מודלים חישוביים
תרגול מס' 3

26 במרץ 2015

נושאי התרגול:

- למת הניפוח לשפות רגולריות.
- משפט מיהיל־נרוד.

1 למת הניפוח לשפות רגולריות

נזכר בלמת הניפוח, שתהווה עבורנו טכניקה להוכחת אי־רגולריות:

למה 1.1 לכל שפה רגולרית L קיים $\ell > 0$ כך שלכל $s \in L$ המקיימת $|s| \geq \ell$ קיים פירוק מהצורה $s = xyz$ כך ש:

1. לכל $i \geq 0$, $xy^iz \in L$.
2. $|y| > 0$.
3. $|xy| \leq \ell$.

איך נשתמש בלמת הניפוח כדי להוכיח ששפה L כלשהי היא אי־רגולריות? נניח בשלילה ש־ L רגולרית ויהא ℓ המובטח לנו. נבחר מילה $s \in L$ באורך גדול מ־ ℓ ונראה שלכל חלוקה xyz של s כך ש־ $|y| > 0$ ו־ $\ell \geq |xy|$ קיים $i \geq 0$ כך ש־ $xy^iz \notin L$, בסתירה ללמת הניפוח.

תרגיל 1

הוכח כי שפת הפלינדרומים מעל $\Sigma = \{0,1\}$, שנסמנה ב־ L_1, אינה רגולרית.

פתרון

נניח בשלילה ש־ L_1 רגולרית ויהא ℓ קבוע הניפוח המובטח לנו. נבחר

$$s = 0^\ell 1 0^\ell \in L_1$$

ונשים לב שאכן $|s| \geq \ell$. כעת נשים לב שכל חלוקה של s ל־ xyz כך ש־ $|xy| \leq \ell$ ו־ $|y| > 0$ מקיימת ש־ y מורכבת כולה מאפסים. לכן, $xy^2z \notin L_1$, בסתירה.

תרגיל 2

הוכח כי $L_2 = \left\{0^{n^2} \mid n \geq 0\right\}$ מעל $\Sigma = \{0\}$ אינה רגולרית.

פתרון

נניח בשלילה ש־ L_2 רגולרית ויהא ℓ קבוע הניפוח המובטח לנו. נבחר

$$s = 0^{\ell^2} \in L_2$$

ונשים לב שאכן $|s| \geq \ell$. נקח חלוקה של s ל־ xyz כך ש־ $|xy| \leq \ell$ ו־ $|y| > 0$. נסמן $|y| = k \leq \ell$ ונסתכל על

$$w = xy^2z = 0^{\ell^2+k}$$

נניח בשלילה ש־ $w \in L_2$. אזי, קיים n כך ש־ $\ell^2 + k = n^2$. אזי:

$$\ell^2 < \ell^2 + k \leq \ell^2 + \ell = \ell\left(\ell+1\right) < \left(\ell+1\right)^2$$

מכאן קיבלנו ש־ $\ell^2 < n^2 < \left(\ell+1\right)^2$, כלומר שקיים n המקיים $\ell < n < \ell+1$, בסתירה. אזי, L_2 אינה רגולרית.

2 משפט מיהיל־נרוד

נזכר כי עבור שפה L מעל הא"ב Σ, היחס $\sim_L$ מעל מילים ב־ $\Sigma^\star$ מוגדר כך: $x \sim_L y$ אם לכל $z \in \Sigma^\star$ מתקיים ש־ $xz \in L \Leftrightarrow yz \in L$. זהו יחס שקילות ולכן מחלק את $\Sigma^\star$ למחלקות שקילות. משפט מיהיל־נרוד יהווה עבורנו כלי נוסף להוכחת אי־רגולריות:

משפט 2.1 $L \subseteq \Sigma^\star$ רגולרית אס"ם קבוצת המנה $\Sigma^\star/_{\sim_L}$ סופית.

דוגמא 1

עבור השפה $L_3 = L\left(ba^\star b^\star\right)$, לכל זוג מילים מהזוגות הבאים, האם הם באותה מחלקת שקילות?

