מודלים חישוביים - הרצאה 4

שרון מלטר, אתגר 17

16 באפריל 2024

בהרצאה הקודמת סיימנו לדבר על אס"דים ואסל"דים. היום נתחיל לדבר על אוטומט מחסנית.
לפני שנתחיל עם האוטומט, נתייחס רגע לביטוי שאנחנו משתמשים בו הרבה- 'זיכרון'. זיכרון יכול להיות סטטי או דינאמי. זיכרון סטטי הוא זיכרון קבוע, שלא ניתן להגדילו, ולכן הוא סופי. זיכרון דינמי הוא זיכרון שניתן להגדיל. הוא קבוע בכל זמן נתון, אך מכיוון שניתן תמיד להגדילו, הוא אינסופי.
לרוב נדבר על זיכרון דינמי.

1 אוטומט מחסנית

אוטומט מחסנית יכול להיות דטרמיניסטי או לא דטרמיניסטי, אך אנחנו נלמד רק את הגירסא הלא דטרמיניסטית (ולגירסא זו תמיד נתכוון כאשר נדבר על אוטומט מחסנית)
מדוע? מכיוון שהוא לא שקול לאוטומט מחסנית דטרמיניסטי- התאוריה של א"מ לא דטרמיניסטי יותר 'סימפטית' ואנחנו נתעניין בעיקר במה א"מ לא דטרמיניסטי לא יכול לעשות. כמובן שכל מה שא"מ לא דטרמיניסטי לא יכול לעשות גם א"מ דטרמיניסטי לא יכול.

אוטומט מחסנית יכול לבצע פעולות של $push$, pop שיכולות להכניס ולהוציא תווים ממחסנית. מותר לנו לבחור את התו הרצוי, אך הוא לא בהכרח יהיה בראש המחסנית אם נרצה להוציא אותו. כלומר, ההוצאה לא תמיד תתאפשר. אבל תכונה זו תסייע לנו להפריד בין מילים שאנו רוצים שיילכו במסלולי חישוב שונים.
בהמשך נוכיח שניתן לקבל את השפה $L = \left\{a^n b^n | n \geqslant 0\right\}$ בעזרת אוטומט מחסנית.

אוטומט מחסנית סגור לפעולות $\bigcirc, \cup, *$ אך לא לפעולות $\cap, ^c$. חלק מתכונות אלה נוכיח בהמשך.

1.1 הגדרה פורמלית

ככל שנתקדם בקורס לא יהיו לנו הגדרות פורמליות כמו של אס"ד, אך כן נגדיר באופן ברור יותר את באסל"ד, מעבר כלשהו אומר לנו שניתן לעבור למצב הבא אם נקרא תו מסוים.
בא"מ לעומת זאת, התווים שמופיעים בקלט אינם חייבים להיות התווים שמכניסים או מוציאים למחסנית. אזי יש לנו שתי שפות א"ב; Σ, א"ב הקלט ו- Γ, א"ב המחסנית.
מעברים כתובים כך במחסנית;

$$\sigma\gamma_1 \Rightarrow \gamma_2, \quad \sigma \in \Sigma_\epsilon, \quad \gamma_1, \gamma_2 \in \Gamma_\epsilon$$

כלומר, אם קוראים את התו σ, וגם ניתן להוציא את התו γ_1 אז ניתן לבצע הכנסה של γ_0. (אפשר, לא חובה)
ניתן להגדיר גם מעבר ללא התייחסות להוצאה אפשרית או ללא ביצוע של הכנסה, למשל;

$$a, \epsilon \Rightarrow 1, \quad b, 1 \Rightarrow \epsilon$$

בסך הכל;
א"מ M הוא שישייה

$$M = (Q, \Sigma, \Gamma, q_0, \delta, F)$$

כאשר;
Q- קבוצה סופית (של מצבים)
Σ- א"ב הקלט
Γ- א"ב המחסנית
$q_0 \in Q$- המצב ההתחלתי

$F \subseteq Q$ ־ קבוצת המצבים המקבלים.
ופונקציית המעבר־

$$\delta(q, \sigma, \gamma_1) \to (q_2, \gamma_2)$$

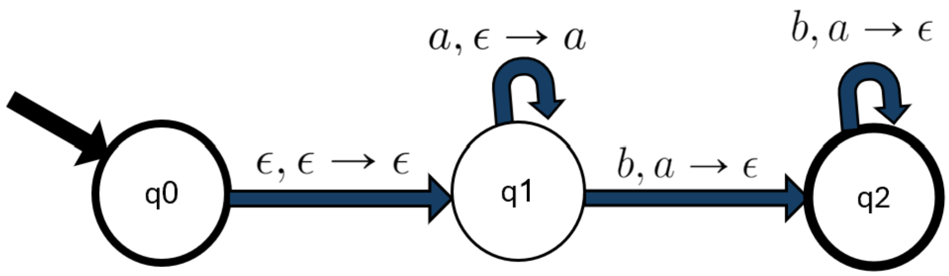
$$\delta : Q \times \Sigma_\epsilon \times \Gamma_\epsilon \to P(Q \times \Gamma_\epsilon)$$

