מרצה: ד"ר מירה גונן

לתיקונים/הערות/הארות - נעם דומוביץ - $ndomovich@gmail.com$

אוטומטים ושפות פורמליות

הגדרות:

- <u>א"ב</u> - $\sum$ קב' סופית לא ריקה של אותיות
- <u>מילה</u> = סדרה סופית מעל א"ב $\sum$

דוגמה: הא"ב יהיה בינארי$\sum = \{0,1\}$ ומילה תהיה $w = 0100$ - הערה: תמיד נקראה משמאל לימין. (האות השניה היא 1)

- <u>שפה</u>: קבוצה של מילים, מעל א"ב $\sum$, נסמן ב L
- <u>מילה ריקה</u>: ε = מילה ללא אותיות
- <u>שפה ריקה</u> = $L = \emptyset$, נשים לב ש $L = \emptyset \neq \{\varepsilon\}$

דוגמה לשפה: $\sum = \{0,1\}$, $L = \{010, 11, 1011\}$, הערה $\varepsilon \notin L$ אא"כ הגדרנו

- $\sum = \{0,1\}$, L = שפת כל המילים באורך זוגי. (כאן $\varepsilon \in L$)
- <u>אורך של מילה</u> w , $|w|$ = מס' האותיות ב w
- <u>$\sum^*$</u> = אוסף כל המילים מעל א"ב $\sum$, $\varepsilon \in \sum^*$ כל L מעל $\sum$ מקיימת $L \subseteq \sum^*$
- $|\sum^*| = \aleph_0$ קב' של סדרות סופיות מעל קב' סימנים סופית (לא ריקה).
- מכיון ש $L \subseteq \sum^*$ נקבל $|L| \leq \aleph_0$

פעולות על מילים

- <u>שרשור</u> של מילים בהנתן $w_1, w_2 \in L$ המילה $w_1 \cdot w_2$ היא המילה המתקבלת מכתיבת אותיות w_2 אחרי w_1
  - דוגמה: $w_2 = 11$, $w_1 = 01$ השרשור: $w_1 \cdot w_2 = 0111$ וזה שונה מ $w_2 \cdot w_2 = 1101$.
  - שרשור מילים אינו פעולה חילופית, כלומר $w_1 \cdot w_2 \neq w_2 \cdot w_1$
  - שרשור היא פעולה קיבוצית, כלומר : $(w_1 \cdot w_2) \cdot w_3 = w_1 \cdot (w_2 \cdot w_3)$
  - שרשור של מילים לא בהכרח משמר סגירות בשפה, כלומר יתכן $w_1 \cdot w_2 \notin L$
- <u>חזקה</u> של מילים:
  - הגדרה פשוטה: $i \in \mathbb{N}$, w^i: $w^i = w \cdot w \cdot w \cdot w ...$
  - <u>הגדרה רקורסיבית</u>: $\begin{matrix} w^0 = \varepsilon \\ w^1 = w \\ w^i = w \cdot w^{i-1} \end{matrix}$
  - ע"פ הגדרה זו, נקבל: $w \cdot \varepsilon = \varepsilon \cdot w = w$ וזה מאוד הגיוני.
- <u>היפוך</u> $(reverse)$: w^R
  - הגדרה פשוטה: $\begin{matrix} w = a_1, a_1, ..., a_n \\ w^R = a_n, a_{n-1}, ...a_1 \end{matrix}$
  - <u>הגדרה רקורסיבית</u>: $\begin{matrix} w = \overbrace{a_1, a_2, ...}^{t} a_n \\ w^R = a_n t^R \end{matrix}$. במקרה ו $w = \varepsilon$ אז $w^R = w$

פעולות על שפות

תהיינה L_1, L_2 שפות מעל א"ב $\sum$ אז :

- איחוד - $L_1 \cup L_2 = \{w | w \in L_1 \vee w \in L_2\}$
- חיתוך - $L_1 \cap L_2 = \{w | w \in L_1 \wedge w \in L_2\}$
- משלים - $\bar{L_1} = \sum^* \backslash L_1$ וכן $L_1 \backslash L_2 = \{w | w \in L_1 \wedge w \notin L_2\}$
- שרשור שפות $L_1 \cdot L_2 = \{w | \exists u \in L_1, \exists v \in L_2, w = u \cdot v\}$

לדוגמה: $\sum = \{a, b\}$, $L_1 = \{a, ab, \varepsilon\}$, $L_2 = \{bb, a\}$, ואז $L_1 \cdot L_2 = \{abb, aa, abbb, aba, bb, a\}$

- תכונות של שרשור שפות:
  - $L \cdot \{\varepsilon\} = L$
  - $L \cdot \emptyset = \emptyset$
- פעולת שרשור על שפות **אינה חלופית,** יתכן - $L_1 \cdot L_2 \neq L_2 \cdot L_1$ בדוגמה הקודמת מתקיים ש $bba \notin L_2 \cdot L_1$, ומכאן שהן שונות.
- פעולת שרשור על שפות היא **כן קיבוצית** $(L_1 L_2) L_3 = L_1 (L_2 L_3)$

  הוכחה ע"י הכלה דו-כיוונית:

  כיוון ראשון: נראה $(L_1 L_2) L_3 \subseteq L_1 (L_2 L_3)$
  - יהיה $w \in (L_1 L_2) L_3$ ע"פ הגדרת השרשור קיימות $u \in L_1 L_2$ וקיימת $t \in L_3$ כך ש $w = u \cdot t$
  - לפי הגדרה קיימת $x \in L_1, y \in L_2$ כך ש $u = x \cdot y$
  - ובסה"כ קיבלנו $w = (xy)t = x(yt)$ ע"פ הגדרת קיבוציות למילים.
  - כעת $x \in L_1, y \cdot t \in L_2$ ולכן $w \in L_1 (L_2 L_3)$

  כיוון שני, אותו דבר.

שיעור 2 - 25/10/18

בשיעור שעבר ראינו: שרשור , חזקה, נחדד כמה נקודות באיטרציה:

- $L^* = \bigcup_{i \in \mathbb{N}} L^i = L^0 \cup L^1 \cup ...$
- $L = \sum$
- $L^* = \sum^*$
- $\sum^* = \bigcup_{i \in \mathbb{N}} \sum^i$
- תמיד: $L^* \subseteq \sum^* \Leftarrow |L^*| \leq \aleph_0$
- $L = \{\varepsilon\} \leftarrow L^* = \{\varepsilon\}$
- $L = \emptyset \Leftarrow L^* = \{\varepsilon\} \neq 0$ לכן L^* יכולה להיות סופית

איטרציה הערות:

שאלה: האם קיימות דוגמות נוספות עבורן L^* היא סופית?

תשובה: תהי L שפה כך שקיימת $w \in L$ $|w| \geq 1$, $L = \{a.......\}$, נשים לב שיכול להיות $w, w \cdot w, w^3, ...$, ולכן קיבלנו $\aleph_0$ מילים שונות ששיכות ל L^* לכן $|L^*| \geq \aleph_0$

- הראנו קודם $|L^*|\leq\aleph_0$ ובסה"כ $|L^*|=\aleph_0$

<u>דוגמה:</u> $\sum=\{0,1\}$, $L=\{00,1\}$, אז ... $L^*=\{\varepsilon\}\cup\{00,1\}\cup\{0000,001,100,11\}\cup\{0000000,00001,...\}$,

<u>שאלה:</u> האם יתכן ש $L^*=L$, <u>תשובה:</u> ראינו קודם ש $L=\{\varepsilon\}\leftarrow L^*=\{\varepsilon\}$ ומכאן שכן.

<u>שאלה:</u> האם יתכן בעבור $L^*\neq\{\varepsilon\}$? <u>תשובה:</u> נשים לב שאם ניקח את $L=\sum^*$ אז יתקיים $L^*=(\sum^*)^*=\sum^*$
<u>הוכחה למעבר</u> - נתבונן בהגדרה: $(\sum^*)^*=(\sum^*)^0\cup\boldsymbol{(\sum^*)^1}\cup.......\cup\boldsymbol{(\sum^*)^*}$ כלומר : $\boldsymbol{(\sum^*)^1}\subseteq(\sum^*)^*\subseteq\boldsymbol{(\sum^*)^*}$

אז הראנו שני מקרי קצה, האם ישנן דוגמאות נוספות:
יהיה $\sum=\{0,1\}$, תהיה שפת כל המילים שמס' האפסים שלו זוגי, $L^*=\{\varepsilon\}\cup\{0,1\}=L\cup L^2\cup.....L^i$ מהגדרת השרשור כל L^i כלשהי יהיה גם זוגי ומכאן שמוכל ב L (ע"פ הגדרת השפה) , ולכן $L^*=L$ אז מצאנו דוגמה יותר מעניית לשאלתנו, אך נשים לב שבאמת ברוב המקרים ב L^* יהיה יותר מילים מ L

<u>טענה:</u> $\underline{(L^*)^*=L^*}$- מסקנה מידית : אין טעם לבצע איטרציה יותר מפעם אחת.

הוכחה:
<u>נראה $L^*\subseteq(L^*)^*$:</u>

- על פי הגדרה .... $(L^*)^*=\bigcup_{i\in\mathbb{N}}=(L^*)^0\cup\boldsymbol{(L^*)^1}\cup....$
- ולפי הגדרת האיחוד נקבל: $L^*\subseteq(L^*)^*$ (מתורת הקבוצת $A\subseteq A\cup B$)

<u>כיוון שני - $(L^*)^*\subseteq L^*$:</u>

- תהי $w\in(L^*)^*$ לפי הגדרת האיטרציה קיים i כך ש $w\in(L^*)^i$,
- אז לפי הגדרת החזקה קיימות $u_1,u_2,....,u_i\in L^*$ כך ש: $w=u_1u_2....u_i$
  כעת נרצה להציג את המלים ע"פ L.
- כל אחת מהמלים u_j כאשר $1\leq j\leq i$ "נפרק" לשרשור של מילים ב L,
- נפרק את u_j ל $t_{j_1},t_{j_2},...,t_{j_{k_j}}$ זה נכון כי $u_j\in L^*$ ולכן לכל u_j קיים טבעי k_j כך ש $u_j\in L^{k_j}$ (k_j באה להצביע על אורך שונה במקום לבחור אות נוספת).

  לכן:

$$w=t_{11},t_{12},....t_{1k_1},t_{21},t_{22},....t_{2k_2},......t_{i1},t_{i2},....t_{ik_i}\in L^{k_1+k_2+....+K_i}\subseteq L^*$$

$\square$

דוגמה - L- שפת כל המילם באורך זוגי מעל $\sum=\{0,1\}$ אז $w\in(L^*)^*$ תהיה

$$\begin{array}{cccccc} & t_{11},t_{12} & t_{21} & t_{31} & t_{32},t_{33} & t_{41}\\ w= & 0000 & \varepsilon & 11 & 001111 & 0\\ & u_1 & & u_2 & u_3 & u_4\end{array}$$

אז $w_i\in L^*$ ולכן $w\in L^{2+1+3+1}\subseteq L^*$

היפוך: - L^R , הגדרה: $L^R=\{w|\,w^R\in L\}$.

