תוכן העניינים

1 שפות 1

2 אוטומט סופי דטרמיניסטי 4

3 אוטומט לא דטרמיניסטי 6

4 שפות רגולריות 9

5 שפות חסרות הקשר 15

6 אוטומט מחסנית 16

1 שפות

הגדרה 1: אותיות ואלפבית

- **אלפבית** היא קבוצה סופית לא ריקה של תווים (סימנים) שונים.
- כל תו באלפבית נקרא **אות**.
- אות היא היחידה הבסיסית (סימבול) ממנה מורכב האלפבית. המונח "אות" הוא כללי וכולל כל סימון מוסכם: אותיות (בעברית, בלטינית או ביוונית), ספרות, תווים מיוחדים (כגון פסיק או סוגריים), צורות גיאומטריות וכדומה.
- אלפבית בדרך כלל מסומנת ב- Σ.
- אותיות האלפבית מסומנות באותיות יווניות קטנות: $\sigma, \tau, \ldots$.
- נתון אלפבית Σ. אורך האלפבית יסומן $|\Sigma|$ ויוגדר מספר האותיות ב- Σ.

הגדרה 2: מילה (מחרוזת)

- נתון אלפבית Σ. **מילה**, w מעל האלפבית Σ היא **סדרה סופית של אותיות** של Σ.
  מילת הריקה, תסומן ε ותוגדר המילה שאינה מכילה שום אות.

  כלומר, אם Σ אז מילה, w מעל Σ יכולה להיות מילת הריקה, $w = \varepsilon$, או מילה בת אות בודדת, $w = \sigma$ כאשר $\sigma \in \Sigma$, או מילה בת k אותיות $k \geq 2$:
  $$w = \sigma_1\sigma_2\cdots\sigma_k \ , \qquad \sigma_i \in \Sigma \ , \quad 1 \leq i \leq k \ .$$
- מילה נכתבת כרצף של אותיות זו לצד זו, ללא פסיקים וללא רווחים ביניהן.
- מילים מסומנות על ידי אותיות קטנות מסוף האלפבית האנגלי: $u,\ v,\ w,\ x,\ y$.
- לעתים מילה נקראת מחרוזת.

- נתונה מילה w מעל אלבפית Σ. אורך המילה יסומן $|w|$ ויוגדר מספר האותיות ב- w.

משפט 1:

אורך מילת הריקה הוא אפס:

$$|\varepsilon| = 0 \ .$$

הגדרה 3: Σ^*

יהי Σ אלפבית. Σ^* מוגדר להיות הקבוצה של כל המילים שניתן ליצור מאותיות ב- Σ.

משפט 2: $\varepsilon \in \Sigma^*$

לכל אלפבית Σ, אם ε מילה הריקה אז מתקיים: $\varepsilon \in \Sigma^*$.

הגדרה 4: Σ^+

יהי Σ אלפבית.

$$\Sigma^+ = \Sigma^* - \{\varepsilon\} \ .$$

במילים: Σ^+ מוגדר להיות הקבוצה של כל המילים הלא ריקות שניתן ליצור מאותיות ב- Σ.

הגדרה 5: שרשור

נתון אלפבית Σ. תהי

$$x = \sigma_1 \cdots \sigma_m \ , \quad \sigma_i \in \Sigma$$

מילה מעל Σ, ותהי

$$y = \tau_1 \cdots \tau_n \ , \quad \tau_i \in \Sigma$$

מילה מעל Σ. **השרשור** של x ו- y מסומן $x \circ y$ ומוגדר להיות המילה

$$x \circ y = \sigma_1 \quad \cdots \quad \sigma_m \tau_1 \quad \cdots \quad \tau_n \ .$$

לפעמים פעולת השרשור מסומנת בלי הסימן $\circ$. כלומר $xy \equiv x \circ y$.

משפט 3: שורשור של מילה עם מילה הריקה

עבור כל אלפבית Σ, אם x מילה מעל Σ אז

$$x \circ \varepsilon = \varepsilon \circ x = x \ .$$

הגדרה 6: חזקה

יהי w מילה מעל האלפבית Σ. לכל $n \in \mathbb{N}$ (כלומר לכל מספר טבעי n) הפעולת החזקה על המילה w מסומנת w^n ומוגדרת באופן הבא:

- אם $n = 0$ אז $w^0 = \varepsilon$.
- אם $n = 1$ אז $w^1 = w$.

- אם $n \geq 2$ אז
$$w^n = \underbrace{w \circ w \quad \cdots \quad \circ w}_{n \text{ פעמים}} \ .$$

הגדרה 7: שפה

שפה L מעל אלפבית Σ היא תת-קבוצה של Σ^*. כלומר:
$$L \subseteq \Sigma^* \ .$$

הגדרה 8: פעולת שרשור על שפות

תהיינה L_1 ו- L_2 שפות מעל אלפבית Σ. השרשור של L_1, L_2 היא שפה שמסומנת $L_1 \circ L_2$ ומוגדרת באופן הבא:
$$L_1 \circ L_2 = \left\{ u \circ w \;\middle|\; u \in L_1 \ , \ w \in L_2 \right\} \ .$$
במילים: כל מילה בשפה $L_1 \circ L_2$ היא שרשור של מילה ב- L_1 ומילה ב- L_2.