1.2 דוגמאות

$$\Sigma = \Big\{a, b\Big\}, \quad \Gamma = \Big\{\Big\}$$

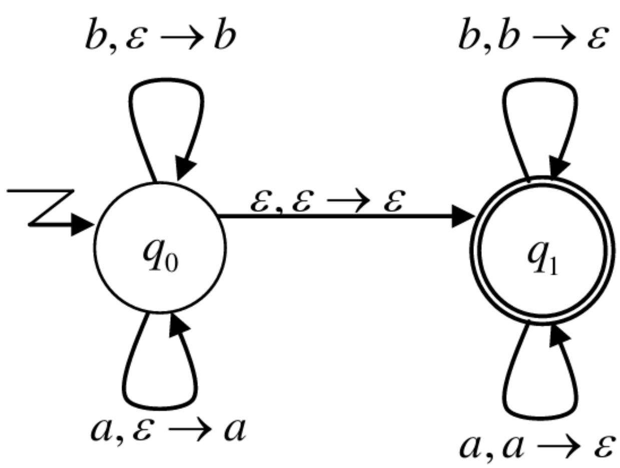
.1

$$L = \Big\{a^n b^n | n \geqslant 0\Big\}$$



.2

$$L_1 = \Big\{w = u_1 \bigcirc u_2 | \quad u_2 = u_1^R, \quad |u_1| = |u_2|\Big\}$$



כמו כן, גם את שלושת השפות הבאות ניתן לקבל מא"מ:

$$L = \Big\{w = u_1 \bigcirc u_2 | \quad u_1 \neq u_2, \quad |u_1| = |u_2|\Big\}$$

$$L = \Big\{w = u_1 \bigcirc u_2 | \quad u_1 = u_2^R, \quad |u_1| = |u_2|\Big\}$$

$$L = \left\{ w = u_1 \bigcirc u_2 \middle| \quad u_1 = u_2, \quad |u_1| = |u_2| \right\}$$

לא מדברים על סגירות, אלא על איך בונים א"מ ל־4 השפות $(L_1, ..., L_4)$ שצולמו. יש לכתוב חלק זה לשתי השפות שנותרו.

1.3 סגירות אוטומט מחסנית

1.3.1 שרשור

למדנו איך אפשר לקבל שרשור של שתי שפות דרך האס"דים שלהם. אותה השיטה לא תעבוד עם כל אוטומט מחסנית, מכיוון שייתכן שהוא לא יהיה ריק לאחר קריאת המילה הראשונה (מהשרשור).
הטריק שלנו לבדיקת ריקנות של א"מ הוא להוסיף אליו סימון מיוחד כלשהו, למשל $\$$, בהתחלה. כך אם אנו קוראים אותה, נדע שהגענו לתחתית המחסנית. יש לוודא שהמחסנית שלנו ריקה, אחרת ייתכן שלא נפעל כנדרש עם המילה השנייה.

1.4 שפות חסרות הקשר

שפות רגולריות הן שפות שיכולות להתקבל ע"י אס"ד, ושפות חסרות הקשר הן שפות שיכולות להתקבל ע"י א"מ. מתארים שפות אלה בעזרת דקדוק חסר הקשר. כלומר, שפה היא חסרת הקשר אם"ם קיים דקדוק חסר הקשר שמתאר אותה.

דקדוק חסר הקשר הוא רשימה של חוקים.
יש הבחנה בין שני תווים; טרמינלים ־ מסומנים באותיות קטנות ונונטרמינלים ־ שמסומנים באותיות גדולות.
כל חוק הוא מהצורה $A \to \alpha$ כאשר A הוא תו נונטרמינלי ו־ α היא מילה שיכולה להכיל מילים טרמינליות ונונטרמינליות.
השפה של הדקדוק חסר הקשר היא כל המילים שיכולות להיגזר ממנו ומורכבות רק מטרמינלים.

