8. מכונה אוניברסלית

הגישה שנקטנו עד כה ביחס לחישוב היתה חקירתו דרך מודל חישוב שאינו אלא אוסף אינסופי של מכונות. כל מכונה מחשבת פונקציה אחת ויחידה. על מנת לחשב פונקציה שידוע כי היא ניתנת לחישוב, היה צורך לבנות מכונה ספציפית (במודל) אשר מחשבת אותה. מבחינה מעשית הדבר דומה לבנית מחשב (special purpose) עבור כל משימה. אנו יודעים כמובן שאין צורך לבנות מחשב לצורך כל משימה, ודי לבנות מחשב (general purpose) הניתן לתכנות באופן שיתאים לכל משימה. גישה זו שאומצה ב-40 השנים האחרונות (החל ממודל החישוב של פון-נוימן), יסודה בתגליתו של טיורינג (Turing) על קיומן של מכונות אוניברסליות. תגלית זו נוסחה ותנוסח גם כאן במונחים של מכונות טיורינג, אך היא תקפה לגבי כל מודל סביר וכללי של חישוב.

8.1 קידוד וסימולציה

בבסיס תגליתו של טיורינג ישנן שלוש אבחנות:

(1) מכונות טיורינג ניתנות לתיאור סופי. (אבחנה זו תשאר תקפה לגבי כל מודל סביר של חישוב, שכן "הגדרנו" מודל כזה, כמכיל מכונות שכל אחת מהן ניתנת לתיאור סופי).
מבין כל התאורים האפשריים של מכונת טיורינג נבחר (שרירותית) בתאור הבא: תהי $M = (Q, q_0, F, \Gamma, \Sigma, \not{b}, \delta)$, מכונת טיורינג ונניח ללא הגבלת הכלליות כי, $Q = \{1,2,3,\ldots,|Q|\}$, $q_0 = 1$, $F = \{2,3\}$, $\Sigma = \{0,1\}$, $\Gamma = \{0,1,2\}$ (2 מקודד את הרווח - $\not{b}$).
המכונה M, תקודד ע"י המחרוזת, $\delta_1 00 \delta_2 00 \cdots 00 \delta_t 000$ כאשר δ_i, היא מחרוזת מן הצורה, $1^q 01^\sigma 01^p 01^\tau 01^d$, כאשר $\delta(q,\sigma) = (p,\tau,d)$ (L מקודד כ-$d=1$ ואילו R מקודד כ-$d=2$). שים לב כי המחרוזת מקודדת במפורש, רק את פונקצית המעברים δ וכל שאר רכיבי המכונה מוסקים מתוך תאור זה (ומתוך ההסכם על F, Σ, Γ ו- Q).

(2) הקונפיגורציות הרגעיות של מכונת טיורינג, ניתנות לתאור סופי (גם אבחנה זו תקפה לגבי כל מודל סביר של חישוב). נבחר לקודד את הקונפיגורציה $\sigma_1 \cdots \sigma_n q \sigma_{n+1} \cdots \sigma_m$ $(q \in Q,\ \sigma_1, \ldots, \sigma_m \in \Gamma)$ ע"י המחרוזת $1^{\sigma_1+1} 01^{\sigma_2+1} 0 \cdots 1^{\sigma_n+1} 001^q 01^{\sigma_{n+1}+1} 0 \cdots 1^{\sigma_m+1}$. שים לב כי מקום הראש מזוהה ע"י שני אפסים סמוכים.

(3) בהנתן תיאור (קידוד) של מכונת טיורינג וקונפיגורציה רגעית של המכונה, ניתן לחשב את הקונפיגורציה העוקבת של אותה מכונה. במילים אחרות, הפונקציה f, המוגדרת ע"י $f(<M>,\alpha) = \beta$ כאשר α היא קונפיגורציה רגעית של M, ומתקיים $\alpha \overset{M}{\vdash\!\!-} \beta$, ניתנת לחישוב. כלומר, ישנה מכונת טיורינג, אשר בקבלה כקלט קידוד של מכונת טיורינג כלשהי, $<M>$, וקונפיגורציה רגעית של אותה מכונה (כלומר של M) מחשבת את הקונפיגורציה העוקבת של M. שים לב כי f היא פונקציה חלקית. f אינה מוגדרת כאשר α אינה קונפיגורציה רגעית של M או כאשר α הינה קונפיגורציה סופית של M.

