# מבוא לתורת הקבוצות ואוטומטים - 234129
# תרגיל בית 5 – פתרון

סמסטר חורף 2023-2024

המתרגל האחראי על התרגיל: רועי גרוס

תאריך הגשה: 08.04.24

נושאי התרגיל: אוטומט אי־דטרמיניסטי, אוטומט חזקה, למת הניפוח לשפות רגולריות

הוראות לפתרון התרגיל:

- בכל מקרה בו אתם מתבקשים להוכיח, עליכם לכתוב הוכחה פורמלית מלאה. עליכם לפרט את כל שלבי הפתרון.
- בכל מקרה בו אתם מנסים להפריך, נסו למצוא דוגמה נגדית קצרה ופשוטה ככל האפשר. זכרו (1) להגדירה במפורש, (2) להוכיח כי היא מקיימת את הנתונים, ו-(3) להוכיח כי אינה מקיימת את מסקנת הטענה.
- אם נדרש הסבר או נימוק בלבד, עליו להיות קצר אך בעל מבנה כללי של הוכחה (למשל הסבר/נימוק מדוע שתי קבוצות שוות צריך לכלול שני כיווני הכלה, וכו').

הוראות הגשה:

- ההגשה אלקטרונית דרך אתר הקורס.
- יש להגיש קובץ PDF בלבד.
- על הפתרון להיות **מוקלד**! אין להגיש פתרונות בכתב יד. איורים וציורים ניתן לצייר ביד.
- ההגשה היא בקבוצות של 3 סטודנטים.

תזכורת: הבדיקה היא מדגמית, ראו את סכמת הציונים באתר הקורס.

# בהצלחה!

## שאלה 1

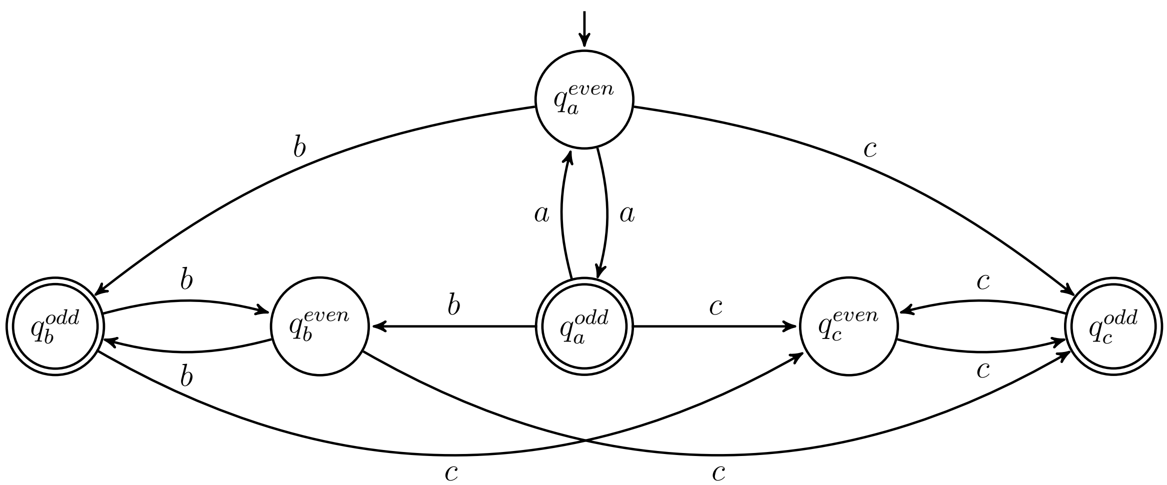
בנו אוטומט סופי אי־דטרמיניסטי לכל אחת מהשפות הבאות מעל $\Sigma = \{a, b, c\}$:

**(א)** שפת כל המילים מהצורה $\{a\}^*\{b\}^*\{c\}^*$ שאורכן אי־זוגי.

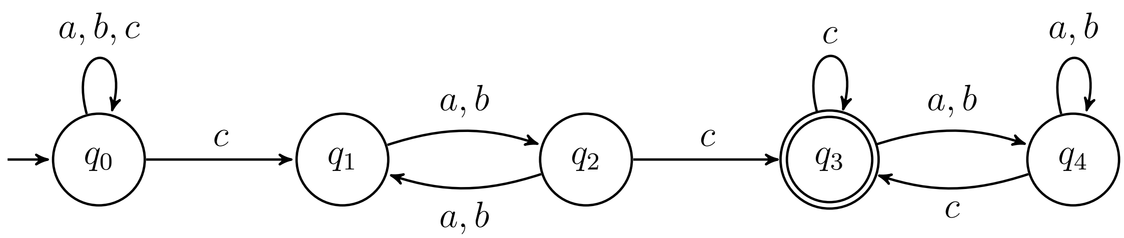
**(ב)** $\{w_1cw_2c\cdots w_nc \mid n \geq 2,\ \forall 1 \leq i \leq n :\ w_i \in \{a,b\}^*,\ \exists 2 \leq i \leq n :\ |w_i| \text{ אי־זוגי}\}$