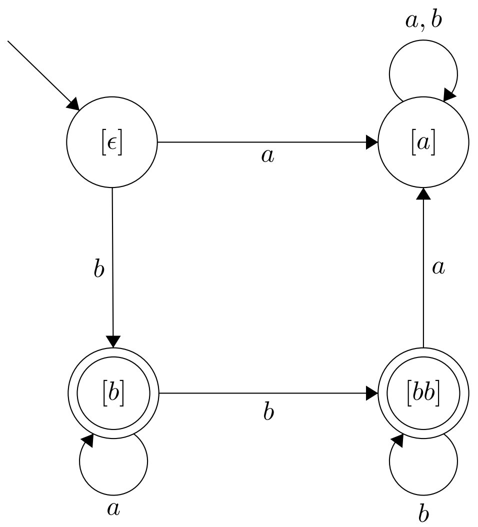
- aa ו־ ab? כן. שתי המילים אינן בשפה וגם אינן רישא של אף מילה בשפה.
- ba ו־ baa? כן. כל המשך חוקי של ba יהיה המשך חוקי של baa ולהיפך.
- baa ו־ bab? לא. עבור $z = a$, $baaa \in L_3$ אך $baba \notin L_3$.
- a ו־ ϵ? לא. עבור $z = ba$, $aba \notin L_3$ אך $ba \in L_3$.

עבור אותה שפה, מהן מחלקות השקילות?

1. $[\epsilon]_{\sim_{L_3}} = \{\epsilon\}$.
2. $[b]_{\sim_{L_3}} = L\left(ba^\star\right)$.
3. $[bb]_{\sim_{L_3}} = L\left(ba^\star bb^\star\right)$.
4. $[a]_{\sim_{L_3}} = L\left(\left(\epsilon + ba^\star bb^\star\right)a\left(b+a\right)^\star\right)$.

ואכן ציפינו למספר סופי של מחלקות שקילות. כיצד ניתן להשתמש בכך כדי לבנות אוטומט המקבל את L_3?

- המצבים יהיו מחלקות השקילות.
- המצב ההתחלתי יהיה מחלקת השקילות המכילה את ϵ.
- המצבים המקבלים יהיו מחלקות שקילות המוכלות ב־ L_3.



תרגיל 3

הוכיחו כי השפה $L_4 = \left\{ ww \mid w \in \{0,1\}^\star \right\}$ אינה רגולרית.

פתרון

לכל $i \geq 1$ נסמן $x_i = 0^i 1$. נשים לב שלכל $i \neq j$ מתקיים ש־ $\langle x_i, x_j \rangle \notin \sim_{L_4}$ (למשל, עבור $z = x_{\min\{i,j\}}$). אם כך, ליחס $\sim_{L_4}$ יש אינסוף מחלקות שקילות, ולפי משפט מיהיל־נרוד L_4 אינה רגולרית.

תרגיל 4

הוכיחו כי השפה $L_5 = \left\{ w \in \{0,1\}^\star \mid \#_0 (w) = \#_1 (w) \right\}$ אינה רגולרית.

פתרון

בהרצאה הוכחנו טענה זו לפי למת הניפוח. כעת נראה שתי דרכים שונות. תחילה, בעזרת משפט מיהיל־נרוד. לכל $i \geq 1$ נסמן $x_i = 0^i$. עבור $i \neq j$ מתקיים כי $x_i 1^i \in L_5$ אך $x_j 1^i \notin L_5$. אם כך, ליחס $\sim_{L_5}$ יש אינסוף מחלקות שקילות, ולפי משפט מיהיל־נרוד L_5 אינה רגולרית. שימו לב כי מחלקות השקילות הן מהצורה הבאה, עבור $k \in \mathbb{Z}$:

$$C_k = \left\{ w \in \{0,1\}^\star \mid \#_a (w) - \#_b (w) = k \right\}$$

הוכיחו זאת.

כעת, נוכיח כי L_5 אינה רגולרית בעזרת תכונות סגור. נניח בשלילה כי L_5 רגולרית ונשים לב כי

$$L_5 \cap L\left(0^\star 1^\star\right) = \left\{ 0^i 1^i \mid i \geq 0 \right\}$$

מכיוון שגם L_5 וגם $L\left(0^\star 1^\star\right)$ רגולריות (מכיוון שהיא מוגדרת ע"י ביטוי רגולרי), החיתוך שלהן הוא רגולרי. אך השפה $\left\{ 0^i 1^i \mid i \geq 0 \right\}$ אינה רגולרית, בסתירה. לכן, L_5 אינה רגולרית.

תרגיל 5

יהי $\Sigma = \{\sigma_1, \ldots, \sigma_n\}$ ונגדיר L_6 מעל Σ בצורה הבאה:

$$L_6 = \{ w \in \Sigma^\star \mid \forall \sigma \in \Sigma . \#_\sigma (w) \leq 1 \}$$

מצאו את מחלקות השקילות של $\sim_{L_6}$ וקבעו האם L_6 רגולרית. הוכיחו תשובתכם.