תרגיל 8.1: כתוב מכונת טיורינג אשר מחשבת את הפונקציה f הנ"ל.

הדרכה: נוח כמובן לכתוב מכונה N, בעלת שני סרטים, המחשבת את f. בהנתן קלט $(\langle M\rangle,\alpha)$, המכונה N תעתיק את $\langle M\rangle$ (תאור המכונה M) לסרט השני, תבדוק אם α יכולה להיות מבחינה תחבירית קונפיגורציה רגעית של M (כלומר, האם α מהווה קידוד חוקי המתאים לקונפיגורציה ב-M), תסרוק את α בחיפוש אחר מקום הראש, תמצא את החמישיה המתאימה בתאור פונקצית המעברים של M (שינוי מצב, הזזת ראש, שינוי אות סרט) ותבצע את השינוי המתבקש. למעשה זוהי פשוט הפעלת מכונה אחת ע"י מכונה שניה ולמעשה כבר עשינו דברים דומים באופן סמוי בעבר.

הערה: שים לב שגם את האמור ב-(3) ניתן להכליל על כל מודל סביר וכללי של חישוב.

האבחנות (1), (2) ו-3), מובילות באופן טבעי להגדרת הפונקציה האוניברסלית ולטענה שפונקציה זו ניתנת לחישוב.

בניה 8.1: **הפונקציה האוניברסלית**, U, היא פונקציה חלקית בעלת שני ארגומנטים, המוגדרת כדלקמן: $U(\langle M\rangle,x)=f_M(x)$, כאשר x הינו בא"ב הקלט של M, המכונה M עוצרת על קלט x עם פלט $f_{M(x)}$. הפונקציה, U, אינה מוגדרת על הזוג $(\langle M\rangle,x)$, כאשר x אינו בא"ב הקלט של M, או כאשר M אינה עוצרת על x.

משפט 8.1: הפונקציה האוניברסלית U ניתנת לחישוב.

ההוכחה נובעת כמעט מיידית מתרגיל 8.1. כל מה שצריך לעשות הוא לבדוק האם x שייך לא"ב הקלט של M, להעביר את הקלט $(\langle M\rangle,x)$ אל $(\langle M\rangle,q_0x)$, כאשר q_0 הוא המצב ההתחלתי של M ולהפעיל את N (מתרגיל 8.1) בצורה איטרטיבית (כל פעם על הפלט הקודם שלה) עד לעצירה. שים לב כי במקרה ש-x בא"ב של מכונה M, המכונה האוניברסלית תעצור אמם M עוצרת על x. פרטי ההוכחה מושארים כתרגיל.

כל מכונה אשר מחשבת את הפונקציה האוניברסלית U תקרא **מכונה אוניברסלית**. המכונה האוניברסלית שאנחנו בנינו, עובדת ע"י סימולציה של המכונה שאת תאורה היא מקבלת כקלט וסימולציה זו נעשית צעד אחר צעד. אין חס וחלילה להסיק מכך שכל מכונה אוניברסלית חייבת לפעול באופן כזה (למרות שזו דרך פעולה מאד סבירה). בפרט ישנה למשל מכונה אוניברסלית הפותחת ב-1000 מסעי סרק ואחר כך מבצעת סימולציה של צעדי המכונה הניתנת כקלט. ישנה מכונה אוניברסלית המבצעת את הסימולציה "בקפיצות". כלומר, כל פעם קוראת סביבה של מקום הראש של M (המכונה הניתנת כקלט) ומבצעת מספר צעדים של M "במכה אחת".