דוגמאות לשפות:

- שפת כל המילים המתחילות ב ab מעל $\sum=\{a,b\}$: סימון ע"י שרשור $L=\{abu \mid u\in\sum^*\}$, ניתן גם כשרשור שפות: $L=\{ab\}\cdot\sum^*$ או בצורה יותר מסובכת : $L=\{a\}\{b\}\sum^*$
- שפת כל המילים שמכילות a,b מעל $\sum=\{a,b\}$: שוב סימון ע"י שפות: $L=\sum^*\{ab\}\sum^*$
- שפת כל המלים שמכילות a,b ולא bb מעל $\sum=\{a,b\}$:
  - ניסיון 1: $L=\{ab,a\}^*$ לא טוב - $a,\varepsilon\in L$.
  - ניסיון 2: $L=\{ab\}^*$ לא טוב $aab\notin\{ab\}^*$ ושייכת לשפה המקורית.
  - ניסיון 3: $L=\{ab,a\}^*\{ab\}\{ab,a\}^*$ לא טוב כי bab לא מתקבלת.
  - ניסיון 4 (מירה): $L=\left(\sum^*\{ab\}\sum^*\right)\setminus\left(\sum^*\{bb\}\sum^*\right)$ - מצוין!

שיעור 3 - 01/11/18

אוטומט סופי

הקדמה:

אוטומט סופי הוא מודל חישובי שבו אוסף סופי של מצבים עם כללי מעבר ממצב אחד לשני.

הגדרה - לא פורמלית: אוטומט סופי הוא מודל מתמטי מופשט המתאר מערכת שמגיבה לקלטים.

- באוטומט סופי שני סוגי מצבים – מצבים מקבלים ומצבים לא מקבלים .
- תוצאת החישוב מוגדרת ע"י סוג המצב אליו מגיע האוטומט בסיום קריאת הקלט :
  - נאמר שהאוטומט **מקבל** קלט w אם הוא מסיים את קריאת w במצב מקבל .
  - אחרת נאמר שהוא **דוחה** את w.
- תגובתו של אוטומט לאות קלט הינה פונקציה של המצב הנוכחי ושל האות
- פונקציה זו קובעת מצב יחיד אליו האוטומט הדטרמיניסטי יעבור .

לאוטומט סופי דטרמיניסטי יש חמישה מרכיבים והם :

- <u>א"ב</u> – כל אותיות הקלט האפשריות עבור האוטומט. מספר האותיות בא"ב זה חייב להיות סופי וגדול מ-0.
- <u>מצבים</u> – כל המצבים שבהם יכול האוטומט להימצא. מספר המצבים חייב להיות סופי וגדול מ-0.
- <u>מצב התחלתי</u> – המצב שממנו מתחיל האוטומט את מסלול החישוב על כל מילת קלט . קבוצת מצבים מקבלים – קבוצה מתוך קבוצת המצבים, המכילה 0 מצבים או יותר .
- <u>פונקציית מעברים</u> – לכל זוג של מצב ואות, פונקציה זו מתאימה מצב (אחד ויחיד) שאליו עובר האוטומט כאשר במצב זה נקראת אות זו.

אוטומט סופי דטרמינסטי A נתון על ידי $A=\left(\sum_A,Q_A,q_{0A},F,\delta_A\right)$, כאשר:

- $\sum_A$ - א"ב קלט
- Q_A - קבוצה סופית לא ריקה של מצבים
- q_{0A} - מצב התחלתי $q_{0A}\in Q_A$

- F_A - קבוצת מצבים מקבלים $F_A \subseteq Q_A$
- δ_A - פונקצית מעברים, $\delta_A : Q_A \times \sum_A \to Q_A$

פונקציית המעבר - פירוט:

- לוקחת שני ארגומנטים: מצב ואות
- $\delta(q,a)$ = המצב שהאוטומט עובר אליו כשהוא במצב q וקלט a .

שיעור 4 - 8/11/18

נרחיב את פונקציית המעברים למחרוזות

נרחיב את הגדרת δ למצב ומחזורות:

נגדיר באינדוקציה על אורך המחרוזת:

- בסיס $\delta(q,\varepsilon) = q$
- צעד: $\delta(q,wa) = \delta(\delta(q,w),a)$ כאשר w היא מחרוזת ו - a היא אות.

שפה של אוטומט דטרמיניסטי:

הגדרות:

- אם $A = \left(\sum_A, Q_A, q_{0A}, F_A, \delta_A\right)$ הוא אוטומט אז $L(A)$ היא השפה שמקבל האוטומט. = אוסף כל המחרוזות w כך ש $\delta(q_0,w)$ שייך ל F . כלומר: $L(A) = \{w \in \sum^* \mid (\delta_0,w) \in F\}$

  כלומר אוסף כל המחרוזות שמקבל האוטומט - אוסף כל המחרוזות המהוות מסלול ממצב ההתחלה למצב מקבל.

- לכל $q \in Q$ שפת המצב $L_A(q)$ היא השפה הבאה: $L_A(q) = \{w \in \sum^* | \delta(q_0 = w)\}$

  מתקיים $L(A) = \bigcup\limits_{q \in F} L_A(q)$

טענה: מתקיים $\delta(q, w_1 \cdot w_2) = \delta(\delta(q,w_1), w_2)$

<u>הוכחה:</u> באינדוקציה על $|w_2|$

<u>בסיס:</u>

- אם $w_2 = a$ אז זו בדיוק פונקציית המעברים מורחבת למילים (ע"פ הגדרה).
- אם $w_2 = \varepsilon$ אז $\delta(q_0, w_1 \cdot \varepsilon) \overset{1}{=} \delta(q_0,w_1) \overset{2}{=} \delta(\delta(q_0,w_1),\varepsilon)$

  1. $w\varepsilon = w$ .2 הגדרת בסיס פונקציית מעבר

<u>צעד:</u>

נניח שהטענה נכונה ל $\delta(q,w_1w_2) = \delta(\delta(q,w_1),w_2)$ נכונה לכל $|w_2| = n$ ונוכיח בעבור w_2a , $a \in \sum$:

$$\delta(q,w_1w_2a) \overset{1}{=} \delta\left(\underbrace{\delta(q,w_1w_2)},a\right) \overset{2}{=} \delta\left(\delta\left(\underbrace{\delta(q,w_1)},w_2\right),a\right) \overset{3}{=} \delta(\delta(q',w_2),a)$$

$$\overset{1}{=} \delta(q',w_2a) \overset{3}{=} \delta(\delta(q,w_1),w_2a)$$

1. הגדרת פונקציית המעברים. 2. הנחת האינד' 3. פונקציית מעברים רגילה: $\delta(q,w_1) = q'$.

דוגמה: אוסף כל המילים מעל $\{0,1\}$שאין בהן 11.

צ"ל להראות ש: $L(A)=L$ (הכלה דו־כיוונית)

נראה $L(A)\subseteq L$ - נוכיח באינדוקציה על האורך של w :

טענה נסבך מעט - בעצם נרצה להוכיח:

1. אם $\delta\left(q,w\right)=q_0$ אז w לא מכילה 11....
2. אם $\delta\left(q,w\right)=q_1$ אז w לא מכילה 11....

בסיס:

תהיה $w=\varepsilon$ אז $|w|=0$, $w\in L\left(q_0\right)$ אין ב w 11 וכמובן $w=\varepsilon$.

צעד:

נניח שהטענה נכונה בעבר $|w|\le n$ ונוכיח בעבור wa . $a\in\{0,1\}$ ולכן נפצל למקרים:

- אם $\delta\left(q,wa\right)=q_0$ אז $\delta\left(\delta\left(q_0,w\right)a\right)=q_0$ ומכיון שבאוטומט כל הקשתות הנכנסות לq_0 מתוייגות באפס, בהכרח $a=0$, ולכן בודאות wa מסתיימת ב 0 , לפי הנחת האינדוקציה w לא מכילה 11 ולכן גם wa לא מכילה 11 . (נשים לב שאכן $\delta\left(q,w\right)\neq q_2$, כי $\delta\left(q_0,wa\right)=q_0$, ואין קשתות מ q_2 שנכנסות ל q_0).
- אם $\delta\left(q,wa\right)=q_1$ אז $\delta\left(\delta\left(q_0,w\right)a\right)=q_1$, מכיון שבאוטומט הקשת היחידה הנכנסת ל q_1 היא מ q_0 , מתקיים בהכרח ש $\delta\left(q_0,w\right)=q_0$ ו $a=1$.
- לפי הנחת האינד' על w , w לא מכילה 11 , ומסתיימת ב 0 (או ε) לכן wa לא מכילה 11

כיוון שני: $L\subseteq L(A)$ - נניח בשלילה שאם w לא מתקבלת ע"י האוטומט אז w מכילה 11 :

- הדרך היחידה ש w לא תתקבל היא אם האוטומט יגיע ל q_2
- הדרך היחידה להגיע ל q_2 , $\delta\left(q_0,w\right)=q_2$ היא אם $w=x1y$ אם x מגיע ל q_1 ו y הסיפא של w שמגיעה ל q_2 בפעם הראשונה.
- אם $\delta\left(q_0,x\right)=q_1$ אז בהכרח $x=z1$ עבור z כלשהו
- לכן $w=z11y$, כלומר w מכיל 11, וזו סתירה.

שיעור 5 - 11/18/15

חזרה על הוכחת $\delta\left(q,w_1w_2\right)=\delta\left(\delta\left(q,w_1\right),w_2\right)$

שפות רגולריות

שאלות שמנחות אותנו, האם ישנן שפות שאס"ד לא מזהה? מה התכונה של אלו שהוא כן מקבל וכד'

הגדרה: שפה L היא רגולרית אם היא מתקבלת ע"י אוטומט סופי סטרמיניסטי

דוגמאות:

- השפה הריקה $\emptyset$ - רגולרית
- השפה $\{\varepsilon\}$ - רגולרית
- לכל a השייכת ל $\sum$ השפה $\{a\}$ רגולרית
- לכל מילה w ב $\sum^*$ השפה $\{w\}$ רגולרית

הוכחה - בניית אוטומט

קיום שפה לא רגולרית משיקולי עוצמות

$|\text{automats sets on} \sum| = \aleph_0 \leftarrow$ זו סדרה סופית של סמנים מעל קב' סופית

$L \subseteq \sum^*$ והרי $|\sum^*| = \aleph_0$. ומכאן ש: | קבוצת השפות מעל $\sum| = |P(\sum^*)| = 2^{\aleph_0}$

ולכן לא רק שיש שפה לא רגולרית, אלא מהוכחה נראה שרב השפות אינן רגולריות.