הגדרה 9: פעולת החזקה על שפות

תהי L שפה מעל Σ. הפעולה L בחזקה n היא שפה שמסומנת L^n ומוגדרת באופן הבא:

- אם $n = 0$ אז $L^0 = \{\varepsilon\}$.
- אם $n = 1$ אז $L^1 = L$.
- אם $n \geq 2$ אז:
$$L^n = \underbrace{L \circ \quad \cdots \quad \circ L}_{n \text{ פעמים}} \ .$$

הגדרה 10: סגור קליין

תהי L שפה מעל אלפבית Σ.
$$L^* = \bigcup_{n=0}^{\infty} L^n \ ,$$
$$L^+ = \bigcup_{n=1}^{\infty} L^n \ .$$

הגדרה 11: פעולת ההיפוך של מילה

פעולת ה reverse של מילה הופכת את הסדר של אותיות המילה ,מהסוף להתחלה.

פורמלי, תהי w מילה מעל א"ב Σ. הפעולת ההיפוך על w מסומנת w^r ומוגדרת באופן הבא:

- אם $w = \varepsilon$ אז $w^r = \varepsilon$.
- אם $w = \sigma$ כאשר $\sigma \in \Sigma$ אות בודדת אז $w^r = \sigma = w$.
- אם $w = \sigma_1\sigma_2 \ \cdots \ \sigma_{n-1}\sigma_n$ אז $w^r = \sigma_n\sigma_{n-1} \ \cdots \ \sigma_2\sigma_1$.

לפעמים פעולת ההיפוך על מילה w מסומנת $\text{reverse}(w)$.

משפט 4: פעולת ההיפוך היא רפלסיבית

תהי w מילה מעל אלפבית Σ. מתקיים:

$$(w^r)^r = w \ .$$

ניתן לנסח את המשפט באופן הבא:

$$x = w^r \qquad \Longleftrightarrow \qquad w = x^r \ .$$

הגדרה 12: פעולת ההיפוך של שפה

תהי L שפה. ההיפוך של L היא שפה שמסומנת L^r (או לפעמים $\text{reverse}(L)$) ומוגדרת באופן הבא:

$$L^r = \left\{ w^r \;\middle|\; w \in L \right\} \ ,$$

כשאר w^r מסמן ההיפוך של המילה w בהתאם עם ההגדרה 11.

במילים: L^r היא שפת ההיפוך של המילים של L.

הגדרה 13: שפות הריישות סייפות ותתי המילים

תהי $L \subseteq \Sigma^*$ שפה. נגדיר השפות הבאות שנגזרות מתוך L.

- **שפת הריישות של** L:

$$\text{prefix}(L) = \left\{ x \;\middle|\; x \in \Sigma^*,\ \exists y \in \Sigma^* :\ xy \in L \right\} \ .$$

- **שפת הסייפות של** L:

$$\text{suffix}(L) = \left\{ y \;\middle|\; y \in \Sigma^*,\ \exists x \in \Sigma^* :\ xy \in L \right\} \ .$$

- **שפת תתי המילים של** L:

$$\text{subtext}(L) = \left\{ y \;\middle|\; y \in \Sigma^*,\ \exists x, z \in \Sigma^* :\ xyz \in L \right\} \ .$$

2 אוטומט סופי דטרמיניסטי

הגדרה 14: אוטומט סופי דטרמיניסטי (DFA)

<u>הגדרה פורמלית</u>

אוטומט סופי דטרמיניסטי (באנגלית: Deterministic Finite Automaton - DFA) הינו חמישייה:

$$A = (Q, \Sigma, \delta, q_0, F) \ .$$

המרכיבים $Q, \Sigma, \delta, q_0, F$ הם מתוארים למטה.

- Q: קבוצת **מצבים** סופית.
- Σ: **האלפבית** שעליו מוגדרות המילים המוזנות לאוטומט.

- δ: **הפונקצית המעברים**,
$$\delta : Q \times \Sigma \to Q \ .$$
ניתן להגדיר את פונקציית המעברים בצורת נוסחה או בצורת תרשים מעברים.
- F: קבוצת **מצבי קבלה** (או מצבי סיום) $F \subseteq Q$.
- q_0: **מצב ההתחלה**, $q_0 \in Q$. זהו מצב שבו האוטומט נמצא בתחילתו של כל תהליך חישוב.

<u>הרצה</u>

- אוטומט מקבל כקלט מילה מעל האלפבית Σ.
- הקלט נקרא אות אות משמאל לימין.
- הקלט נקרא רק פעם אחת.
- מתחילים במצב ההתחלתי q_0.
- בכל צעד, האוטומט קורא אות ועובר ממצב הנוכחי למצב הבא לפי הפונקצית המעברים.
- החישוב מסתיים כשהקלט נגמר, כלומר, כאשר האות האחרונה של הקלט נקראת.

הגדרה 15: קבלה של מילה ע"י אוטומט

יהי A. אוטומט. אומרים כי A **מקבל מילה** w אם לאחר סריקת כל האותיות של w, האוטומט מגיע לאחד ממצבי הקבלה שלו.
אומרים כי המילה w מתקבלת על ידי האוטומט A.

הגדרה 16: פונקציית המעבר המורחבת δ^*

הפונקציית המעבר המורחבת, מסומנת $\delta^* : Q \times \Sigma^* \to Q$ היא פונקציה שמוגדרת באופן הבא (הגדרה רקורסיבית!)

א) $\delta^*(q, \varepsilon) = q$

ב) לכל $a \in \Sigma$: $\delta^*(q, a) = \delta(q, a)$.

ג) לכל $w \in \Sigma$: $\delta^*(q, wa) = \delta^*(\delta^*(q, w), a)$.