## פתרון שאלה 1

**(א)** נגדיר אוטומט $A = (Q, \Sigma, I, R, F)$ כאשר $Q = \{q_a^{odd}, q_a^{even}, q_b^{odd}, q_b^{even}, q_c^{odd}, q_c^{even}\}$, $I = \{q_a^{even}\}$, $F = \{q_a^{odd}, q_b^{odd}, q_c^{odd}\}$ ו־$R$ מוגדר ע"י השרטוט:




**(ב)** נגדיר אוטומט $A = (Q, \Sigma, I, R, F)$ כאשר $Q = \{q_0, q_1, q_2, q_3, q_4\}$, $I = \{q_0\}$, $F = \{q3\}$ ו־$R$ מוגדר ע"י השרטוט:




הרעיון: נישאר במצב ההתחלתי עד שנראה את ה-$c$ הראשון מכאן ננחש מתי אנחנו קוראים את המילה מהאורך האי זוגי , עבורה נצטרך שני מצבים בשביל לבדוק שאנחנו אי זוגיים מכאן ואילך אם המילה מסתיימת ב-$c$ נעבור למצב מקבל.

## שאלה 2

תהי $L$ שפה רגולרית מעל א"ב $\Sigma$.

הוכיחו באמצעות בניית אוטומט שהשפה הבאה רגולרית:

$$L' = \{u\sigma v \mid u, v \in \Sigma^*,\ \sigma \in \Sigma,\ \sigma uv \in L\}$$

## פתרון שאלה 2

$L$ רגולרית, ולכן קיים אוטומט סופי דטרמינסטי $A = (Q, \Sigma, q_0, \delta, F)$ כך שמתקיים $L(A) = L$.

נבנה אוטומט אי־דטרמיניסטי $B$ עבור $L'$. לצורך כך ננחש איזו אות זזה, ונסמלץ את הריצה של $A$. מתישהו במהלך הריצה נצטרך לקרוא את האות שניחשנו ולהתעלם ממנה.

נגדיר $B = (Q', \Sigma, I, R, F')$ כאשר

$$Q' = Q \times (\Sigma \cup \{\$\})$$

כאשר אנו מניחים כי $\$ \notin \Sigma$.

$$I = \{(\delta(q_0, \sigma), \sigma) \mid \sigma \in \Sigma\}$$

כלומר, מתחילים מהמצב שהיינו מגיעים אליו לאחר קריאת $\sigma$, וזוכרים אותה להמשך.

$$F' = F \times \{\$\}$$

נגדיר את יחס המעברים:

$$\begin{aligned} R = \{((q, \tau), \sigma, (p, \tau)) \mid q, p \in Q,\ \tau \in \Sigma \cup \{\$\},\ \delta(q, \sigma) = p\} \\ \cup \{((q, \sigma), \sigma, (q, \$)) \mid q \in Q,\ \sigma \in \Sigma\} \end{aligned}$$

ננמק מדוע $L(B) = L'$.

$\Leftarrow$: תהי $w \in L'$. לפי ההגדרה מתקיים כי $w = u\sigma v$ כך ש-$\sigma uv \in L$. נתחיל במצב התחלתי $(\sigma, \delta(q_0, \sigma)) \in I$. במהלך הריצה על $u$ לא נעבור למצבי Q אלא נשאר בהעתק של האוטומט שהתחלנו בו. לאחר מכן נקלוט $\sigma$ ונעבור לעותק של $Q$ שמתויג עם $\$$, ומכאן נמשיך עד למצב מקבל. החישוב מתחיל במצב התחלתי, כל המעברים חוקיים, ממבנה המילה נגיע למצב מקבל. לפיכך, $w \in L(B)$.