פתרון

לכל מילה $w \in \Sigma^\star$ נגדיר וקטור v_w באורך n כך ש־ $v_w(i)$ הוא מספר המופעים של σ_i ב־ w. נראה כי מחלקות השקילות של $\sim_{L_6}$ הן אוסף הוקטורים הבינאריים באורך n ומחלקה אחת נוספת, שאותה נסמן ב־ C_{no}. דהיינו, אם $w \in L_6$ אז v_w הוא וקטור בינארי וכל המילים בשפה בעלי אותו יצוג בינארי נמצאים באותה מחלקת שקילות. עבור כל $w \notin L_6$ מתקיים $C_{no} = [w]_{\sim_{L_6}}$. נסמן את מחלקות השקילות שאינן C_{no} ב־ $C_0, \ldots, C_{2^n-1}$ בהתאם לקידוד העשרוני של הוקטור הבינארי. נוכיח כי אלו אכן מחלקות השקילות של היחס. לשם כך, נוכיח:

1. לכל $w_1, w_2 \in C_i$ מתקיים $w_1 \sim_{L_6} w_2$.
2. לכל $w_1, w_2 \in C_{no}$ מתקיים $w_1 \sim_{L_6} w_2$.
3. לכל $w_1 \in C_i$ ו־ $w_2 \in C_{no}$ מתקיים $w_1 \nsim_{L_6} w_2$.
4. לכל $w_1 \in C_i$ ו־ $w_2 \in C_j$ עבור $i \neq j$ מתקיים $w_1 \nsim_{L_6} w_2$.
5. כל $w \in \Sigma^\star$ שייך למחלקה כלשהי ואף מחלקה אינה ריקה.

ברגע שנוכיח טענו אלו, נוכל לסכם כי L_6 רגולרית – שכן מספר מחלקות השקילות הוא $2^n + 1$. אם כך:

1. לכל $\sigma \in \Sigma$ מתקיים ש־ $\#_\sigma(w_1) = \#_\sigma(w_2)$ וכן $\#_a(w_1) \leq 1$. לכן, לכל $z \in \Sigma^\star$ אם קיים $\sigma \in \Sigma$ כך ש־ $\#_\sigma(z) > 1$ אזי $w_1 z \notin L_6$ וגם $w_2 z \notin L_6$. אחרת, אם קיים $\sigma \in \Sigma$ כך ש־ $\#_\sigma(z) = 1$ ו־ $\#_\sigma(w_1) = \#_\sigma(w_2) = 1$ אזי $w_1 z \notin L_6$ וגם $w_2 z \notin L_6$. אחרת, $\#_\sigma(w_1 z) = \#_\sigma(w_2 z) \leq 1$, ולסיכום: $w_1 z \in L_6$ אם"ם $w_1 z \in L_6$ ולכן $w_1 \sim_{L_6} w_2$.
2. לכל $w \in C_{no}$ קיים $\sigma \in \Sigma$ כך ש־ $\#_\sigma(w) > 1$ ולכן $w \notin L_6$. לכל, לכל $u \in \Sigma^\star$ גם $\#_\sigma(wu) > 1$ ולכן $wz \notin L_6$. כלומר, לכל $z \in \Sigma^\star$ מתקיים $w_1 z \notin L_6$ וגם $w_2 z \notin L_6$ ולכן $w_1 \sim_{L_6} w_2$.
3. לכל $\sigma \in \Sigma$ מתקיים $\#_\sigma(w_1) \leq 1$. עבור $z = \epsilon$ מתקיים $w_1 z \in L_6$ אך $w_2 z \notin L_6$, לפי טענה 2. אם כך, $w_1 \nsim_{L_6} w_2$.
4. קיים $\sigma \in \Sigma$ כך ש־ $\#_\sigma(w_1) \neq \#_\sigma(w_2)$. נניח בה"כ כי $\#_\sigma(w_1) = 0$. אם כך, $\#_\sigma(w_2) = 1$ ולכן עבור $z = \sigma$ מתקיים ש־ $w_1 z \in L_6$ אך $w_2 z \notin L_6$ ולכן $w_1 \nsim_{L_6} w_2$.
5. תחילה, ברור כי לכל $w \in \Sigma^\star$ קיים v_w שהוא או בינארי (דהיינו, מתאים לאחד מה־ C_i־ים) או לא (מתאים ל־ C_{no}). ברור גם כי $\sigma_1 \sigma_1 \in C_{no}$ ולכל $b \in \{0,1\}^n$ קיימת $w \in \Sigma^\star$ כך ש־ $v_w = b$ (למשל, ע"י מעבר בסדר לקסיקוגרפי על b ועבור כל 1 במקום ה־ i נשרשר את σ_i, החל מהמילה הריקה).