הגדרה 17: שפה של אוטומט

יהי $A = (Q, \Sigma, \delta, q_0, F)$ אוטומט סופי דטרמיניסטי. השפה של A תסומן $L(A)$ ותוגדר:
$$L(A) = \left\{ w \in \Sigma^* \;\middle|\; \delta^*(q_0, w) \in F \right\} \ .$$

הגדרה 18: אוטומטים שקולים

שני אוטומטים A, B נקראים **שקולים** אם
$$L(A) = L(B) \ .$$

סימון: $A \equiv B$.

משפט 5:

יהי A אוטומט סופי. אז קיים אוטומט סופי B כך ש:
$$L(B) = \overline{L(A)} \ .$$

הגדרה 19: שפה רגולרית

שפה L נקראת **שפה רגולרית** אם קיים אוטומט A כך ש- $L = L(A)$.

משפט 6:

אם L רגולרית אזי $\bar{L}$ רגולרית.

הגדרה 20: אוטומט המכפלה

יהיו $A_1 = (Q_1, \Sigma_1, \delta_1, q_{01}, F_1)$ ו- $A_2 = (Q_2, \Sigma_2, \delta_2, q_{02}, F_2)$ שני אוטומטים. אם $\Sigma_1 = \Sigma_2$ אז **האוטומט המכפלה** של A_1 ו- A_2 הוא אוטומט:
$$A = (Q, \Sigma, \delta, q_0, F)$$
כך ש:

- $Q = Q_1 \times Q_2$
- $\Sigma_1 = \Sigma_2 = \Sigma$
- $\delta : \Sigma \times (Q_1 \times Q_2) \to Q_1 \times Q_2$ כאשר הפונקצית המעברים δ של האוטומט המכפלה מוגדרת:
$$\delta\left((q_1, q_2), \sigma\right) = \left(\delta_1\left(q_1, \sigma\right), \delta_2\left(q_2, \sigma\right)\right) \ .$$
- $q_0 = (q_{01}, q_{02})$
- $F = F_1 \times F_2$.

3 אוטומט לא דטרמיניסטי

הגדרה 21: אוטומט לא-דטרמיניסטי סופי (NFA)

אוטומט לא-דטרמיניסטי סופי (באנגלית: Non-Deterministic Finite Automaton - NFA) הינו חמישייה:
$$A = (Q, \Sigma, \Delta, q_0, F) \ .$$
המרכיבים $Q, \Sigma, \Delta, q_0, F$ הם מתוארים למטה.

- Q: קבוצת **מצבים** סופית.
- Σ: **האלפבית** שעליו מוגדרות המילים המוזנות לאוטומט.
- Δ: **הפונקצית המעברים**,
$$\Delta : Q \times \Sigma \to P\left(Q\right) \ ,$$

כאשר $P(Q)$ היא קבוצת כל תתי-הקבוצות של Q. ניתן להגדיר את פונקציית המעברים בצורת נוסחה או בצורת תרשים מעברים.

- F: קבוצת **מצבי קבלה** (או מצבי סיום) $F \subseteq Q$.
- q_0: **מצב ההתחלה,** $q_0 \in Q$. זהו מצב שבו האוטומט נמצא בתחילתו של כל תהליך חישוב.

ההגדרהשל אוטומט לא-דטרמיניסטי זהה בכל להגדרת אוטומט דטרמיניסטי מלבד בפונקצייתהמעבר שמעתיקה זוג (q,σ) לקבוצת מצבים $\delta(q,\sigma) \subseteq Q$. משמעות הדבר היא שאם אוטומט לא-דטרמיניסטי נמצא במצב q כאשר הוא קורא תו σ הוא יוכל לעבור לכלאחד מהמצבים בקבוצה $\delta(q,\sigma)$.

הגדרה 22: קבלה ע"י אוטומט לא-דטרמיניסטי סופי

אומרים כי מילה w מתקבלת על ידי אוטומט לא-דטרמיניסטי סופי N אם לפחות אחת מסדרות החישוב של w מסתיימת במצב קבלה.

הגדרה 23: פונקציית המעבר המורחבת של אוטומט לא-דטרמיניסטי Δ^*

הפונקצית המעבר המוחרבת של אוטומט סופי לא-דטרמיניסטי תסומן

$$\Delta^* : Q \times \Sigma^* \to P(Q)$$

ותוגדר באופן הבא.

א) $\Delta^*(q,\varepsilon) = \{q\}$.

ב) לכל $\sigma \in \Sigma$:

$$\Delta^*(q,\sigma) = \Delta(q,\sigma) \ .$$

ג) לכל $w \in \Sigma^*$ ולכל $\sigma \in \Sigma$:

$$\Delta^*(q,w\sigma) = \Delta(\Delta^*(q,w) \ , \ \sigma) \ .$$

זוהי הגדרה רקורסיבית!

הגדרה 24: השפה של אוטומט סופי לא דטרמיניסטי

השפה של N הינה הקבוצה:

$$L(N) = \left\{ w \in \Sigma^* \;\middle|\; \Delta^*(q_0,w) \cap F \neq \emptyset \right\} \ .$$

כלומר: מילה w מתקבלת ע"י אוטומט סופי לא-דטרמיניסטי N אם ורק אם הקבוצה $\Delta^*(q_0,w)$ כוללת בתוכה לפחות מצב קבלה אחד.