$\Rightarrow$: תהי $w \in L(B)$, אז קיימת ריצה מקבלת ב־$B$. הריצה מתחילה במצב התחלתי מהצורה $(\sigma, p)$ עבור $\sigma \in \Sigma, p \in Q$. מתישהו במהלך הריצה אנחנו עוברים למצבי $Q$ שמתויגים עם $\$$. נסמן את המילה המתקבלת עד אז כ־$u$. האות שנקראה במעבר היא $\sigma$. נסמן את האותיות שנקראו אחר כך ב־$v$. נשים לב כי באוטומט המקורי הרצנו את $\sigma uv$ ומאחר שהגענו למצב מקבל אז נוכל להסיק כי מתקיים $\sigma uv \in L$ ולכן לפי ההגדרה $w = u\sigma v \in L'$.

## שאלה 3

יהי $A = (Q, \Sigma, q_0, \delta, F)$ אוטומט סופי דטרמיניסטי, ותהי $D \subseteq Q$ תת־קבוצה של מצבים. בהינתן מילה $w = \sigma_1\sigma_2...\sigma_n$, נתבונן בריצה $\pi = q_0 \xrightarrow{\sigma_1} q_1 \xrightarrow{\sigma_2} q_2... \xrightarrow{\sigma_n} q_n$ של $A$ על $w$. נאמר ש־$\pi$ היא **ריצה טובה** אם $D \subseteq \{q_0, q_1, ..., q_n\}$, כלומר אם הריצה מבקרת בכל המצבים ב־$D$.

<u>הוכיחו באמצעות בניית אוטומט</u> שהשפה הבאה רגולרית:

$$L = \{w \in L(A) \mid \text{הריצה של } A \text{ על } w \text{ היא ריצה טובה}\}$$

## פתרון שאלה 3

נשים לב שאנחנו רוצים לסמלץ את ריצת האוטומט A על מילה ולזכור את כל המצבים מ־$D$ שבהם ביקרנו וכך לדעת אם לקבל או לא.

נבנה אוטומט דטרמיניסטי $B = (Q', \Sigma, q_0', \delta', F')$ עבור $L$, כאשר

$$Q' = Q \times \mathcal{P}(Q)$$

$$q_0' = (q_0, \{q_0\})$$

$$F' = F \times \{S \mid D \subseteq S\}$$

נגדיר את פונקציית המעברים:

לכל $q \in Q$, $S \subseteq Q$, $\sigma \in \Sigma$:

$$\delta'((q, S), \sigma) = (\delta(q, \sigma), S \cup \{\delta(q, \sigma)\})$$

ננמק מדוע $L(B) = L$.

<u>$\Leftarrow$</u>: תהי $w \in L$. אזי $w \in L(A)$ וגם הריצה של $A$ על $w$ היא ריצה טובה. כלומר, כל המצבים ב־$D$ מופיעים בריצה של $A$ על $w$. מהגדרת $\delta'$, בכל פעם שמבקרים במצב ב־$A$, הוא "מתווסף" למצב ב־$B$, כלומר האוטומט מתקדם לפי $\delta$ של $A$ וזוכר את המצב ש־$A$ הגיע אליו. מאחר שיש לפחות ביקור אחד של $A$ בכל מצב ששייך ל־$D$, בסיום הריצה נגיע למצב שזוכר את כל המצבים ב־$D$ (ואולי עוד כמה שלא), כלומר $D$ מוכלת בקבוצת המצבים שביקרנו בהם. מאחר ש־$w \in L(A)$, הרכיב הראשון יהיה מצב מקבל ב־$F$, וסה"כ קיבלנו ש־$w \in L(B)$.

<u>$\Leftarrow$</u>: תהי $w \in L(B)$. אזי $\delta'(q_0', w) = (p, S)$ עבור $p \in F$ ו־$S \subseteq Q$ כך ש־$D \subseteq S$. מאחר שמצב ב־$B$ מכיל מצב $q \in Q$ ברכיב השני רק לאחר ביקור ב־$q$ לראשונה (מאופן הגדרת $\delta'$), אם הגענו לקבוצת מצבים שמכילה את $D$, בהכרח הקלט הוביל לביקור בכל המצבים ב־$D$. ברכיב הראשון הגענו ל־$p \in F$, ולכן קיימת ריצה מקבלת מתאימה ב־$A$, כלומר $w \in L(A)$. סה"כ קיבלנו $w \in L'$.

## שאלה 4

**(א)** <u>הוכיחו / הפריכו</u>: השפה $L_1 = \{a^i b^j \mid i,j \in \mathbb{N},\ i - j = 2024\}$ רגולרית.

**(ב)** <u>הוכיחו / הפריכו</u>: השפה $L_2 = \{a^i b^j \mid i,j \in \mathbb{N},\ i + j = 2024\}$ רגולרית.

## פתרון שאלה 4

**(א)** הטענה אינה נכונה.

**דרך 1 – למת הניפוח**: יהי $n \geq 1$ טבעי. נבחר $w = a^{2024+n} b^n$.

מתקיים כי $w \in L$ לפי הגדרה, וגם $|w| \geq n$.