$L = \{0^n 1^n | \geq 1\}$ **אינה שפה רגולרית**

<u>הוכחה:</u>

- נניח בשלילה כי קיים אוטומט $A = (Q, \sum, q_0, \delta, F)$ כך ש $L(A) = L$
- נתבונן בסדרת המצבים .... $q_0, q_1,$ כך ש $q_i = \delta(q_0, 0^i)$
- כיון שמספר המצבים סופי ואורך הסדרה לא, קיימים $i < j$ כך ש $q_i = q_j$ כלומר $\delta(q_0, 0^i) = \delta(q_0, 0^j)$
- עתה, עפ"י תכונות השרשור נקבל כי

$$\delta\left(q_0, 0^i 1^i\right) = \delta\left(\delta\left(q_0, 0^i\right), 1^i\right) = \delta\left(\delta\left(q_0, 0^j\right), 1^i\right) = \delta\left(q_0, 0^j 1^i\right)$$

- אבל $\delta(q_0, 0^i 1^i) \in F$ - מההנחה שיש אוטומט כזה
- ואילו $\delta(q_0, 0^j 1^i) \notin F$ בסתירה לשיוויון הנ"ל.

סגירות השפות הרגולריות תחת פעולות בוליאניות

תהיינה L_1, L_2 שפות רגולריות.

הכלה - כלומר האם בהכרח $L_1 \subseteq L_2$ רגולרית?

- לא. ד"נ:
  - נגדיר :

$$L_1 = \{a^n b^n | n > 0\} \quad L_2 = \{\sum^* \text{where } \sum = \{a, b\}\ \}$$

  - והרי $L_1 \subseteq L_2$ אבל הוכחנו ש L_1 אינה רגולרית
- דוגמה נוספת L_1 רגולרית האם המכילה רגולרית , לא - ד"נ:

$$L_1 = \{01\} \quad L_2 = \{a^n b^n | n > 0\}$$

משלים - כלומר אם L רגולריות האם $\bar{L}$ רוגולרית ?

<u>טענה נכונה - הוכחה</u>: הרעיון נהפוך את קבוצת המקבלים ללא מקבלים, ואת הלא מקבלים למקבלים.

- בהינתן אוטומט $A = (Q, \sum, q_0, \delta, F)$ נבנה אוטומט $\bar{A}$ עבור $\bar{L}$

$$\bar{A} = (Q, \sum, q_0, \delta, Q \backslash F)$$

- נראה כי $\bar{L} = L(\bar{A})$:

$$w \in \bar{L} \iff w \notin L = L(A) \iff \delta(q_0, w) \notin F \iff \delta(q_0, w) \in Q \backslash F \iff w \in L(\bar{A})$$

חיתוך - הטענה נכונה

<u>הרעיון:</u> נבנה את אוטומט המכפלה

- תהיינה L_1, L_2 שפות רגולריות מעל $\sum$ ויהיו $\begin{Bmatrix} A_1 = (Q_1, \sum, q_{01}, \delta_1, F_1) \\ A_2 = (Q_2, \sum, q_{02}, \delta_2, F_2) \end{Bmatrix}$, האוטומטים המקבלים את השפות L_1, L_2 בהתאמה.
- כעת צ"ל שקיים אוטומט A כך ש $L(A) = L_1 \cap L_2$
  - $A = (Q, \sum, q_0, \delta, F)$
  - $Q = Q_1 \times Q_2$
  - $q_0 = (q_{01}, q_{02})$
  - $\forall q_1 \in Q_1, q_2 \in Q_2, \delta\left((q_1, q_2), a\right) = \delta_1\left((q_1, a\right), \delta_2\left((q_1, a\right)$
  - $F = F_1 \times F_2$
- כלומר, צ"ל $L(A) = L_1 \cap L_2$
- נשתמש בלמה הבאה, לכל $w \in \sum^*$, $(q_1, q_2) \in Q$ מתקיים :

$$\delta\left(\left(q_1, q_2\right), w\right) = \delta\left(\delta_1, \left(q_1, w\right), \delta_2\left(q_2, w\right)\right)$$

<u>הוכחה - באינדוקציה על $|w|$:</u>

- <u>בסיס:</u> נניח כי $w = \varepsilon$, $|w| = 0$

$$\delta\left(\left(q_1, q_2\right), w\right) = \left(q_1, q_2\right) = \delta\left(\delta_1, \left(q_1, w\right), \delta_2\left(q_2, w\right)\right)$$

- <u>צעד:</u> נניח לכל $|u| < n$ ונוכיח ל $w = ua$, כך ש $|w| = n$, $|u| = n-1$
- מנתון $\delta\left(\left(q_1, q_2\right), w\right) = \delta\left(\left(q_1, q_2\right), ua\right)$
- מהגדרת פונקצייה המעברים: $\delta\left(\left(q_1, q_2\right), ua\right) = \delta\left(\left((q_1, q_2), u\right), a\right)$
- מהנחת האינדקוציה: $\delta\left(\left((q_1, q_2), u\right), a\right) = \delta\left(\delta_1, \left(q_1, u\right), \delta_2\left(q_2, u\right), a\right)$
- מהגדרת δ : $\delta\left(\delta_1, \left(q_1, u\right), \delta_2\left(q_2, u\right), a\right) = \left(\delta_1\left(\delta_1, \left(q_1, u\right), a\right), \delta_2\left(\delta_2\left(q_2, u\right), a\right)\right)$
- ומהגדרת פונ' המעברים למילים של שני אוטומטים:

$$\left(\delta_1\left(\delta_1, \left(q_1, u\right), a\right), \delta_2\left(\delta_2\left(q_2, u\right), a\right)\right) = \left(\delta_1\left(q_1, w\right), \delta_2\left(\delta_2, w\right)\right)$$

- כעת נראה $L(a) = L_1 \cap L_2$:

$$\begin{aligned} w \in L(A) = \iff \ & \delta\left(q_0, w\right) \in F = F_1 \times F_2 \\ \iff \ \delta\left(q_0, w\right) = \delta\left(\left(q_{01}, q_{02}\right), w\right) & \stackrel{1}{=} \left(\delta_1\left(q_{01}, w\right), \delta_1\left(q_{02}, w\right)\right) \in F_1 \times F_2 \\ \stackrel{2}{\iff} \ & \left(\delta_1\left(q_{01}, w\right) \in F_1,\ \delta_1\left(q_{02}, w\right)\right) \in F_2 \\ \iff \ & w \in L_1, w \in L_2 \iff w \in L_1 \cap L_2 \end{aligned}$$

1. מהלמה. 2. בהוכחת האיחוד נשנה את ה',$\wedge$=' ל $\vee$

איחוד - דה-מורגן...

שיעור 6 - 22/11/18

ראינו בשיעור שעבר: סגירות שפות רגולירות - למשלים, חיתוך - ע"י בניית אוטומט מכפלה.

איחוד L_1, L_2 רגולריות $L_1 \cup L_2$ רגולרי?

הוכחה 1 :

- ניתן להשתמש באוטומט המכפלה גם לאיחוד - ע"י שינוי קבוצת המצבים המקבלים - $F = \{(q_i, q_j) \,|q_i \in F_1 \text{or } q_i \in F \}$, לכן האיחוד שפה רגולרית

הוכחה 2:

- דרך נוספת ע"י דה מורגן:

$$L_1 \cup L_2 = \overline{\overline{L_1 \cup L_2}} = \overline{\overline{L_1} \cup \overline{L_2}}$$

- L_1, L_2 רגולרית ולכן מסגירות למשלים $\bar{L_1}, \bar{L_2}$ רגולריות
- מסגירות לחיתוך $\bar{L_1} \cap \bar{L_2}$ ושוב מסגירות למשלים נקבל ש $\overline{\overline{L_1} \cup \overline{L_2}}$ רגולרית

הוכחה 3 :

הוכחת החיתוך המלאה עם שינויי סימן.

מה קורה עם איחוד/חיתוך של מס' כלשהו של שפות רגולריות?

- איחוד/חיתוך של מספר סופי של שפות רגולריות הוא שפה רגולרית - הוכחה באינ' על מס' השפות
- האם איחוד/חיתוך של $\aleph_0$ שפות רגולריות הוא שפה רגולרית? לא
  - איחוד: נחפש ד"נ כך ש $\bigcup L_i - \text{not regular}$, אבל $L_i - \text{regular}$
  - לכן נבחר $\bigcup\limits_{i \in \mathbb{N} \cup \{0\}} \{a^n b^n | n > i\}$, כאשר - $L_i = \left\{a^i, b^i\right\}$
  - חיתוך: אם הטענה היתה נכונה אז מסגירות למשלים, וחיתוך הינו מקבלים שגם האיחוד רגולרי - סתירה.

אוטומט סופי לא דטרמינסטי

- אוטומט סופי לא דטרמניסטי יכול להמשיך לכמה מצבים, או להתקע
- מעבר ממצב ואות נתונה יכול להיות לקבוצת מצבים.
- מתחילים ממצב התחלתי נתון.
- לכל מילה יתכנו מספר חישובים אפשריים, או אף חישוב אפשרי
- האוטומט מקבל את קיימת סדרת מעברים המובילה למצב מקבל

שיעור 7 - 29/11/18

חזרה על אסל"ד - שחור אדום משעור שעבר.

אסל"ד - הגדרה פורמלית:

$A=(Q,\sum,q_0,\delta,F)$ כאשר:

- Q קבוצת מצבים
- $\sum$ א"ב נתון
- δ פונקציית מעברים
- $q_0 \in Q$ מצב התחלתי
- פונקציית המעברים δ:
  - $\delta(q,a)$ היא **קבוצת** מצבים.
  - הרחבה למחרוזת:
  - בסיס: $\delta(q,\varepsilon)=\{q\}$
  - צעד: $\delta(q,wa)=\bigcup_{p\in\delta(q,w)}\delta(p,a)$
  - סימון - להרחבה לקבוצת מצבים: $\delta(\boldsymbol{P},w)=\bigcup_{\boldsymbol{q}\in\boldsymbol{P}}\delta(q,w)$
- מחזורות מתקבלת אם $\delta(q_0,w)$ מכיל לפחות מצב מקבל אחד, כלומר , קיים מסלול חישוב על w המתחיל ב q_0 ומסתיים במצב מקבל.
- השפה של האוטומט היא קבוצת המחרוזות שהוא מקבל:

  כלומר $L(A)=\{w\in\sum^*|\delta(q_0,w)\cap F\neq\emptyset\}$

משפט : אס"ד שקול לאסל"ד

<u>כיון ראשון:</u> ניתן להפוך כל אס"ד לאוטומט לאסל"ד:

אם $\delta_D(q,a)$ נגדיר $\delta_N(q,a)=\{p\}$ (נגדיר כל מעבר, כקבוצה וסיימנו)

<u>כיוון שני:</u> ניתן להפוך כל אסל"ד לאס"ד.

- יהי $A=(Q,\sum,q_0,\delta_N,F_N)$
- הרעיון נבנה את אוטומט החזקה - האס"ד יראה כך:
  - $P(Q)$ תהיה קבוצת המצבים
  - $\sum$ יהיה א"ב
  - $\{q_0\}$מצב התחלתי
  - קבוצת מצבים מקבלים F_D : כל תתי הקבוצות של Q המכילות איבר מ F
  - פונקציית המעבר $\delta_D(\{q_1,....,q_k\},a)$ היא הקבוצה המתקבלת מאיחוד על כל $\delta_N(q_i,a)$ $1\le i\le k$

$$\delta_D\left(\underbrace{\{q_1,...,q_k\},a}_{\in P(Q)}\right)=\underbrace{\bigcup_{i=1}^{k}\delta_N(q_ia)}_{\in Q}$$

הוכחת השקילות :

נראה באינדוקציה על $|w|$ נניח כי $\delta_N(q_0,w)=\delta_D(\{q_0\},w)$

בסיס:

אם $w=\varepsilon$, $\delta_D(\{q_0,\varepsilon\})=\{q_0\}=\delta_N(q_0,\varepsilon)$ מהגדרה בעבור מילה ריקה = להשאר באותו מצב.