משפט 7:

לכל שתי מילים $u,w \in \Sigma^*$:

$$\Delta^*(q,uw) = \Delta^*(\Delta^*(q,u),w) \ .$$

הגדרה 25: גרירה

נרשום $q \xrightarrow{w} q'$ אם קיימת סדרת חישוב של המילה w שמתחילה במצב q ומסתיימת במצב q'.

משפט 8: שקילות בין DFA ו- NFA

אם N אוטומט לא-דטרמיניסטי ואם השפה שלו היא $L(N)$ אז קיים אוטומט דטרמיניסטי D כך ש:

$$L(D) = L(N) \ .$$

במילים אחרות: לכל אוטומט לא-דטרמיניסטי N קיים אוטומט דטרמיניסטי D השקול אליו.

הגדרה 26: אוטומט לא-דטרמיניסטי עם מעברי ε

אוטומט לא-דטרמיניסטי עם מעברי ε מתאפיין ע"י חמשת המרכיבים של אוטומט לא-דטרמיניסטי של ההגדרה 21:

$$M = (Q, \Sigma, \Delta, q_0, F)$$

עם הבדל קטן (אך משמעותי) בפונקציית המעבר:

$$\Delta : Q \times (\Sigma \cup \{\varepsilon\}) \to P(Q) \ ,$$

לרשימת התווים Σ נוספת המילה הריקה ε.

הגדרה 27: ε - סגור

לכל מצב q נסמן על ידי $C_\varepsilon(q)$ את קבוצת כל המצבים p כך שקיים מסלול מהמצב q למצב p אשר כל צלעותיו מסומנות ε.

$$C_\varepsilon(q) \triangleq \left\{ p \in Q \ \middle| q \xrightarrow{\varepsilon} r_1 \xrightarrow{\varepsilon} \cdots \xrightarrow{\varepsilon} r_i \xrightarrow{\varepsilon} p \text{ קיים מסלול} \right\}$$

הקבוצה $C_\varepsilon(q)$ תמיד תכלול בתוכה את q.

הגדרה 28: ε - סגור

יהי $M = (Q, \Sigma, \delta, q_0, F)$ אוטומט לא-דטרמיניסטי עם מעברי- ε. אם $P \subseteq Q$ תת-קבוצה של מצבים אז

$$C_\varepsilon(P) = \bigcup_{q \in P} C_\varepsilon(q) \ .$$

הגדרה 29: פונקצית המעבר המורחבת של אוטומט עם מעברי- ε

יהי $M = (Q, \Sigma, \Delta, q_0, F)$ אוטומט לא-דטרמיניסטי עם מעברי- ε. ראשית לכל מצב $q \in Q$ נגדיר:

$$\Delta^*(q, \varepsilon) = C_\varepsilon(q) \ .$$

לכל מילה $w \in \Sigma^*$ נגדיר $\Delta^*(q, w)$ באמצעות האלגוריתם הבא.

על כל מילה $w = \sigma_1 \cdots \sigma_k$ ומצב $q \in Q$ הפונקציה Δ^* תפעיל כך:

1) מחשבת $P = \Delta^*(q, \varepsilon) = C_\varepsilon(q)$.

2) לכל $1 \le i \le k$:

- מחשבת $P \leftarrow \bigcup_{r \in P} \Delta(r, \sigma_i)$

- מחשבת $P \leftarrow \bigcup_{r \in P} C_\varepsilon(r)$

3) מחזירה P.

הגדרה 30: קבלה ע"י אוטומט לא-דטרמיניסטי עם מעברי ε

השפה $L(M)$ המתקבלת על ידי אוטומט לא-דטרמיניסטי עם מעברי- ε היא
$L(M) = \{w \in \Sigma^* \mid \text{מכילה מצב של } F \ \Delta^*(q_0, w)\}$.

משפט 9:
שקילות בין אוטומט אי-דטרמיניסטי עם מעברי- ε לבין אוטומט א"ד ללא מעברי ε

תהי L שפה. אם קיים אוטומט לא-דטרמיניסטי עם מעברי- ε, A_ε המקבל את L אז קיים אוטומט לא-דטרמיניסטי ללא מעברי- ε, A_N המקבל את L.

משפט 10:
שקילות בין אוטומט אי-דטרמיניסטי עם מעברי- ε לבין אוטומט א"ד ללא מעברי ε

תהי L שפה. אם קיים אוטומט לא-דטרמיניסטי עם מעברי- ε, A_ε המקבל את L אז קיים אוטומט לא-דטרמיניסטי ללא מעברי- ε, A_N המקבל את L.

4 שפות רגולריות

הגדרה 31: אטומט המכפלה

יהיו
$$M_1 = (Q_1, \Sigma, \delta_1, q_{0,1}, F_1)$$
$$M_2 = (Q_2, \Sigma, \delta_2, q_{0,2}, F_2)$$
שני אוטומטים דטרמיניסטיים המקבלים את השפות L_1, L_2 בהאמה. **האוטומט המכפלה** תסומן $M_1 \times M_2$ ויוגר
$$M_1 \times M_2 = (Q_1 \times Q_2, \Sigma, \delta, (q_{0,1}, q_{0,2}), F_1 \times F_2)$$
כאשר לכל $q_1 \in Q_1$ ולכל $q_2 \in Q_2$ ולכל $\sigma \in \Sigma$:
$$\delta((q_1, q_2), \sigma) = (\delta_1(q_1, \sigma) \ , \ \delta_2(q_2, \sigma)) \ .$$

הגדרה 32:

יהי Σ אלפבית נתון (קבוצה סופית של תווים) ותהיינה L_1, L_2 שפות מעל האלפבית הזה. נמנה מספר פעולות שבאמצעותן ניתן לייצר שפות חדשות משפות נתונות כזכור:

- **האיחוד** של השפות L_1, L_2 מוגדר
$$L_1 \cup L_2 = \{w \mid w \in L_1 \ \vee \ w \in L_2\} \ .$$
- **החיתוך** של השפות L_1, L_2 מוגדר
$$L_1 \cap L_2 = \{w \mid w \in L_1 \ \wedge \ w \in L_2\} \ .$$

- **השרשור** של השפות L_1, L_2 מוגדר
$$L_1L_2 = \left\{uv \;\middle|\; u \in L_1 \;\wedge\; v \in L_2\right\} \ .$$
- **החזקה** של השפה L מוגדרת:
$$\begin{aligned} L^0 &= \{\varepsilon\} \ , \\ L^1 &= L \ , \\ L^2 &= LL \ , \\ L^{n+1} &= L^nL \ . \end{aligned}$$
- **הסגור קליין** מוגדר
$$L^* = \bigcup_{n=0}^{\infty} L^n \ .$$
- **הסגור קליין פלוס** מוגדר
$$L^+ = \bigcup_{n=1}^{\infty} L^n \ .$$

משפט 11: סגירות שפות רגולריות תחת חיתוך

אם L_1, L_2 שפות רגולראיות אז $L_1 \cap L_2$ היא שפה רגולרית.

משפט 12: סגירות שפות רגולריות תחת איחוד

אם L_1, L_2 שפות רגולראיות אז $L_1 \cup L_2$ היא שפה רגולרית.

משפט 13: סגירות שפות רגולריות תחת הפרש

אם L_1, L_2 שפות רגולראיות אז $L_1 \setminus L_2$ היא שפה רגולרית.

משפט 14: סגירות שפות רגולריות תחת הפרש

אם L_1, L_2 שפות רגולראיות אז $L_1 \Delta L_2 = (L_1 \setminus L_2) \cup (L_2 \setminus L_1)$ היא שפה רגולרית.

משפט 15: סגירות שפות רגולריות תחת שרשור

אם L_1, L_2 שפות רגולראיות אז L_1L_2 היא שפה רגולרית.

משפט 16: סגירות שפות רגולריות תחת סגור קליין

אם L שפות רגולראיות אז L^* היא שפה רגולרית.

משפט 17:

כל שפה סופית היא רגולרית.

משפט 18:

אם L שפה רגולרית אז גם L^r שפה רגולרית, כאשר L^r היא השפת ההיפוך L שמוגדרת:
$$L^r = \left\{w^r \middle| w \in L\right\} \ .$$

משפט 19: סגירות שפות רגולריות

תהיינה L_1, L_2, L שפות רגולריות. אזי השפות הבאות גם רגולריות:

1) החיתוך: $L_1 \cap L_2$
2) האיחוד: $L_1 \cup L_2$
3) ההפרש: $L_1 \setminus L_2$
4) ההפרש סימטרי: $L_1 \Delta L_2$
5) השרשור: $L_1 L_2$
6) הסגור קליין:
7) השרשור: L^*
8) ההיפוך: L^r
9) הרישא: $\text{prefix}(L) = \left\{x \in \Sigma^* \middle| \exists y \in \Sigma^* : \ xy \in L\right\}$

המצבים המקבלים לכל פעולה (חיתוך, איחוד, הפרש והפרש סימטרי) מתוארים בטבלה למטה:

פעולה	שפה	מצבים מקבלים
חיתוך	$L_1 \cap L_2$	$F_1 \times F_2$
איחוד	$L_1 \cup L_2$	$(F_1 \times Q_2) \cup (Q_1 \times F_2)$
הפרש	$L_1 \setminus L_2$	$F_1 \times (Q_2 \setminus F_2)$
הפרש	$L_1 \Delta L_2$	$(F_1 \times (Q_2 \setminus F_2)) \cup ((Q_1 \setminus F_1) \times F_2)$

הגדרה 33: ביטוי רגולרי

מחרוזת r הינה ביטוי רגולרי אם היא מאחת הצורות הבאות:

בסיס האינדוקציה:

- σ, כאשר σ אות בודדת בא"ב.
- ε
- $\emptyset$

פעולות:

- r_1r_2, כאשר r_1 ו-r_2 ביטויים רגולריים.
- $r_1|r_2$, כאשר r_1 ו-r_2 ביטויים רגולריים.
- r_1^*, כאשר r_1 ביטוי רגולרי.
- (r_1), כאשר r_1 ביטוי רגולרי.

הגדרה 34:

יהי r ביטוי רגולרי. **השפה של** r , נסמן $L(r)$ הינה השפה הבאה:

בסיס האינדוקציה:

- אם $r = \sigma$, כאשר σ אות בודדת בא"ב אזי $L(r) = \{\sigma\}$
- אם $r = \varepsilon$ אזי $L(r) = \{\varepsilon\}$.
- אם $r = \emptyset$ אזי $L(r) = \emptyset$.

פעולות:

- אם $r = r_1 \circ r_2$ אזי $L(r) = L(r_1) \circ L(r_2)$
- אם $r = r_1|r_2$ אז $L(r) = L(r_1) \cup L(r_2)$.
- אם $r = r_1^*$ אז $L(r) = L(r_1)^*$.
- אם $r = (r_1)$ אז $L(r) = L(r_1)$.