הערה: בכל הדיון הנ"ל הנחנו כי המכונה M, שתאורה ניתן כקלט לפונקציה האוניברסלית, היא מכונה המקבלת קלט יחיד. הנחה זו מוצדקת בכך שניתן לקודד כל סדרה של קלטים $(x_1,x_2,\ldots,x_n)$ לקלט יחיד $x_1\#x_2\#\cdots\#x_n$. אם

$x_1, \ldots, x_n \in \{0,1\}^*$ ואם רוצים שגם הקלט היחיד המתקבל, יהיה בקידוד בינארי, ניתן להשיג זאת ע"י $d(x_1)01d(x_2)01 \cdots 01d(x_n)$, כאשר $d(\alpha)$ הוא קידוד המכפיל כל אות ב- α. כלומר, $d(\sigma_1 \cdots \sigma_l) = \sigma_1\sigma_1 \cdots \sigma_l\sigma_l$. אנחנו נעדיף להתעלם מפרטי הקידוד ולרשום $(x_1, \ldots, x_n)$ בהבנה שזהו קידוד במחרוזת בינארית אחת של סדרה בת n מחרוזות. הערה זו חשובה, משום שבעתיד נתעניין בערך הפונקציה האוניברסלית, על זוגות $(<M>,x)$, כאשר M היא מכונה אוניברסלית.

מומלץ מאד לפתור את תרגיל 8.2, לפני המשך הקריאה.

<u>תרגיל 8.2:</u> תהי M_U מכונה אוניברסלית כלשהי.

א. הוכח: $U(<M_U>,(<M>,x)) = U(<M>,x)$.

ב. האם הפלט של M_U על קלט $(<M_U>,(<M>,x))$ זהה לפלט של M_U על קלט $(<M>,x)$?

ג. תהי M מכונת טיורינג כלשהי, המקיימת, $U(<M>,(<M'>,x)) = U(<M'>,x)$ לכל מכונה M' ולכל מחרוזת x. הוכח כי M הינה מכונה אוניברסלית.

ד. תהי M מכונת טיורינג כלשהי, המקיימת, $U(<M>,(<M>,x)) = U(<M>,x)$ לכל מחרוזת x. האם M הינה מכונה אוניברסלית? במידה ולא, תן דוגמא נגדית.

ההשלכות המעשיות והתאורטיות של קיום מכונות אוניברסליות הן מרחיקות לכת עד כדי כך, שקשה להבחין בהן, משום שהשלכות אלו נהפכו מכבר לחלק ממה שאנו מקבלים כמובן מאליו (לדוגמא, מחשב general purpose). כדי להבחין בהשלכות, רצוי לעבור לדיון בנושא שלגביו אין לנו עדיין הרגלים. נושא שכזה, שהינו בעל חשיבות בפני עצמו, הינו סיבוכיות קולמגורוב.

8.2 סיבוכיות קולמגורוב

המוטיבציה להגדרת הסיבוכיות של קולמגורוב, היא שאלה העומדת בבסיס הפילוסופיה של המדע והיא השאלה בדבר אופיו של "הסבר מדעי". אינטואיטיבית, הסבר שכזה צריך להיות <u>תמציתי</u> ועם זאת הוא צריך לאפשר שחזור וניבוי של כל התופעות הנצפות על ידינו. הדרישה שההסבר יאפשר שחזור וניבוי של תופעות נצפות הינה ברורה ומקובלת, כאן ברצוננו להדגיש את תמציתיות ההסבר. תמציתיות ההסבר היא המבחינה אותו מתיאור מפורט גרידא, שהרי לא יעלה על הדעת לקרוא לתאור כל מצבי העולם בעבר, בהווה ובעתיד, "הסבר", למרות שתאור כזה ודאי מאפשר שחזור וניבוי של תופעות. מצד שני, ברור שהסברים קצרים מאד יתכנו רק עבור תופעות פשוטות מאד, ואילו לתופעות מסובכות יותר, יתכנו רק הסברים ארוכים יותר. מכאן הקישור בין "סיבוכיות התופעה", לבין אורך ההסבר הקצר ביותר עבורה. לפני שנפנה לפיתוח הטכני, נשים לב כי הדיבורים על הסבר, הניחו קיום שפה שבה הסברים אלו נכתבים (בקפיצה קדימה נראה שמושג המכונה האוניברסלית, יבטיח קיום שפה אוניברסלית של

הסברים. בשפה זו ניתן יהיה לכתוב הסברים בכל שפה אחרת).

<u>הגדרה 8.1:</u> תהי M מכונת טיורינג ו- f_M הפונקציה המחושבת על ידי M. **הסיבוכיות של מחרוזת** x על-פי מכונה M, תוגדר כאורך המחרוזת הקצרה ביותר y כך ש- $f_M(y)=x$. סיבוכיות זאת תסומן ב- K_M. כלומר,

$$K_M(x)=\min_k\{\exists y\ |y|=k\ \wedge\ f_M(y)=x\}$$

אם לא קיים y, כך ש- $f_M(y)=x$, נגדיר $K_M(x)=\infty$.

