון) היא סדרה סופית של אותיות (ז.א. זאת קבוצה סדורה של אברים בא"ב).

<u>שפה</u> (מעל א"ב נתון) היא קבוצה (סופית או אינסופית) של מילים.

<u>מחלקת שפות</u> (מעל א"ב נתון) היא קבוצה (סופית או אינסופית) של שפות.

<u>שים לב!</u> כל אחד מן המושגים הנ"ל הוא מקטגוריה שונה. (במיוחד שים לב שמחלקה היא קבוצה של קבוצות של מילים).

לשם הנוחיות נניח קיום סדר של אברי הא"ב (כלומר נניח קיומה של רלציה טרנזיטיבית ולא רפלקסיבית "<" בין אברי הסדרה).

נגדיר **סדר לקסיקוגרפי קאנוני** על מילים מעל א"ב Σ:

יהי $w_1,w_2\in\Sigma^*$ נכתוב $w_1<w_2$ כאשר $|w_1|<|w_2|$ (האורך של w_1 קטן מהאורך של w_2) או כאשר, $|w_1|=|w_2|$, וכן מתקיים $w,w_1',w_2'\in\Sigma^*$ (כך ש- $w_2=w\tau w_2'$ ו- $w_1=w\sigma w_1'$ וכן $\sigma,\tau\in\Sigma$) וגם $\sigma<\tau$.
כלומר, מילים ארוכות יותר הן גם "גדולות" יותר, ובין מילים בעלות אורך שווה המילה הקודמת בסדר המילוני "קטנה" יותר.

נראה כי השפה Σ^* (השפה המכילה את כל המילים מעל Σ) היא קבוצה בת-מניה, ע"י שנראה העתקה f חד-חד ערכית ועל ממנה לטבעיים (כלומר נראה "מספור" של אברי השפה Σ^*):

המילה הריקה תקבל את המספר 1.

המילה $w=\sigma_n\cdots\sigma_2\sigma_1$ תקבל את המספר:

$$f(w)=1+\sum_{i=0}^{|w|-1}|\Sigma|^i+\sum_{i=1}^{|w|}(\mathrm{ord}(\sigma_i)-1)\cdot|\Sigma|^{i-1}$$

כאשר ord הינה פונקציה אשר מחזירה עבור האות ה k-ית בא"ב את הערך k.

<u>הסבר:</u> ה-1 הוא עבור המילה הריקה, הביטוי שאחריו סופר את כל המילים הקצרות מ-w והביטוי האחרון הוא ערכו המספרי של w (לפי בסיס $|\Sigma|$).

<u>תרגילים:</u>

(1) הוכח כי כל שפה אינסופית היא קבוצה בת-מניה.

(2) הוכח כי קיימות מחלקות-שפות שאינן קבוצות בנות-מניה (ואינן סופיות).
(רמז: הראה התאמה חד-חד ערכית ועל בין מחלקת שפות לבין הממשיים באינטרוול $[0,1]$).

(3) האם מחלקת כל השפות הסופיות היא בת-מניה? דייק בטיעונך!
(רמז: סדר את השפות, על פי האורך הכולל של אבריהן).

<u>סימונים:</u>

$|w|$ - האורך של w.

Σ^* - אוסף כל המילים מעל הא"ב Σ.

III. קידוד חד-פענח

תהיינה Σ_1 ו-Σ_2 שתי קבוצות סופיות (א"ב). נאמר כי h היא קידוד חד-פענח של Σ_1 ע"י Σ_2 אם מתקיים:

(1) $h : \Sigma_1 \to \Sigma_2^*$

(2) לכל $\sigma_1, \ldots, \sigma_n, \tau_1, \ldots, \tau_m \in \Sigma_1$ אם $h(\sigma_1) \cdots h(\sigma_n) = h(\tau_1) \cdots h(\tau_m)$ אז $n = m$ וגם לכל $1 \le i \le n$ מתקיים, $\sigma_i = \tau_i$.

ההרחבה של h למילים (מעל Σ_1^*) מוגדרת באופן טבעי כדלקמן:

$$h(\sigma_1 \cdots \sigma_n) = h(\sigma_1) \cdots h(\sigma_n)$$

מקרה פשוט של קידוד חד-פענח הוא קוד פרפיקסי. בקוד כזה לכל $\sigma \neq \tau$ מתקיים $h(\sigma)$ אינו פרפיקס (רישא) של $h(\tau)$. מקרה פשוט של קוד פרפיקסי הוא קידוד ע"י מילים שוות אורך: דהיינו לכל $\sigma,\tau \in \Sigma_1$ מתקיים $|h(\tau)| = |h(\sigma)|$ (וכמובן h היא חד-חד ערכית).