הגדרה 35: אוטומט סופי מוכלל $GNFA$

<u>הגדרת החמשת מרכיבים</u>

אוטומט מוכלל הוא אוטומט לא דטרמיניסטי עם מעברי- ε:
$G = (Q, \Sigma, \delta, q_s, q_f)$

כשאר:

1) Q: קבוצת **מצבים** סופית.

2) Σ: **האלפבית** של הקלט.

3) δ: **הפונקצית המעברים,**

$$\delta : (Q \setminus \{q_f\}) \times (Q \setminus \{q_s\}) \to R$$

כאשר R האוסף של כל הביטויים הרגולריים.

δ מקבל כקלט שני מצבים q_i, q_j ומחזירה הביטויי הרגולרי שמסמן את הצלע היוצאת מ- q_i ונכנסת ל- q_j.

4) q_s המצב ההתחלתי היחיד.

5) q_f המצב המקבל היחיד.

הכללי הביצוע של $GNFA$

$GNFA$ מתבצע לפי הכללים הבאים:

- כל צלע מסומנת על ידי ביטוי רגולרי באלפבית Σ.
- קיים מצב התחלתי אחד ויחיד q_s אשר אליו לא נכנסת שום צלע.
  קיימת צלע היוצאת מ- q_s לכל מצב אחר.
- קיים מצב קבלה אחד ויחיד q_f אשר ממנו לא יוצאת שום צלע. על כל מצב אחר קיימת צלע ממנו הנכנסת למצב המקבל.
- קיימת קשת בין כל שני קדקודים, וקיימת וקשת בין כל קודקוד לעצמו, למעט העובדה שאין כניסות למצב ההתחלתי ואין יציאות מהמצב המקבל.
- המעבר ממצב q_i למצב q_j הוא אפשרי רק על ידי קליטה של מחרוזת התואמת את הביטוי הרגולרי המסמן את הצלע היוצאת מ־ q_i ל־ q_j.
- אם קיימת צלע היוצאת ממצב q_i ונכנסת למצב q_j המסומנת ע"י ביטוי רגולרי r אז $\delta(q_i, q_j) = r$.

קבלה של מילה ע"י $GNFA$

קבלה של מילה w ע"י אוטומט $GNFA$ מוגדרת באופן הבא. יהי $GNFA\ G = (Q, \Sigma, \delta, q_s, q_f)$.
אומרים כי G מקבל מילה $w \in \Sigma^*$ אם ורק אם קיים פירוק $w = w_1 w_2 \cdots w_k$, כאשר לכל $1 \leq i \leq k$: $w_i \in \Sigma^*$, וקיימים סדרת מצבים $q_0, q_1, \ldots, q_k \in Q$ כך ש:

1) $q_0 = q_s$ הוא המצב ההתחלתי,

2) $q_k = q_f$ הוא המצב המקבל,

3) לכל $1 \leq i \leq k$: $w_i \in L(r_i)$ כאשר $r_i = \delta\left(q_{i-1,q_i}\right)$.
   כלומר r_i הוא הביטוי על הצלע מ- q_{i-1} ל- q_i.

משפט 20:

לכל אוטומט סופי $A \in$ קיים אוטומט סופי מוכלל B שקול. כלומר:
$L(A) = L(B)$.

הגדרה 36: שיטת דילול מצבים

נתון אוטומט סופי מוכלל G. האלגוריתם הבא מקבל כקלט G ונותן כפלט אוטומט סופי מוכלל חדש G', עם רק שני מצבים.

$CONVERT(G) =$ " על קלט $G = (Q, \Sigma, \delta, q_s, q_f)$, מבצע השיטת הרקורסיבית הבאה:

שלב 1) יהי k מספר המצבים של G.

שלב 2) אם $k = 2$ אזי G מורכב ממצב התחלתי q_s, מצב המקבל q_f, וצלע יחיד מ- q_s ל- q_f המסומנת עם ביטוי רגולרי, r. $\Leftarrow$ מחזיר r.

שלב 3) אם $k > 2$, בוחר מצב $q' \in Q$ כאשר $q' \neq q_s$ וגם $q' \neq q_f$.

יהי G' ה- $GNFA$ הבא: $G' = (Q', \Sigma, \delta', q_s, q_f)$ כאשר

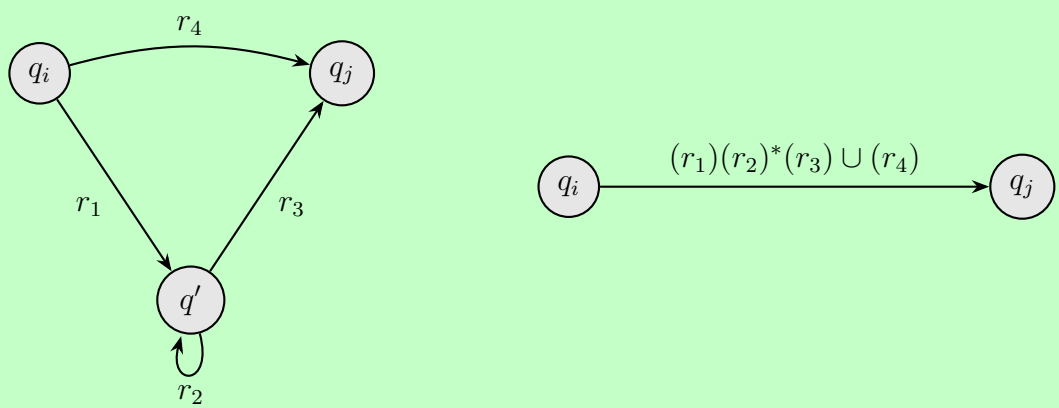
$Q' = Q \setminus \{q'\}$

ולכל $q_i \in Q' \setminus \{q_f\}$ ולכל $q_j \in Q' \setminus \{q_s\}$,

$$\delta'(q_i, q_j) = (r_1)(r_2)^*(r_3) \cup (r_4) \ ,$$

כאשר

$$r_1 = \delta(q_i, q'), \quad r_2 = \delta(q', q'), \quad r_3 = \delta(q', q_j), \quad r_4 = \delta(q', q') \ .$$