נביט בפירוק כללי של המילה $w = xyz$ כך ש־$|xy| \leq n$ וגם $|y| \geq 1$.

מאחר ש־$n$ האותיות הראשונות הן $a$, נקבל כי

$$x = a^s \qquad y = a^t \qquad z = a^{2024+n-s-t} b^n$$

עבור $s,t \in \mathbb{N}$ כך ש־$t \geq 1$ ו־$s + t \leq n$.

נבחר $i = 2$. מתקיים:

$$xy^2z = a^{2024+n+t} b^n \notin L_1$$

שכן $2024 < 2024 + t = (2024 + n + t) - n$ כי $t \geq 1$.

משלילת למת הניפוח, $L_1$ לא רגולרית.

**דרך 2 – תכונות סגור**: נניח בשלילה כי השפה $L_1$ היא שפה רגולרית, ונשים לב שמהגדרת השפה נובע כי $L_1 = \{a^{2024} a^i b^i \mid i \in \mathbb{N}\}$. נשרשר אותה עם השפה $L' = \{b^{2024}\}$ ונקבל:

$$L_1 \cdot L' = \{a^{2024} a^i b^i b^{2024} \mid i \in \mathbb{N}\} = \{a^i b^i \mid i \in \mathbb{N},\ i \geq 2024\}$$

השפה $L'' = \{a^i b^i \mid i \in \mathbb{N},\ i < 2024\}$ היא שפה סופית ולכן רגולרית. מסגירות לאיחוד, השפה $(L_1 \cdot L') \cup L''$ רגולרית, אך מתקיים $(L_1 \cdot L') \cup L'' = \{a^i b^i \mid i \in \mathbb{N}\}$, וזאת שפה שאנחנו יודעים שאיננה רגולרית - סתירה!

**(ב)** הטענה נכונה. השפה היא סופית ולכן היא רגולרית, נראה מפורשות שהשפה סופית

$$L_2 = \left\{a^i b^j \mid i,j \in \mathbb{N},\ i + j = 2024\right\} \subseteq \{w \in \Sigma^* \mid |w| = 2024\}$$

(הצדיקו את המעבר האחרון). הקבוצה האחרונה סופית, ולכן גם $L_2$ סופית בתור תת־קבוצה של קבוצה סופית, ולכן לפי משפט $L_2$ רגולרית.

## שאלה 5

שפה $L$ מעל א"ב $\Sigma$ נקראת **שפת שרשרת** אם קיים $n \in \mathbb{N}$ כך שלכל $w \in L$ עבורה $|w| \geq n$, קיימת סדרת מילים אינסופית ... $w_0, w_1,$ כך ש־ $w_0 = w$, לכל $i \in \mathbb{N}$ מתקיים $w_i \in L$ וגם

$$0 < |w_{i+1}| - |w_i| \leq n$$

הוכיחו: אם $L$ שפה רגולרית, אז $L$ שפת שרשרת.

## פתרון שאלה 5

$L$ רגולרית ולכן מקיימת את למת הניפוח. יהי $n \geq 1$ הקבוע הטבעי המובטח מהלמה.

תהי $w \in L$ מילה כך ש־ $|w| \geq n$. לפי למת הניפוח, קיים עבורה פירוק $w = xyz$ כך ש־ $|xy| \leq n$, $|y| \geq 1$ ולכל $i \in \mathbb{N}$ מתקיים $xy^iz \in L$.

נגדיר סדרת מילים אינסופית $w_0, w_1, \ldots$ כך:

$$\forall i \in \mathbb{N} : \quad w_i = xy^{i+1}z$$

מתקיים $w_0 = xyz = w$, ומלמת הניפוח נובע כי לכל $i \in \mathbb{N}$ מתקיים $w_i = xy^{i+1}z \in L$.

לכל $i \in \mathbb{N}$ מתקיים גם

$$|w_{i+1}| - |w_i| = |xy^{i+2}z| - |xy^{i+1}z| = |y|$$

ומהלמה נובע ישירות $n \geq |y| > 0$. לכן, $n \geq |w_{i+1}| - |w_i| > 0$.

קיבלנו אפוא כי $L$ שפת שרשרת, לפי הגדרה.