המכונה M מאפשרת לעבור מההסבר y לתיאור מלא של התופעה x. העובדה שמכונה M מסוגלת לשחזר תופעה רק על סמך הסבר ארוך, אינה שוללת אפשרות קיום מכונה M' המסוגלת לשחזר את התופעה x, מתוך הסבר קצר בהרבה. יתר על כן, קל לראות כי לכל תופעה x, ולכל הסבר y, קיימת מכונה M, כך ש- $f_M(y)=x$. אם כך המצב, איך נקבע את הסיבוכיות של מחרוזת באופן בלתי תלוי במכונה הספציפית? התשובה טמונה במכונות אוניברסליות.

<u>משפט 8.2:</u> תהי M_U מכונה אוניברסלית. לכל מכונה M, קיים קבוע c, כך שלכל $x\in\{0,1\}^*$ מתקיים:

$$K_{M_U}(x)\le K_M(x)+c.$$

לפיכך, את הסיבוכיות, כפי שהיא נמדדת ביחס למכונה אוניברסלית (כלשהי), לא ניתן לשפר ע"י מעבר למכונה אחרת, אלא ע"י קבוע חיבורי שאינו תלוי במחרוזת שאת סיבוכיותה מודדים.

<u>הוכחת משפט 8.2:</u> תהי y מחרוזת באורך $K_M(x)$ כך ש- $f_M(y)=x$ (y זה מובטח על סמך הגדרת K_M). על פי הגדרת הפונקציה האוניברסלית, נובע כי $U(<M>,y)=x$. מכיוון ש-M_U הינה מכונה אוניברסלית, מתקיים $M_U(<M>,y)=x$. שים לב כי ע"י קידוד מתאים (קצת יותר מתוחכם מהקידוד שניתן בתחילת הפרק) ניתן לדאוג לכך ש- $|<M>,y|=O(|<M>|)+|y|$ (מהו קידוד זה?). לפיכך, על סמך הגדרת K_{M_U} נקבל כי $K_{M_U}(x)\le c_M+|y|$, כאשר c_M הוא קבוע התלוי רק באורך הקידוד של M. □

שתי טענות נוספות בעלות ענין הנוגעות לסיבוכיות קולמגורוב הן:

(1) הסיבוכיות של כל מחרוזת היא לכל היותר אורכה.

(2) לרוב המחרוזות סיבוכיות גבוהה (קרובה לארכן).

בשני המשפטים הבאים, ננסח ונוכיח טענות אלו במדויק. מנקודה זו ואילך, נקבע אחת ולתמיד מכונה אוניברסלית כלשהי, ונמדוד את הסיבוכיות ביחס אליה. הסיבוכיות של מחרוזת x תסומן $K(x)$ (המכונה האוניברסלית מושמטת מהסימון). שים לב כי מהגדרה 8.1 נובע כי לכל זוג מכונות שקולות, M_1 ו-M_2, מתקיים, $K_{M_1}(x)=K_{M_2}(x)$,

ובפרט טענה זו נכונה עבור כל שתי מכונות אוניברסליות. לכן, בהגדרת הסיבוכיות של מחרוזת אין משמעות איזו מכונה אוניברסלית בחרנו.

<u>משפט 8.3:</u> קיים קבוע c כך שלכל מחרוזת x מתקיים, $K(x) \le |x| + c$.

<u>הוכחה:</u> השתמש במשפט 8.2 לגבי המכונה M_{id} אשר מחשבת את פונקצית הזהות $\mathrm{id}(x) = x$. □

<u>משפט 8.4:</u> לכל זוג מספרים טבעיים n, m מתקיים: $\frac{|\{x \in \{0,1\}^m : K(x) \le n\}|}{|\{0,1\}^m|} < \frac{1}{2^{m-n-1}}$.