לפני אחרי

שלב 4) $CONVERT(G')$. "

משפט 21:

תהי L שפה. L רגולרית אם ורק אם קיים ביטוי רגולרי r כך ש:

$$L = L(r) \ .$$

משפט 22: למת הניפוח

אם L שפה רגולרית אז קיים מספר טבעי p **(קבוע למת הניפוח)** כך שאם $w \in L$ וגם $|w| \geq p$ אז קיים פירוק $w = xyz$ כך שהתנאים הבאים מתקיימים:

1) $|y| > 0$.

2) $|xy| \leq p$.

3) לכל $i \geq 0$ מתקיים $xy^iz \in L$.

5 שפות חסרות הקשר

הגדרה 37: דקדוק חסר הקשר

הגדרה פורמלית

דקדוק חסר הקשר (CFG) G מורכב מארבעת הדברים
$$G = (V, \Sigma, R, S) \ ,$$
כאשר:

- V קבוצת **משתנים** סופית.
- Σ אלפבית סופית
- R אוסף כללי יצירה
$$A \to \alpha$$
כאשר $A \in V$ משתנה בודד בצד שמאל ו- $\alpha \in (V \cup \Sigma)^*$ מחרוזת של משתנים וטרמינילים בצד ימין.
- $S \in V$ המשתנה ההתחלתי.

יצירה של מילה על ידי דקדוק חסר קשר

1) כתבו את המשתנה ההתחלתי S.

2) מצאו משתנה וכלל אשר מתחיל אם משתנה זה, והחליפו אותו עם המחרוזת בצד ימין של הכלל.

3) חזרו על שלבים 1 ו- 2 עד שלא נשאר אף משתנים של V.

הגדרה 38: גזירות

יהי $G = (V, \Sigma, R, S)$ דקדוק חסר הקשר ויהי $(A \to \alpha) \in R$ כלל יצירה של הדקדוק.
לכל $x, y \in (V \cup \Sigma)^*$ נסמן
$$xAy \xrightarrow{G} x\alpha y \ .$$
עבור מחרוזת $x, y \in (V \cup V)^*$ נסמן
$$x \xrightarrow{G*} y$$
אם ניתן לעבור מ- x ל- y על ידי 0 או יותר גזירות של G.

הגדרה 39: שפת הדקדוק

יהי $G = (V, \Sigma, R, S)$.
השפה של G הינה
$$L(G) = \left\{ w \in \Sigma^* \mid S \xrightarrow{G*} w \right\} \ .$$
אומרים כי שפה L היא **חסרת הקשר** אם קיים דקדוק חסר הקשר G כך ש- $L = L(G)$.

משפט 23: למת הניפוח לשפות חסרות הקשר

תהי L שפה חסרת הקשר.
קיים מספר טבעי p (קבוע למת הניפוח) כך שלכל $w \in L$ $(|w| \geq p)$ קיים פירוק $w = tuxyz$ כך שהתנאים הבאים מתקיימים:

1) $|uy| > 0$,
2) $|uxy| \leq p$,
3) לכל i טבעי: $tu^ixy^iz \in L$.

6 אוטומט מחסנית

הגדרה 40: אוטומט מחסנית (PDA)

אוטומט מחסנית הינו שישייה:
$$P = (Q, \Sigma, \Gamma, \Delta, q_s, F) \ .$$

- Q: **קבוצת מצבים סופית**:

  אלה הן המצבים בהם האוטומט עשוי להמצא בכל שלב ושלב.

- Σ: **אלפבית קלט סופי**:

  זהו האלפבית שמעליו מוגדרות המחרוזות המוזנות לאוטומט.
  נסמן ע"י Σ_ε את קבוצת האלפבית יחד עם המילה הריקה ε:
  $\Sigma_\varepsilon = \Sigma \cup \{\varepsilon\}$.

- Γ: **אלפבית מחסנית**:

  זוהי קבוצת התווים שניתן לדחוף או לשלוף מהמחסנית (stack) של האוטומט.
  נסמן ב־ Γ_ε את קבוצת האלפבית יחד עם המילה הריקה ε:
  $\Gamma_\varepsilon = \Gamma \cup \{\varepsilon\}$,
  וב- Γ^* את כל המילים שניתן ליצור באמצעות Γ.

- q_s מצב התחלתי

- F מצבים מקבלים. $F \subseteq Q$.