צעד:

1. תהי $w=xa$:

$$\delta_D(\{q_0,w\})\overset{1}{=}\delta_D(\{q_0,xa\})\overset{2}{=}\delta_D(\delta_D\{q_0,x\},a)\overset{3}{=}\delta_D(\delta_N(q_0,x),a)$$

$$\overset{4}{=}\delta_N(\delta_N(q_0,x),a)\overset{5}{=}\delta_N(q_0,w)$$

2. הגדרת פונקציית מעברים לאסל"ד. 3. הנחת האינדוקציה. 5. פונקיית מעבר לאס"ד

$$\delta_N(q,xa)=\bigcup_{p\in\delta_N(q,x)}\delta_D(p,a)=\delta_D(\delta_N(q,x),a)\ \ .4$$

כלומר:

$$w\in L(D)\iff\delta_D(\{q_0\},w)\in \boldsymbol{F}\iff\delta_{\boldsymbol{D}}(\{q_0\},w)\cap \boldsymbol{F_N}\neq\boldsymbol{\phi}$$

$$\iff\delta_{\boldsymbol{N}}(q_0,w)\cap F_N\neq\phi\iff w\in L(N)$$

$\square$

אסל"ד עם מסעי ε

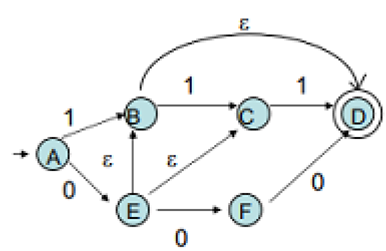
- מסעי ε מאפשרים מעבר ממצב למצב על קלט ε
- $\delta:Q\times\sum\cup\{\varepsilon\}\to P(Q)$

סגור של קבוצת מצבים

- סגור של מצב q = קבוצת המצבים שניתן להגיע אליהם מ q ע"י שימוש במסעי ε בלבד נסמן $CL(q)$
- הסגור של קבוצת מצבים P = איחוד כל הסגורים $CL(q)$ כאשר q שייך ל P
  - בסיס: $\delta'(q,\varepsilon)=CL(q)$
  - צעד: $\delta'(q,xa)=\bigcup_{p\in\delta'(q,\varepsilon)}CL(\delta(p,\varepsilon)$

שיעור 8 - 06/12/18

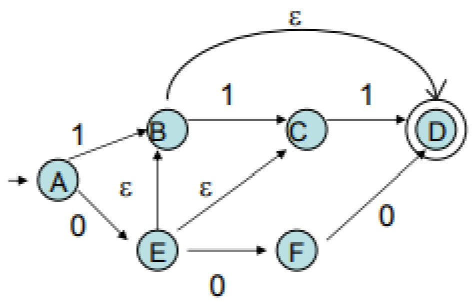
בשיעור שעבר הראינו מודל נוסף אסלד עם מסעי ε , לדוגמה:



הרחבת מילים של פונקציות מעברים:

באינדוקציה:

- <u>בסיס:</u> $\delta' = (q,\varepsilon) = CL(q)$
- <u>צעד:</u> $\delta'(q,xa) = \bigcup\limits_{p\in(q,x)} CL(\delta(p,a))$
- <u>הרעיון:</u> $\delta'(q,w)$ היא קבוצת המצבים שאליהם ניתן להגיע מ q ע"י קריאת w ושימוש אפשרי במסעי ε בכל שלב.



•בדוגמא: $\delta'(A,\varepsilon) = CL(A) = \{A\}$

$\delta'(A,0) = \cup_{p\in\delta'(A,\varepsilon)} CL(\delta(p,0))=$

$CL(\delta(A,0))= CL(\{E\}) = \{B, C, D, E\}$

$\delta'(A,01) = \cup_{p\in\delta'(A,0)} CL(\delta(p,1))=$

$\cup_{p\in\{B, C, D, E\}} CL(\delta(p,1))=$

$CL(C) \cup CL(D) = \{C,D\}$

הגדרה: $\sum^{+}$ כל המילים בשפה המכילות לפחות אות אחת (המילה הריקה בחוץ)

<u>הגדרה פורמלית:</u>

שפה של אוטומט לא דטרמיסנטי עם מסעי ε היא קבוצת w עבורן $\delta'(q_0,w)$ מכילה מצב מקבל.

<u>טענה: אסל"ד ללא מסעי ε $\iff$ אסלד עם מסעי ε :</u>

$\Leftarrow$ כל אוטומט לא דטרמינסטי ללא מסעי ε הוא גם אסלד עם ε , כנדרש.

$\Rightarrow$ הרעיון: נחבר את ה ε עם מסע רגיל.

- נתון אסל"ד עם מסעי ε שלו קבוצת מצבים Q , א"ב $\sum$, מצב התחלתי מצבים q_0 , קבוצת מצבים מקבלים F_E ופונקצית מעברים δ_E
- נרצה לבנות אסל"ד רגיל:
  - נוכל לקחת את $Q,\sum,q_0$ כמות שהם.
  - נגדיר F_N ו δ_N :
    - קבוצת המצבים שניתן להגיע אליהם מ q ע"י שימוש במסעי ε , וקריאת a =$\delta_N(q,a) = \delta'_E(q,a) = \bigcup_{CL(q)}(\delta_E(p,a))$
    - F_N אם ε לא שייך לשפת האוטומט אז $F_N = F_E$

      אם ε שייך לשפת האוטומט , אז $F_N = F_E \cup \{q_0\}$
- לכל $w\in\sum^{+}$ נוכיח באינדוקציה על $|w|$ ש: $\delta'_N(q_0,w) = \delta'_E(q_0,w)$

  <u>בסיס:</u> עבור $|w|=1$, $w=\sigma$ אז:

$$\delta'_N(q_0,\sigma) \overset{1}{=} \delta'_N(q_0,\sigma) \overset{2}{=} \delta'_E(q_0,\sigma)$$

1. הגדרת δ'_N . 2. הגדרת δ_N

<u>צעד:</u> נניח כי נכון עבור <u>u השונה מ־ε</u> , ונוכיח נכונות עבור $w=u\sigma$ ונוכיח עבור $w=u\sigma$.

$$\delta'_N(q_0,u\sigma) \overset{1}{=} \delta_N(\delta'_N(q_0,u),\sigma) \overset{2}{=} \delta_N(\delta'_E(q_0,u),\sigma) \overset{3}{=} \delta'_E(\delta'_E(q_0,u),\sigma) \overset{1}{=} \delta'_E(q_0,u\sigma)$$

1.הגדרת פונקצית מעברים מורחבת . 2.הנחת האינדוקציה. 3. הגדרת האוטומט/הנחת אינדו'

- נותר להראות ששני האוטומטים מקבלים את אותן מילים, כלומר:

$$\delta'_N(q_0,w)\cap F_N\neq\emptyset \iff \delta'_E(q_0,w)\cap F_E\neq\emptyset$$

  - עבור $w=\varepsilon$, נובעת מהגדרת F_N
  - עבור $w\neq\varepsilon$ נובעת מההוכחה באינדוקציה והגדרת F_N

סגירות תחת פעולות רגולריות

נשתמש בכלי החדש שלנו אסל"ד עם מסעי ε , הרעיון:

- יהיו L_1,L_2 שפות רגולריות לא ריקות
- יהיו $i=1,2$ $A_i=(Q,\sum,q_{0i},\delta_i,\{q_{fi}\})$ אוטומטים המקבלים אותן כך ש $Q_1\cap Q_2\neq\emptyset$ ו q_{fi} אין קשתות יוצאות - נעשה זאת ע"י הוספת מצב מקבל שמקבל רק מסעי ε, וכל מצב מקבל נשלח למצב שנוסף עם מסע ε.

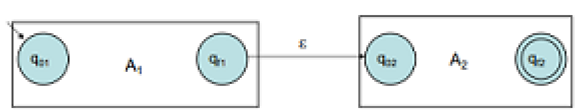
סגירות לאיחוד:

נבצע את אותו רעיון, רק נוסיף את המצב בהתחלה, והקשתות יצאו ממנו.

סגירות לשרשור

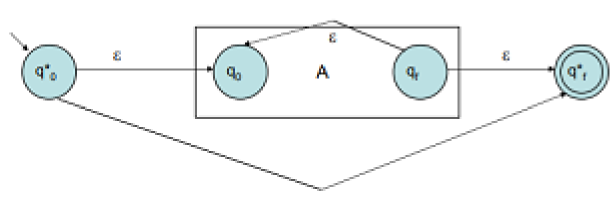
צריך להראות הכלה דו־כיוונית

- $L_1\cdot L_2$ הוא השפה שכל מילה בה מורכבת, ונבנה זאת כך:



סגירות לאיטרציה:

נבנה את האוטומט הבא שיקבל את L^* :



הסיבה שיש "מסלול נפרד" ל מילה ריקה, כיוון שזה מקל על ההוכחה.

הסבר:

כיוון ראשון: $w \in L^* \Rightarrow \exists i \in \mathbb{N}\ w \in L^*$

- עבור $w = u_1, u_2, ..., u_i$

$$\Rightarrow \exists u_1, u_2, ..., u_i \in L \Rightarrow u_1, ..., u_i \in L(A)$$

- קיים מסלל חישוב שמקבל את w:
  - אם $w = \varepsilon$ אז $q_0^+ \to q_f^+$
  - ואם $w \neq \varepsilon$ אז: $q_0^+ \xrightarrow{\varepsilon} q_0$

כיוון השני אם $w \in L(A)$ אז $w \in L^*$

- אפשרות 1 : אם $q_0^+ \underset{\rightarrow}{\varepsilon} q_f^+$ אז $w = \varepsilon$ ואכן $\varepsilon = L^*$
- אפשרות 2:

יהי i מס' הפעמים במסלול חישוב מקבל של wשהגענו ל q_f $(i > 0)$, נסמן ב u_i את המילה שקרא שאוטומט מ q_0 עד הפעם הראשונה שהגיע ל q_f , באופן כללי נסמן ב u_j את המילה שקרא האוטומט מהפעם ה j שהגיע ל q_0 עד הפעם ה j שהגיע ל q_F

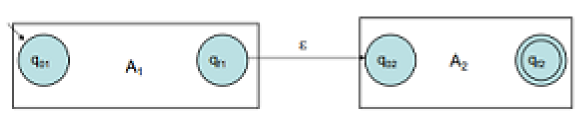
$$q_0^+ \xrightarrow{\varepsilon} q_0 \xrightarrow{u_i} q_f \xrightarrow{\varepsilon} q_0 \xrightarrow{u_2} q_f \dashrightarrow \xrightarrow{\varepsilon} q_0 \xrightarrow{u_i} q_f \xrightarrow{\varepsilon} q_f^+$$

- לכן $w = u_1, u_2, ..., u_i$, מקיימת$w \in L^*$, כנדרש.