<u>הוכחה:</u> הנסוח המסורבל מסתיר עובדה פשוטה יותר והיא: $|\{x \in \{0,1\}^* : K(x) \le n\}| < 2^{n+1}$. עובדה זו נובעת מכך שהמיפוי של תופעות אל הסברים הוא מיפוי אחד לאחד, אחרת לא מתאפשר שחזור תופעות. במילים אחרות, יתכנו לכל היותר 2^{n+1} x-ים אשר מיוצרים ע"י המכונה האוניברסלית, על קלט שארכו אינו עולה על n. □

<u>תרגיל 8.3:</u> הוכח כי הסיבוכיות של המחרוזת xx אינה עולה על $|x| + O(1)$, וכי לרוב ה x-ים באורך נתון, גם אינה נופלת מכך.

<u>תרגיל 8.4:</u> הוכח כי הסיבוכיות של המחרוזת, 1^n, אינה עולה על $\log_2 n + O(1)$ (רמז: חשוב על מעבר מייצוג בינארי לייצוג אונרי).

<u>תרגיל 8.5:</u> הוכח כי הסיבוכיות של המחרוזת, 1^{2^n}, אינה עולה על $\log_2 n + O(1)$ (רמז: חשב מחרוזת זו מתוך n)

<u>תרגיל 8.6:</u> תהי $f : N \to N$ (N היא קבוצת המספרים הטבעיים) פונקציה רקורסיבית כלשהי (ז.א. פונקציה הנתנת לחישוב ע"י מכונת טיורינג). הוכח כי הסיבוכיות של המחרוזת, $1^{f(n)}$, אינה עולה על $\log_2 n + O(1)$.

<u>תרגיל 8.7:</u> תהי $f : \{0,1\}^* \to \{0,1\}^*$ פונקציה כלשהי, הניתנת לחישוב. הוכח כי הסיבוכיות של המחרוזת, $f(x)$, אינה עולה על $K(x) + O(1)$. (רמז: על קלט y מאורך $K(x)$, שחזר תחילה את x ואח"כ חשב את $f(x)$).

תרגיל למתקדמים:

א. תהי $f:\{0,1\}^*\to N$ פונקציה מלאה ולא חסומה (כלומר, לכל n קיימת מחרוזת x כך ש- $f(x)>n$) כלשהי הניתנת לחישוב. עבור כל n נגדיר את x_n^f בתור ה-x הקטן ביותר בסדר לקסיקוגרפי, עבורו $f(x)\geq n$. הראה כי $K(x_n^f)\leq \log_2 n + O(1)$ (רמז: על קלט n, סרוק את כל ה-x-ים בסדר לקסיקוגרפי, עבור כל אחד מהם חשב את $f(x)$ ובדוק האם $f(x)\geq n$).

ב. הוכח בהסתמך על (א) כי הפונקציה $K:\{0,1\}^*\to N$ אינה ניתנת לחישוב.
הדרכה: הראה כי K מלאה ולא חסומה. הנח בשלילה כי K ניתנת לחישוב והסתכל באברי הסדרה x_n^K.

ג. בהסתמך על (ב), הוכח כי אין מכונת טיורינג, אשר על קלט n, מוציאה כפלט את קבוצת המילים $\{x : K(x)\leq n\}$ (רמז: הנח בשלילה כי ישנה מכונה כזו והשתמש בה על מנת לחשב את הפונקציה K).

לקח למתקדמים: הפסוק "המספר הגדול ביותר אשר ניתן לתיאור ע"י פחות ממאה מילים" אינו מהווה פרדוקס (זהו פרדוקס לכאורה, מכיוון שישנו מספר סופי של מספרים "הניתנים לתיאור" ע"י פחות ממאה מילים ולכן יש בינהם גם גדול ביותר, ולעומת זאת בפסוק "המספר הגדול באחד מהמספר הגדול ביותר אשר ניתן לתיאור ע"י פחות ממאה מילים" יש פחות ממאה מילים). פסוק זה, אינו מוגדר היטב (כפי שקורה עם רוב הפרדוקסים). הוא מסתמך על אי הבהירות ביחס למושג "ניתן לתיאור". אם משמעות מושג זה הוא "ניתן לחישוב מתוך תיאור קצר", מתברר כי הפסוק מניח במובלע כי ניתן לחשב את קבוצת המספרים שיש להם תיאור באורך נתון ולהנחה סמויה זו אין בסיס. למעשה, "הפרדוקס" מוכיח זאת בצורה של נפנוף ידיים והתרגיל מוכיח זאת במדויק.