1.4.1 דוגמאות

נראה דוגמאות לדקדוקים חסרי הקשר שמייצגים שפות חסרות הקשר מעל הא"ב $\Sigma = \left\{a, b\right\}$.
1.

$$S \to aSa$$

$$S \to bSb$$

$$S \to \epsilon$$

השפה היא־

$$L = \left\{ w : w = uw^R \right\}$$

(או שפת כל הפולינדרומים באורך זוגי)

2.

$$S \to AB$$

$$A \to aAb$$

$$A \to \epsilon$$

$$B \to bB$$

$$B \to \epsilon$$

השפה היא־

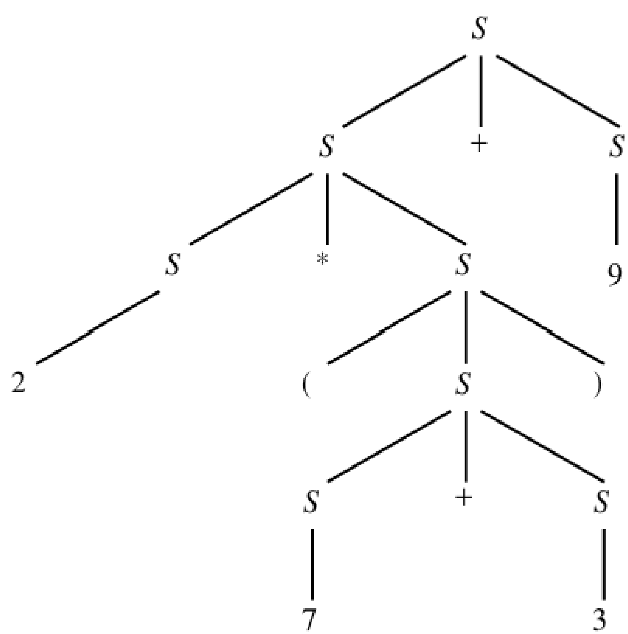
$$L = \left\{ a^n b^m : m \geqslant n \right\}$$

1.4.2 קצת היסטוריה ומוטיבציה

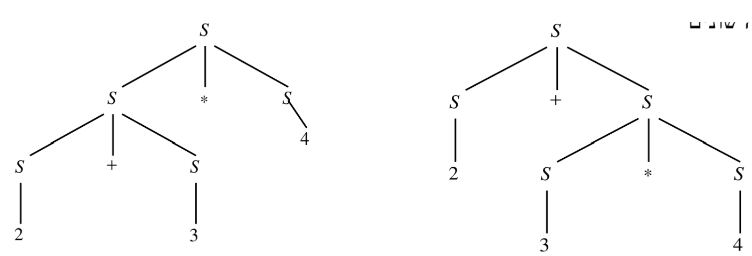
המונחים "שפה ודקדוק חסרי הקשר" באים דווקא מתחום הבלשנות, שזהו תחום חקר השפות הטבעיות. ואלו לא המונחים היחידים שזהו מקורם־
סינטקס: אילו משפטים הם משפטים תקניים בעברית?
משמעות: בהינתן משפט תקין, איך להבין אותו?

למשל; .אדם נשך כלב." הוא משפט תקני (אם כי לא כשר) אבל המשפט "אדם נשך נשך. כלב כלב" הוא לא תקני.
למען הנוחות, בלשנים חיפשו קבוצה סופית של חוקי משפטים (או במילים אחרות, דקדוק) שאם משפט יעבוד לפי, הוא יהיה תקני.
לבסוף התברר ששפות טבעיות אינן חסרות הקשר, כלומר לא ניתן לייצגן באמצעות דקדוק חסר הקשר. אך המושג עבר לשימוש של מדעי המחשב, בו כן יש שפות חסרות הקשר, כגון שפות תכנות ושפת הביטויים החשבוניים.

הנה עץ גזירות, שמבטא מילה של שפה חסרת הקשר, שמבטא ביטוי חשבוני; כאן הביטוי הוא $2*(1+3)+9$.



שימי לב שניתן לקבל עצי גזירה שונים לאותו הביטוי. למשל עבור $2+3*4$;



דרך נוספת להביע ביטוי היא מסלול־

$$S \to S + S \to S * S + S \to 2 * S + S, \to 2 * (S) + S \to 2 * (S) + 9 \to 2 * (7 + S) + 9 \to 2 * (7 + 3) + 9$$

אדם המממש קומפיילר יצטרך תחילה לכתוב דקדוק שגוזר את שפת התכנות. לאחר הגזירה הקומפיילר צריך גם להבין את התכנית, כלומר את המשמעות שלה.

<u>משפט</u> שפה היא חסרת הקשר אם"ם היא יכולה להתקבל על ידי א"מ.
בפרקים הבאים־ הוכחת המשפט :)