שיעור 9 - 13/12/18

סגירות לשרשור

- יהיו L_1, L_2 השרשור $L_1 \cdot L_2$ הוא השפה שכל מילה מורכבת משני חלקים (הראשון מ L_1 , השני מ L_2)
- יהיו $i = 1, 2$ עם $A_i = (\sum, Q_i, q_{0i}, \{q_{xi}\}, \delta_i)$ האוטומטים המקבלים את L_1, L_2
- נניח כי קבוצת המצבים של האוטומטים הינן זרות ולכל אחד מהאוטומטים מצב מקבל יחיד, אשר אין ממנו קשתות יוצאות.
- נבנה את האוטומט הבא שיקבל את $L_1 \cdot L_2$



- $Q = Q_1 \cup Q_2$
- $F = \{Q_{F_2}\}$
- $\delta(q_{f1}, \varepsilon) = [q_{02}]$

- $\delta(q,a)=\begin{cases}\delta\left(q,a\right) & \text{if } q\in Q_1\\ \delta\left(q,a\right) & \text{if } q\in Q_2\end{cases}$ שאר פונקצית המעברים זהה לאלו הקיימות ב A_1, A_2

כעת, יש להראות ש $L(A)=L_1\cdot L_2$

<u>כיוון ראשון: $L(A)\subseteq L_1\cdot L_2$:</u>

- תהי $w\in L(A)$, אז $\delta\left(q_{01},w\right)=\{q_{f2}\}$
- כל מסלול חישוב מקבל של w בהכרח עובר דרך הקשת $q_{f_1}\xrightarrow{\varepsilon}q_{02}$
- המילה שקרא האוטומט מq_{01} ועד q_{f1} תקרא u
- המילה שקרא האוטומט מ q_{02} ועד q_{f2} תקרא v
- קיבלנו $w=u\cdot\varepsilon\cdot v=u\cdot v$ לכן $w\in L_1\cdot L_2$

<u>כיוון שני: $L_1\cdot L_2\subseteq L(A)$</u>

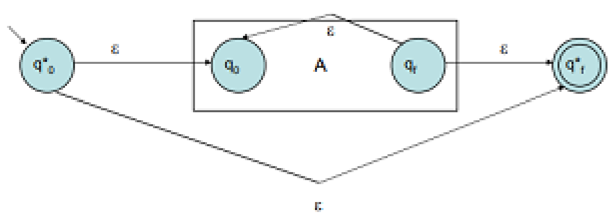
- באופן דומה (מהסוף להתחלה)

סגירות לאיטרציה

$L^*=\bigcup\limits_{i\in\mathbb{N}}L^i$, כאשר $L^i=L\cdot....\cdot L$ ו $L^0=\{\varepsilon\}$

- הוכחה באינדוקציה
- הטענה נכונה לאיחוד סופי
- למשל $L_i=\left\{0^i1^i\right\}$ רגולרית, $\cup L_i=\{a^nb^n|n\in\mathbb{N}\}$ אינה רגולרית

נבנה את האוטומט הבא שיקבל את L^* :



- נוסיף לאוטומט ε שימשיך למצב q_0 , נרצה להוסיף גם מצב מקבל חדש ומצב התחלה חדש, כעת נראה ש: $L^*=L(A)L$

<u>$L^*\subseteq L(A)$:</u>

- אם $\varepsilon\neq w\in L^*$ נראה מסלול חישוב מקבל עבור קיים k כך ש $w=w_1w_2,....w_k$. $k\geq 2$ לכל w_i
- יש מסלול מקבל באוטומט A ,
- $(q_0,w_k)=\{q_f\}$ - כי - $w_i\in L$ - לכן :

$$q_0^+\xrightarrow{\varepsilon}q_0\xrightarrow{w_i}q_f\xrightarrow{\varepsilon}q_0\xrightarrow{w_2}q_f\dashrightarrow\xrightarrow{\varepsilon}q_0\xrightarrow{w_i}q_f\xrightarrow{\varepsilon}q_f^+$$

- וזהו מסלול מקבל ל w ולכן $w \in L(A)$

<u>$L(A) \subseteq L^*$:</u>

- נניח ש A^* מקבל את w , ונתבונן במסלול חישוב מקבל.
- מקרה א': המסלול מבצע ε : מ q_0^* ל q_f^* $\Leftarrow$ כלומר $w = \varepsilon$ ולכן $w \in L^*$
- מקרה ב': המסלול מגיע $k \geq 1$ פעמים ל q_f, ולאחר כל ביקור ב q_f פרט לביקור האחרון מבצע מסע ε ל q_0. במקרה זה ניתן לרשום $w = w_1 w_2, ..., w_k$, כאשר כל w_i מתקבלת על ידי A . לכן במקרה זה $w \in L^*$

סגירות להיפוך

תהי L שפה רגולרית. ההיפוך של L, L^R היא שפת כל המילים ב L הרשומות מהסוף להתחלה

- בהנתן אוטומט המקבל את L נבנה אוטומט המקבל את R ע"י <u>הפיכת כיווני החיצים</u>, ומצב התחלתי והמקבלים.

$$\delta_R(q,a) = \{p \in Q | q \in \delta(p,a)\}$$

- נראה באינדוקציה - ע"י טענת עזר - הרעיון להוסיף מצב התחלתי חדש, ולהוציא ממנו מסעי ε- אינדוקציה על $|w|$:
  - טענת העזר: לכל $w \in \sum^*$ ו $q \in Q$ מתקיים $q \in \delta(q,w) \iff q \in \delta_R(q, w^R)$
    כיוון ראשון: $q \in \delta(q,w) \Rightarrow q \in \delta_R(q, w^R)$
  - <u>בסיס:</u> $\delta(p,a) = \{p\}$ מכאן ש $\delta_R(p, \varepsilon^R) = \{p\}$
    * צעד: נניח כי הטענה נכונה עבור $|u| = n-1$ונוכיח ל $w = ua$, $|w| = n$:
  - יהי $q = \delta(q,w) = \delta(q,ua) = \delta(\delta(p,u),a)$, כלומר:

$$p \xrightarrow{ua} q : \ p \xrightarrow{u} r \xrightarrow{a} q$$

  - באוטומט החדש כל שעשינו זה להפוך את החיצים, ולכן מהנחת האינדוקציה על u נקבל כי $p \in \delta(r, u^R)$ $r \in \delta(q,u)$
  - מכאן נקבל כי $p \in \delta_R(q, au^R) = \delta_R(q, w^R)$
  - כיוון שני $q \in \delta_R(q, w^R) \Rightarrow q \in \delta(q,w) \Rightarrow$ באופן דומה
- כעת: $L(A_R) = L^R$

  <u>כיוון 1:</u>

  - תהי $w \in L^R$ לכן $w^R \in L$ לכן $w^R \in L(A)$,
  - לכן יש מסלול חישוב מקבל מ q ל q_F של w^R
  - לפי טענת האינדוקציה יש מסלול חישוב מקבל של $w = (w^R)^R$ מ q_F ל q_0
  - לכן $w \in L(A_R)$

  <u>כיוון 2 :</u>

  - באופן דומה (מלמטה למעלה)

למת הניפוח לשפות רגולריות

תהי L שפה רגולרית. אז קיים n ב N כך שלכל מילה ב L שאורכה לפחות n קיים פירוק מהצורה $Z = uvw$ כאשר:

- $|uv| \leq n$
- $1 \leq |v|$
- $uv^i w \in L$ לכל $0 \leq i$

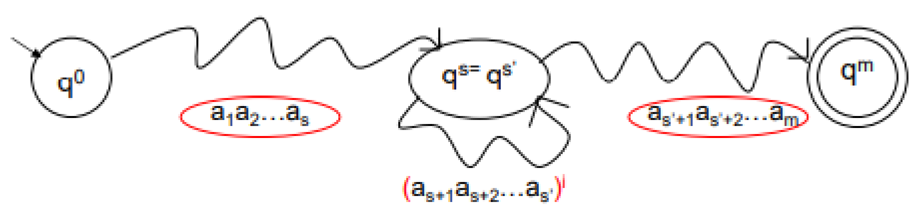
שיעור 10 - 20/12/18

הערות:

- נשים לב שבאופן מעשי הלמה מתייחסת **לשפות אינסופיות בלבד** כי אם לדוגמה נקח את $L = \{0, 011\}$ (שפה סופית) ואת $n = 4$, הלמה תתקיים, באופן ריק.
- ובעבור $n = 3$, כלומר בחרנו n "לא טוב", אז למשל : $011 = uvw$, ואכן $|vu| \leq 3$,את $u = \varepsilon$ אז נקבל נקבל $v = 0$, $w = 1$ ו L , $uv^2w = \varepsilon 001 \notin L$, וכד'.
- חזרה על הוכחת $L = \left\{0^1 1^n | n \in \mathbb{N}\right\}$ - אינה רגולרית, כפי שהוצגה בשיעורים קודמים

הוכחה: מעין עיבוד של הוכחת $0^n 1^n$

- תהי L שפה רגולרית ויהי $A\left(Q, \sum, q_0, \delta, F\right)$ אס"ד המקבל אותה
- נבחר $n = |Q|$
- תהי $z = a_1 a_2, ..., a_m$ מילה כלשהי מL כך ש: $m \geq n$
- לכל $0 \leq i \leq m$ נסמן $q^i = \delta\left(q_0, a_1, a_2, ..., a_i\right)$, כאשר $q^0 = q_0$
- כיוון ש $n = |Q|$ (מעקרון שובך היונים) , קיימים זוג אינדקסים s, s' : $0 \leq s < s'$ כך ש: $q^{s'} = q^s$
- נתבונן בחישוב האוטומט על המילה $z = a_1 a_2 .... a_m$



- מכך ש $z \in L$ כלומר q^m נובע שגם $a_1 ..... a_s a_{s'+1} .... a_m \in L$
- כעת נרצה באינדוקציה ש $a_1 ..... a_s \left(a_{s'+1} a_{s'}\right)^i a_{s'+1} .... a_m \in L$, לכל $i \geq 0$
- נרשום: $\underbrace{a_1 ..... a_s}_{u}, \underbrace{a_{s'+1} ...... a_{s'}}_{v}, \underbrace{a_{s'+1} .... a_m}_{w}$, מתקיים:
  - $|uv| = s' \leq n$ - מאיך שהגדרנו s, s'
  - $|v| \geq 1$ כי $s' > s$
  - נותר להראות שלכל $i \geq 0$: $uv^i w \in L$

- אנו יודעים ש $\delta\left(q^{s},v\right)=q^{s}$ $(*)$

- טענה: לכל i $\delta\left(q^{s},v^{i}\right)=q^{s}$, נוכיח באינדוקציה על i :

  בסיס: $i=0$ אז $v^{0}=\varepsilon$, ו $\delta\left(q^{s},\varepsilon\right)=q^{s}$ מהגדרה.

  צעד: נניח כי $\delta\left(q^{s},v^{i}\right)=q^{s}$, כעת:

$$\delta\left(q^{s},v^{i+1}\right)\overset{\text{def'}}{=}\delta\left(q^{s},v^{i}\cdot v\right)\overset{\text{def'}}{=}\delta\left(\delta\left(q^{s},v^{i}\right),v\right)\overset{\text{ind'}}{=}\delta\left(q^{s},v\right)\overset{*}{=}q^{s}$$

- כלומר הראנו שהליכה מספר כלשהו של פעמים בלולאה, מחזירה אותנו לאותו מצב, פורמלית : $\delta\left(q^{s},v^{i}\right)=q^{s}$ לכל $i\geq 0$

- ואנחנו יודעים ש $\delta\left(q^{s},w\right)\in F$, וכן $\delta\left(q^{0},u\right)=q^{s}$

- כעת:

$$\delta\left(q^{0},uv^{i}w\right)=\delta\left(\delta\left(\delta\left(q^{0},u\right),v^{i}\right),w\right)=\delta\left(\delta\left(q^{s},v^{i}\right),w\right)\overset{\text{lema}}{=}\delta\left(q^{s},w\right)=q^{m}\in F$$

$\square$

הערות:

- ניתן להוכיח את למת הניפוח עם שינויים קלים, למשל $|uw|\leq n$

- מכיון שהלמה חד - כיוונית, אם שפה מקיימת את הלמה, אין זה אומר שהיא רגולרית.

- בד"כ נשתמש בה על מנת להוכיח אי־רגולריות, ולכן נשתמש בדכ **בשלילת** הלמה, כמו בניסוח דלהלן:

נניח בשלילה ש L רגולרית יהי n מס' המובטח מלמת הניפוח. נראה ש**קיימת** מילה $z\in L$ המקיימת $|z|\geq n$ ש**לכל** פירוק $z=uvw$, יסתור את התנאים בלמת הניפוח:

1. $|uv|\leq n$ $(uv>n)$
2. $1\leq|v|$ $(1>|v|)$
3. $uv^{i}w\notin L$ **קיים** $0\leq i$

מדה־מורגן, מספיק להראות , שפירוק המקיים את 1,2 סותר את תנאי 3. (אם ניקח מילה שלא מקיימת את 1 או 2 - אז היא מראש לא בשפה) .

דוגמאות:

- $0^{n}1^{n}$:

  - יהיה $z=uvw$: לפי תנאי 1 ו - 2 : $1\leq t\leq n$ עם $v=0^{t}$
  - עבור $i=0$,המילה $uv^{i}w=uv^{0}w=uw=0^{n-t}1^{n}$
  - (לא חובה) עבור $i=2$, המילה:$uv^{i}w=uv^{2}w=uw=0^{n-t}0^{2t}1^{n}=0^{n+t}1^{n}$

- $L=\{x\in\{a,b\}^{*}\,|\#_{a}(x)=\#_{b}(x)\}$ אינה רגולרית. (# - "כמות ה")

  - נניח בשלילה כי L רגולרית, ויהי n הקבוע שקיומו מובטח בלמת הניפוח.
  - נבחר $z=a^{n}b^{n}$ ברור כי $z\in L$ ו $|z|\geq n$
  - ונמשיך כמו מקודם.

- $L = \{xx | x \in \{a,b\}^*\}$ אינה רולגרית.
  - נניח בשלילה כי L רגולרית ויהי n הקבוע שקיומו מובטח בלמה.
  - נבחר $z = a^n b a^n b$. ברור כי כי $z \in L$ ו $|z| \geq n$
  - יהיה $z = uvw$ פירוק המובטח מהלמה.
  - נבחר את $i = 0$ ונקבל: $uv^iw = uv^0w = a^{n-|v|}ba^nb \notin L$
- $L = \left\{a^{k^2} | k \in \mathbb{N}\right\}$ אינה רגולרית:
  - נניח בשלילה כי L רגולרית והיה n הקבוע שהמובטח בלמה.
  - נבחר $z = a^{n^2} \in L$
  - כייון ש $n^2 \geq n$ מובטח כי $|z| \geq n$ כנדרש.
  - יהיה $z = uvw$ כמובטח מהלמה. נסמן $t = |v|$ שמתנאים $1,2$ $1 \leq t \leq n$
  - עבור $i = 2$ המילה $z_2 = uv^iw = uv^2w = a^{n^2+t}$
    * נותר להראות כי $n^2 + t$ אינו ריבוע שלם, ולכן $a^{n^2+t} \notin L$
    * $n^2 + t > n^2 \Leftarrow n^2 + t \leq n^2 + n < (n+1)^1$
    * כלומר a^{n^2+t} נופל בין שני ריבועים שלמים, ולכן המילה אינה בשפה.

שיעור 11 - 27/12/18

חסר 10 דקות ראשונות

נראה דוגמה לשפה לא רגולרית הניתנת לניפוח - כלומר קיימות שפות לא רגולריות שעבורן מתקיימת למת הניפוח:

השפה $L = \{a\}^* \cup \left\{b^i a^{k^2} | j,k \in \mathbb{N}\right\}$

נראה בעבור L מתקיימת למת הניפוח עם הקבוע $n = 1$

- עבור כל $z \in L$ עם $|z| \geq 1$, נגדיר פירוק uvw כך ש $|v| = 1$ $u = \varepsilon$,
  - אם $z \in \{a\}^*$ אז לכל i גם $z_i = un^iw \in \{a\}^*$
  - אם z מהצורה $b^ik^{k^2}$ כאשר $j > 0$, אז לכל i גם z_i מצורה.

השפה $L = \{a\}^* \cup \left\{b^i a^{k^2} | j,k \in \mathbb{N}\right\}$ <u>אינה רגולרית</u>

- נניח בשלילה כי L רגולרית , אז מסגירות להפרש נקבל:
- $L' = \left\{b^j a^{k^2} | j,k \in \mathbb{N}^+\right\} = L \setminus \{a\}^* = L \cap \overline{\{a\}^*}$ רגולרית
- כעת מסגירות להפיוך נקבל כי גם $(L')^R = \left\{a^{k^2} b^j | j,k \in \mathbb{N}^+\right\}$ רגולרית
- וזו סתירה כי הראינו בדוגמה קודמת כי השפה $L = \left\{a^{k^2} | k \in \mathbb{N}\right\}$ אינה רגלורית (בסתירה ללמת הניפוח) , ולכן נוכל לכתוב כמעט את אותה הוכחה.

ביטויים רגולרים

מודל מסוג אחר לגמרי שניתן לתאר איתו שפות.

הגדרה: אוסף הביטויים הרגולרים מעל א"ב $\sum$ המסומן ב $R_{\sum}$ מוגדר **באינדוקצייה מבנית** באופן הבא:

אטומים:

- $\phi, \varepsilon \in R$
- $\forall \sigma \in \sum, \sigma \in R$

פעולות יצירה:

- אם $r_1, r_2 \in R$ אז $(r_1 + r_2) \in R$, $(r_1 \cdot r_2) \in R$
- אם $r \in R$ אז $(r^*) \in R$

דוגמאות:

- ϕ, ε, a, b
- $(\varepsilon + b), (\varepsilon + b) \cdot b$

הגדרת שפה: נגדיר את הפונקציה L מ R ל 2^{Σ^*}:

- תהי $L[r]$ השפה שמציין הביטוי r
- נגדיר את הפונקציה L מ R ל 2^{Σ^*}
- $L[\phi] = \phi$
- $L[\varepsilon] = \{\varepsilon\}$
- לכל $\sigma \in \sum$ מתקיים $L[\sigma] = \{\sigma\}$
- אם $r_1, r_2 \in R$ אז:
- $L[(r_1 + r_2)] = L[r_1] + L[r_2]$
- $L[(r_1 \cdot r_2)] = L[r_1] \cdot L[r_2]$
- אם $r \in R$ אז $L[(r^*)] = (L[r])^*$

אם r ביטוי רגולרי נסמן ב (r^+) את הביטוי $r \cdot (r^*)$

סדר קדימויות בהשמטת סוגריים:

- $*$
- $\cdot$
- $+$
- לרוב נשמיט את אופרטור השרשור
- לעיתים נשתמש בביטוי הרגלורי לציון השפה

דוגמאות נוספות:

- שאלה: נתונה שפה $\sum=\{a,b,c\}$ האם $\sum^*=a*b*c*$?

נשים לב ש $ca\in a*b*c*$, $ca\notin a*b*c*$, ומכאן שלא.

שאלה: האם $\sum^*=a*+b*+c*$? גם לא. כי $ab\notin a*b*c*$

שאלה: האם $\sum^*=(a+b+c)*$? כן.

- $(a+b)(a+b)^*=(a+b)^*$ כל המילים מלבד ε
- שפות כל המילים באורך זוגי מעל $\sum=\{a,b\}$: $(\sum\sum)^*=((a+b)\cdot(a+b))^*$

שקילות לאוטומטים

טענה: לכל $r\in R$ מעל $\sum$ $L[r]$ היא שפה רגולרית.

הוכחה באינדוקציה מבנית על r (על הצורה שנבנה את הביטוי)

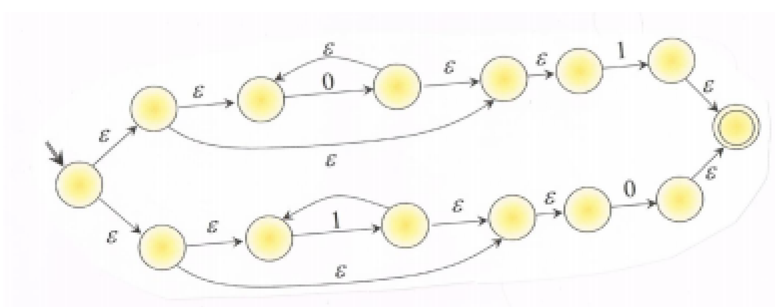
בסיס:

- $L[\phi]=\phi$ רגולרית מהגדרה.
- $L[\varepsilon]=\{\varepsilon\}$ רגולרית מהגדרה.

צעד: הרעיון נובע ישירות מסגירות השפות הרגולריות תחת פעולות רגולריות. כלומר נניח כי נטענה נכונה בעבור ביטויים r_1,r_2 ונוכיח עבור כל האפשרויות לביטווים רגולרים

- רגולרית מסגירות לאיחוד - $L[(r_1+r_2)]=L[r_1]\cup L[r_2]$
- רגולרית מסגירות לשרשור - $L[(r_1\cdot r_2)]=L[r_1]\cdot L[r_2]$
- רגולרית מסגירות לאיטרציה - $L[(r_1^*)]=(L[r_1])^*$

דוגמה לצד זה של ההוכחה: 0^*1+1^*0



שיטת הבניה:

- ראשית נלך ל + כלומר לאיחוד ולכן ישנה הפצלות
- נבנה את את השרשור
- נבנה את האיטרציה

שיעור 12 - 03/01/19

טענה: לכל שפה רגולרית $L \subseteq \sum^*$ **קיים ביטוי רגולרי כך ש** $L[r] = L$

הוכחה:

- L רגולרית לכן קיים DFA כך ש $A = (\sum, \{q1, ..., q_m, \boldsymbol{q_1}, \delta, F\})$ כך ש $L(A) = L$.
- הערות קדם הוכחה:
  - q_1 ולא q_0 , לשם נוחות בהוכחה.
  - בעבר הגדרנו "שפה של מצב" כך: $L(q) = \{w|\delta(q_1, w) = q\}$ - כלומר כל המילים שמסיימים לקרוא אותן במצב q .
  - ומתקיים ש: $L(A) = \bigcup_{q_0 \in F} L(q_j$
- לכל i, j, kנסמן ב $L_{i,j}^k$ את השפה הכוללת את המילים שמובילות את האוטומט מ q_i ל q_j בלי לעבור דרך מצב שמספרו גדול מ k - לא כולל המצב ממנו יצאנו ואליו מגיעים - כלומר:

$$L_{i,j}^k = \{w : \delta(q_i, w) = q_j,\ \forall u, v \neq w, uv = w,\ \delta(q_i, u) = q_l \to l \leq k\}$$

- כיון שאין מצב גדול מ m הרי ש $L_{i,k}^m$ כולל את כל המילים המובילות את האוטומט ממצב q_i ל q_j
- בפרט מתקיים $L(A) = \bigcup_{q_j \in F} L_1^m, j$ (לכן q_1 ולא q_0)
- השפה $L(A)$ הינה איחוד של מספר סופי של שפות
- לכן אם נוכל למצוא לכל $L_{i,j}^k$ ביטוי רגולרי, נוכל למצוא ביטוי רגולרי עבור $L(A)$

  נעשה זאת ע"י הגדרה רקורסיבית:
- עבור $k = 0$: יתכן רק מעבר ישיר מ q_i ל q_j

  - מהגדרת השפה, $L_{i,j}^0$ כל w היא בעלת אורך לכל היותר 1
  - ולכן נכתוב כל האפשרויות למעבר על הקשת למשל: $\varepsilon, (\sigma_1 + \sigma_2 + ... + \sigma_n)$
- עבור $k > 0$:

  ב $L_{i,j}^k$, קיימים 2 סוגי מילים:
  - מילים שלא גורמות ל A לעבור דרך q_k - אלה מילים ששייכות גם ל $L_{i,j}^{k-1}$
  - מילים שכן גורמות ל A לעבור דרך q_k - מילים אלו נחלק ל 3 חלקים uvw:
    - רישא u המובילה את A לביקור ראשון ב q_k . u נמצאת ב $L_{i,k}^{k-1}$
    - חלק אמצעי v הגורם ל A לבצע מספר סיבובים תוך חזרה ל q_k . v נמצא ב $\left(L_{k,k}^{k-1}\right)^*$
    - סיפא w המובילה את A מביקורו האחרון ב q_k ל q_j נמצאת ב $L_{k,j}^{k-1}$
  - כלומר $L_{i,j}^k = \left(L_{i,k}^{k-1}\right) \cdot \left(L_{k,k}^{k-1}\right)^* \cdot \left(L_{k,j}^{k-1}\right)$
  - שזה שקול ל $r_{ij}^k = \left(r_{i,j}^{k-1}\right)\left(r_{kk}^{k-1}\right)^*\left(L_{k,j}^{k-1}\right)$ - (נוסחת הנסיגה)

דוגמה לבניה על בסיס הנוסחה הרקוריסיבית נמצא בסוף מצגת 4

משפט קליני

משפחת השפות הרגולריות היא הקבוצה הקטנה ביותר המכילה את כל השפות הסופיות והסגורה תחת הפעולות הרגולריות

הוכחה:

כיוון 1

- אנחנו יודעים שכל שפה סופית היא רגולרית.
- הוכחנו שקבוצת שפות רגולריות סגורה לפעולות רגולריות (איחוד, איטרציה, שרשור)

כיון 2

- נובע מהמשפט שלכל שפה רגולרית קיים ביטוי רגולרי

זהויות בין ביטויים רוגלורים

- באופן עקרוני צריך להראות הכלה דו־כיוונית
- דוגמה בסוף מצגת 4

שיעור 13 - 10/01/19 - שיעור חזרה

1. האם $L_2 = \{(01)^n (10)^n\}$ שפה רגולרית?
   - לא, נפריך ע"י למת הניפוח
   - הערות של מירה על בעיות נפצות בהוכחה
2. האם $L_3 = \{(01)^n 0 (10)^n | n \in \mathbb{N}\}$ רגולרית?

$$L_3 = \{(01)^n 0 (10)^n | n \in \mathbb{N}\} = \left\{(01)^{2n} 0 | n \in \mathbb{N}\right\}$$

$$\text{for n=2: } (01)(01)0(10)(10) = (01)^4 0$$

ולכן נבנה אוטומט

3. האם השפה $L_1 = \left\{a^i b^m c^k | i \leq m+k\right\}$ רגולרית?
   - די ברור שהשפה אינה רגולרית, השאלה איזו חלוקה כדי לעשות?
   - יהי n הקבוע מלמת הניפוח....
   - נבחר פירוק $z = a^n b^n c^0$ ואכן $n \leq n+0$ לכל פירוק $z = uvw$ כך ש $|u| \geq 1$ $|uv| \leq n$ מתקיים ש $u = a^t \; 1 \leq t \leq n,$
   - $uv^i w = a^{n+(u-1)t} b^n c^0$
   - נקח $i = 2$ נוקבל ש $uv^2 w = a^{n+t} b^n c^0 \in L$
   - $n + t > n + 0$

   אפשרות אחרת: מהנחה בשלילה להשתמש בסגירות L^R , אבל לא ככ העוזר.
4. האם $L = \left\{a^{10} b^n c^n | n \in \mathbb{N}\right\}$ רגולרית? פה נח להשתמש ב L^R
5. תהי L נגדיר $SUb(L) = \{y \in \sum^* | \exists x, z \in \sum^* s.t. xyz \in L\}$, כלומר $sub(L)$ היא השפה המכילה את כל תתי המלים של מילים בשפה L ,הוכח/הפרך:

(א) אם L לא רגולרית אז בהכרח $sub(L)$ לא רגולרית

- ד"נ : $L=\{a^nb^n|n\in\mathbb{N}\}$ לא רגולרית (הוכחנו בכיתה), ו a^*b^* שפה רגולרית
- אפשרות נוספת $L=\left\{a^{n^2}|n\in\mathbb{N}\right\}$ אינה רגולרית, ואז $subs(L)=a^*$ והיא רגולרית

(ב) אם השפות L_2,L_3,L_1 רגולריות ומתקיים ש $L_1\cup L_2=L_3$ וגם $L_1\cap L_2=L_4$ אז L_1 רגולרית

$$L_1=(L_3\backslash L_1)\cup(L_1\cap L_2)$$

- מסגירות לחתוך איחוד ומשלים L_1 רגולרית

6. L תקרא קו־סופית אם משלימתה היא שפה סופית, הוכח/הפרך:

(א) תהי L שפה קו־סופית כלשהי ותהי P שפה לא רגולרית כלשהי אז בהכרח $L\cup P$ רגולרית

- $\overline{L\cup P}=\bar{L}\cap\bar{P}$ לכן $\bar{L}\cap\bar{P}$ ולכן היא רגולרית (הוכחנו שכל שפה סופית היא רגולרית

(ב) תהי L שפה קו־סופית כלשהי ותהי P שפה לא רגולרית כלשהי. אז בהכרח $L\cap P$ רגולרית

- נבחר $L=\sum^*$ ו $P-\{a^nb^n|n\in\mathbb{N}\}$ נקבל ש $L\cap P=P$ ולכן $L\cap P$ לא רגולרית

נספח - הגדרות ומשפטים

הגדרות:

- א"ב - $\sum$ קב' סופית לא ריקה של אותיות
- מילה = סדרה סופית מעל א"ב $\sum$
- שפה: קבוצה של מילים, מעל א"ב $\sum$, נסמן ב L
- מילה ריקה: ε = מילה ללא אותיות
- שפה ריקה : $L = \emptyset$, נשים לב ש $L = \emptyset \neq \{\varepsilon\}$
- אורך של מילה w , $|w|$ = מס' האותיות ב w
- $\sum^*$ = אוסף כל המילים מעל א"ב $\sum$,
  - $\varepsilon \in \sum^*$
  - כל L מעל $\sum$ מקיימת $L \subseteq \sum^*$
- $|\sum^*| = \aleph_0$ קב' של סדרות סופיות מעל קב' סימנים סופית (לא ריקה).

פעולות על מילים

- שרשור של מילים בהנתן $w_1, w_2 \in L$ המילה $w_1 \cdot w_2$ היא המילה המתקבלת מכתיבת אותיות w_2 אחרי w_1
  - שרשור מילים אינו פעולה חילופית, כלומר $w_1 \cdot w_2 \neq w_2 \cdot w_1$
  - שרשור היא פעולה קיבוצית, כלומר : $(w_1 \cdot w_2) \cdot w_3 = w_1 \cdot (w_2 \cdot w_3)$
  - שרשור של מילים לא בהכרח משמר סגירות בשפה, כלומר יתכן $w_1 \cdot w_2 \notin L$
- חזקה של מילים:
  - הגדרה פשוטה: $i \in \mathbb{N}$, w^i: $w^i = w \cdot w \cdot w \cdot w ...$
  - הגדרה רקורסיבית: $$\begin{aligned} w^0 &= \varepsilon \\ w^1 &= w \\ w^i &= w \cdot w^{i-1} \end{aligned}$$
- היפוך $(reverse)$: w^R
  - הגדרה פשוטה: $$\begin{aligned} w &= a_1, a_1, ..., a_n \\ w^R &= a_n, a_{n-1}, ...a_1 \end{aligned}$$
  - הגדרה רקורסיבית: $w = \overbrace{a_1, a_2, ...}^{t} a_n$. במקרה ו $w = \varepsilon$ אז $w^R = w$
    $$w^R = a_n t^R$$

פעולות על שפות

תהיינה L_1, L_2 שפות מעל א"ב $\sum$ אז :

- איחוד- $L_1 \cup L_2 = \{w|w \in L_1 \vee w \in L_2\}$
- חיתוך- $L_1 \cap L_2 = \{w|w \in L_1 \wedge w \in L_2\}$
- משלים - $\bar{L_1} = \sum^* \setminus L_1$ וכן $L_1 \setminus L_2 = \{w|w \in L_1 \wedge w \notin L_2\}$
- שרשור שפות $L_1 \cdot L_2 = \{w|\exists u \in L_1, \exists v \in L_2, w = u \cdot v\}$
- איטרציה : $L^* = \bigcup_{i \in \mathbb{N}} L^i = L^0 \cup L^1 \cup ...$
- היפוך: - L^R , הגדרה: $L^R = \{w|\, w^R \in L\}$.

אוטומט סופי דטרמינסטי

- נאמר שהאוטומט **מקבל** קלט w אם הוא מסיים את קריאת w במצב מקבל .
  - אחרת נאמר שהוא **דוחה** את w.
- אוטומט סופי דטרמינסטי A נתון על ידי $A = \left(\sum_A, Q_A, q_{0A}, F, \delta_A\right)$, כאשר
  - א"ב – כל אותיות הקלט האפשריות עבור האוטומט. מספר האותיות בא"ב זה חייב להיות סופי וגדול מ-0.
    - * נסמן ב $\sum_A$
  - מצבים – כל המצבים שבהם יכול האוטומט להימצא. מספר המצבים חייב להיות סופי וגדול מ-0.
    - * נסמן ב Q_A - קבוצה סופית לא ריקה של מצבים
  - מצב התחלתי – המצב שממנו מתחיל האוטומט את מסלול החישוב על כל מילת קלט . קבוצת מצבים מקבלים – קבוצה מתוך קבוצת המצבים, המכילה 0 מצבים או יותר .
    - * q_{0A} - מצב התחלתי $q_{0A} \in Q_A$
    - * F_A - קבוצת מצבים מקבלים $F_A \subseteq Q_A$
  - פונקציית מעברים – לכל זוג של מצב ואות, פונקציה זו מתאימה מצב (אחד ויחיד) שאליו עובר האוטומט כאשר במצב זה נקראת אות זו.
    - * F_A - קבוצת מצבים מקבלים $F_A \subseteq Q_A$
- פונקציית המעבר:
  - בסיס $\delta\left(q, \varepsilon\right) = q$
  - צעד: $\delta\left(q, wa\right) = \delta\left(\delta(q, w), a\right)$ כאשר w היא מחרוזת ו - a היא אות.

שפה של אוטומט דטרמינסטי:

- אם $A = \left(\sum_A, Q_A, q_{0A}, F_A, \delta_A\right)$הוא אוטומט אז $L(A)$ היא השפה שמקבל האוטומט. = אוסף כל המחרוזות w כך ש $\delta\left(q_0, w\right)$ שייך ל F . כלומר:$L(A) = \{w \in \sum^* \,|\, (\delta_0, w) \in F\}$

  כלומר אוסף כל המחרוזות שמקבל האוטומט - אוסף כל המחרוזות המהוות מסלול ממצב ההתחלה למצב מקבל.
- לכל $q \in Q$ שפת המצב $L_A(q)$ היא השפה הבאה: $L_A(q) = \{w \in \sum^* |\delta\left(q_0 = w\right)\}$

  מתקיים $L(A) = \bigcup_{q \in F} L_A(q)$
- הגדרה: שפה L היא רגולרית אם היא מתקבלת ע"י אוטומטט סופי דטרמיניסטי.
- עבור מצב q באוטומט א"ד נאמר ש q הוא **מצב בור** אם לכל $a \in \sum$ מתקיים $\delta\left(q, a\right) = \emptyset$.

אוטומט סופי לא דטרמיניסטי

$A = (Q, \sum, q_0, \delta, F)$ כאשר:

- Q קבוצת מצבים, $\sum$ א"ב נתון, <u>פונקציית מעברים δ</u> , $q_0 \in Q$ מצב התחלתי, קבוצת מצבים מקבלים $F \subseteq Q$
- פונקציית המעברים δ: $\delta(q,a)$ היא **קבוצת** מצבים - הרחבה למחרוזת:
  - בסיס: $\delta(q,\varepsilon) = \{q\}$
  - צעד: $\delta(q,wa) = \bigcup\limits_{p\in\delta(q,w)} \delta(p,a)$, סימון - להרחבה לקבוצת מצבים: $\delta(\boldsymbol{P},w) = \bigcup\limits_{\boldsymbol{q}\in\boldsymbol{P}} \delta(q,w)$
- השפה של האוטומט היא קבוצת המחרוזות שהוא מקבל:

  כלומר $L(A) = \{w \in \sum^* | \delta(q_0,w) \cap F \neq \emptyset\}$

אסל"ד עם מסעי ε

- אסל"ד עם מסעי ε - מסעי ε מאפשרים מעבר ממצב למצב על קלט ε
  - $\delta : Q \times \sum \cup \{\varepsilon\} \to P(Q)$
- סגור של קבוצת מצבים:
  - סגור של מצב q = קבוצת המצבים שניתן להגיע אליהם מ q ע"י שימוש במסעי ε בלבד נסמן $CL(q)$
  - הסגורים של קבוצת מצבים P = איחוד כל הסגורים $CL(q)$ כאשר q שייך ל P
    * <u>בסיס</u>: $\delta'(q,\varepsilon) = CL(q)$
    * <u>צעד</u>: $\delta'(q,xa) = \bigcup\limits_{p\in\delta'(q,\varepsilon)} CL(\delta(p,\varepsilon)$

הרחבת מילים של פונקציות מעברים:

- באינדוקציה:
  - <u>בסיס</u>: $\delta' = (q,\varepsilon) = CL(q)$
  - <u>צעד</u>: $\delta'(q,xa) = \bigcup\limits_{p\in(q,x)} CL(\delta(p,a))$

<u>הרעיון</u>: $\delta'(q,w)$ היא קבוצת המצבים שאליהם ניתן להגיע מ q ע"י קריאת w ושימוש אפשרי במסעי ε בכל שלב.

- $\sum^+$ כל המילים בשפה המכילות לפחות אות אחת (המילה הריקה בחוץ)

  <u>שפה של אוטומט לא דטרמיסנטי</u> עם מסעי ε היא קבוצת w עבורן $\delta'(q_0,w)$ מכילה מצב מקבל.

ביטויים רגולרים

- <u>הגדרה</u>: אוסף הביטויים הרגולרים מעל א"ב $\sum$ המסומן ב $R_{\sum}$ מוגדר **באינדוקצייה מבנית** באופן הבא:

  <u>אטומים</u>:

  - $\phi, \varepsilon \in R$
  - $\forall\sigma \in \sum, \sigma \in R$

  <u>פעולות יצירה</u>:

- אם $r_1, r_2 \in R$ אז $(r_1 + r_2) \in R$, $(r_1 \cdot r_2) \in R$
- אם $r \in R$ אז $(r^*) \in R$

<u>הגדרת שפה:</u> נגדיר את הפונקציה L מ R ל 2^{Σ^*}:

- תהי $L[r]$ השפה שמציין הביטוי r:
  - $L[\phi] = \phi$
  - $L[\varepsilon] = \{\varepsilon\}$
- לכל $\sigma \in \sum$ מתקיים $L[\sigma] = \{\sigma\}$
- אם $r_1, r_2 \in R$ אז:
  - $L[(r_1 + r_2)] = L[r_1] + L[r_2]$
  - $L[(r_1 \cdot r_2)] = L[r_1] \cdot L[r_2]$
- אם $r \in R$ אז $L[(r^*)] = (L[r])^*$
- אם r ביטוי רגולרי נסמן ב (r^+) את הביטוי $r \cdot (r^*)$
- סדר קדימויות בהשמטת סוגריים:
  - $*$
  - $\cdot$
  - $+$
  - (לרוב נשמיט את אופרטור השרשור)

משפטים

- פעולת שרשור על שפות **אינה חלופית,** יתכן - $L_1 \cdot L_2 \neq L_2 \cdot L_1$
- פעולת שרשור על שפות היא **כן קיבוצית** $(L_1 L_2) L_3 = L_1 (L_2 L_3)$
- <u>טענה: $(L^*)^* = L^*$</u>- מסקנה מידית : אין טעם לבצע איטרציה יותר מפעם אחת.
- טענה: מתקיים $\delta(q, w_1 \cdot w_2) = \delta(\delta(q, w_1), w_2)$
- תהיה קבוצת השפות מעל $\sum$, אז | = $|P(\sum^*)| = 2^{\aleph_0}$ (מסקנה מיידית: רוב השפות אינן רגולריות)
- $L = \{0^n 1^n | \geq 1\}$ אינה שפה רגולרית
- אס"ד שקול לאסל"ד.
- אסל"ד ללא מסעי ε $\iff$ אסלד עם מסעי ε .
- **סגירות השפות הרגולריות תחת פעולות בוליאניות**
  - **הכלה** - יהיו L_1, L_2, שפות כך ש: $L_1 \subseteq L_2$
    - אם L_2 <u>רגולרית</u> אז L_1 <u>אינה בהכרח רגולרית</u> - ד"נ: $L_1 = \{a^n b^n | n > 0\} \quad L_2 = \{\sum^* \text{where } \sum = \{a, b\}\ \}$
    - אם L_1 <u>רגולרית</u> אז L_2 <u>אינה בהכרח רגורלית</u> - ד"נ: $L_1 = \{01\} \quad L_2 = \{a^n b^n | n > 0\}$
  - **משלים** - אם L רגולריות אז $\bar{L}$ רוגולרית.
  - **חיתוך** - אם L_1, L_2 רגולריות אז $L_1 \cap L_2$ רגולרית
  - **איחוד** - אם L_1, L_2 רגולרית אז $L_1 \cup L_2$ רגולרית
  - **שרשור**- אם L_1, L_2 רגולריות אז $L_1 \cdot L_2$ רגולרית.
  - **איטרציה** - אם L_1 רגולרית אז L_1^* רגולרית.
  - **היפוך** - אם L_1 רגולרית אז L_1^R

• למת הניפוח לשפות רגולריות

תהי L שפה רגולרית. אז קיים n ב N כך שלכל מילה ב L שאורכה לפחות n קיים פירוק מהצורה $Z = uvw$ כאשר:

- $|uv| \leq n$
- $1 \leq |v|$
- $uv^i w \in L$ לכל $0 \leq i$

- **שלילת למת הניפוח**

נניח בשלילה ש L רגולרית **יהי** n מס' המובטח מלמת הניפות. נראה ש**קיימת** מילה $z \in L$ המקיימת $|z| \geq n$ ש**לכל** פירוק $z = uvw$, יסתור את התנאים בלמת הניפוח:

1. $(uv > n) |uv| \leq n$
2. $(1 > |v|)\ 1 \leq |v|$
3. $uv^i w \notin L$ **קיים** $0 \leq i$

- <u>שפות לא רגולריות - דוגמאות:</u>
  - $L = \{0^n 1^n | n \in \mathbb{N}\}$, אינה רגולרית
  - $L = \{x \in \{a, b\}^* | \#_a(x) = \#_b(x)\}$ אינה רגולרית. (# - "כמות ה")

- $L = \{xx | x \in \{a,b\}^*\}$ אינה רולגרית.
- $L = \left\{a^{k^2} | k \in \mathbb{N}\right\}$ אינה רגולרית.

- לכל $r \in R$ מעל $\sum$ $L\left[r\right]$ היא שפה רגולרית, ($L[r]$ $\Longleftrightarrow$ קיום אוטמט)
- לכל שפה רגולרית $L \subseteq \sum^*$ קיים ביטוי רגולרי כך ש $L\left[r\right] = L$
- **משפט קליני** - משפחת השפות הרגולריות היא הקבוצה הקטנה ביותר המכילה את כל השפות הסופיות והסגורה תחת הפעולות הרגולריות