- Σ: **פונקציית המעברים**:

$$\Delta : Q \times \Sigma_\varepsilon \times \Gamma_\varepsilon \to P(Q \times \Gamma^*) \ ,$$
כאשר $\Sigma_\varepsilon = \Sigma \cup \{\varepsilon\}$ ו- $\Gamma_\varepsilon = \Gamma \cup \{\varepsilon\}$ ו- $P(Q \times \Gamma^*)$ הקבוצת החזקה של $Q \times \Gamma^*$.
Δ מוגדרת באופן הבא:

- משמעות הדר היא שאם P במצב q וקורא a במילת הקלט, אז $\Delta(q,a,b) = \{(q',c),\ldots\}$ באחד המעברים האפשריים, הוא שולף b מראש המחסנית ודוחף את התו c לראש האוטומט, ואז עובר למצב החדש q'.
- בתרשים מצבים מעבר כזה מסומן כמתואר למטה:

$$q \xrightarrow{a,b \to c} q'$$

- אם $a = \varepsilon$ אז P יכול לעבור מ- q ל- q' בלי לקרוא אף אות במילת הקלט, כמו ב- NFA_ε רגיל.
- אם $b = \varepsilon$ אז P יכול לדחוף אות לראש המחסנית בלי לשלוף התו בראש המחסנית.
- אם $c = \varepsilon$ אז P יכול לשלוף אות מראש המחסנית בלי לדחוף אות לראש המחסנית.

- אוטומט מחסנית הוא למעשה אוטומט לא-דטרמיניסטי רגיל בתוספת התקן זיכרון. המחסנית שמצטרפת אליו היא למעשה התקן זיכרון המאפשר לאוטומט לשמור מידע על פעולות קודמות, ולהשתמש בו בשלבים מאוחרים יותר של החישוב.
- הקורא המצוי בשפות תיכנות יזהה את המחסנית כמבנה נתונים בסיסי ידוע שנקרא לפעמים $LIFO$ שפירושו Last In First Out.
- למרות זאת, מדובר בהתקן זיכרון מוגבל מאוד מאחר ובכל שלב בחישוב, לאוטומט יש גישה רק ל- **תו הראשי** במחסנית. כלומר, האוטומט מסוגל לקרוא רק את התו הנמצא ב- **ראש המחסנית**, ואין לו שום מידע לגבי התווים האחרים במחסנית.

הגדרה 41: קונפיגורציה

יהי $P = (Q,\Sigma,\Gamma,\Delta,q_0,F)$ אוטומט מחסנית. **קונפיגורציה** של P הינה שלישיה (q,u,x):

$$(q \in Q\ ,\ u \in \Sigma^*\ ,\ x \in \Gamma^*)\ ,$$

כאשר:

- q המצב הנוכחי של האוטומט.
- u תוכן הסרט מהמיקום של הראש ולצד ימין של הראש.
- x תוכן המחסנית.

הגדרה 42: גרירה

יהי $P = (Q,\Sigma,\Gamma,\Delta,q_0,F)$ אוטומט מחסנית, ותהיינה c_1, c_2 קונפיגורציות של P.
נסמן $c_1 \vdash c_2$ (c_1 גורר את c_2) אם כשנמצאים ב- c_1 ניתן לעבור ל- c_2 ע"י מעבר יחיד.

הגדרה 43: גרירה כוכבית

יהי $P = (Q,\Sigma,\Gamma,\Delta,q_0,F)$ אוטומט מחסנית, ותהיינה c_1, c_2 קונפיגורציות של P.
נסמן $c_1 \vdash^* c_2$ אם ניתן לעבור מ- c_1 ל- c_2 באפס או יותר צעדים.

הגדרה 44: קבלת מילה ע"י אוטומט מחסנית

יהי $P = (Q, \Sigma, \Gamma, \Delta, q_s, F)$ אוטומט מחסנית. השפה של P מוגדרת להיות:

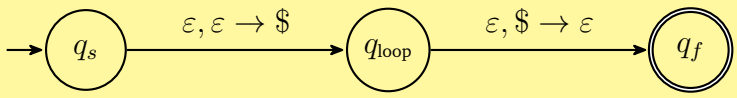
$$L(P) = \left\{ w \;\middle|\; \exists q_f \in F \; \exists u \in \Gamma^* \;:\; (q_s, w, \varepsilon) \vdash^* (q_f, \varepsilon, u) \right\} .$$

משפט 24:

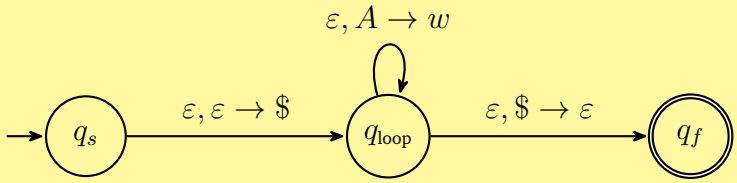
תהי L שפה חסרת הקשר. אז קיים אוטומט מחסנית P שקול שמקבל את L.

בפרט, אם $G = (V, \Sigma, R, S)$ הדקדוק היוצר השפה L. אזי האוטומט המסחנית P שקול המקבל L ניתן לבנות על פי האלגוריתם הבא.

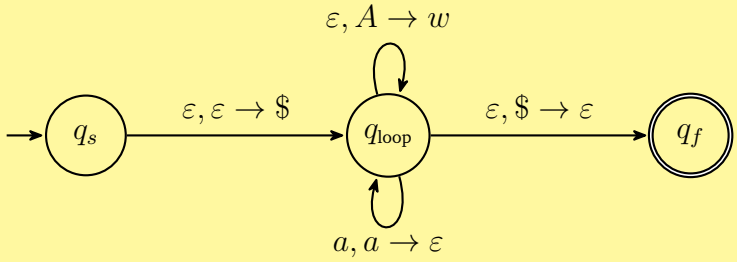
שלב 1) מתחילים עם מחסנית אוטומט 3 מצבים הבא:



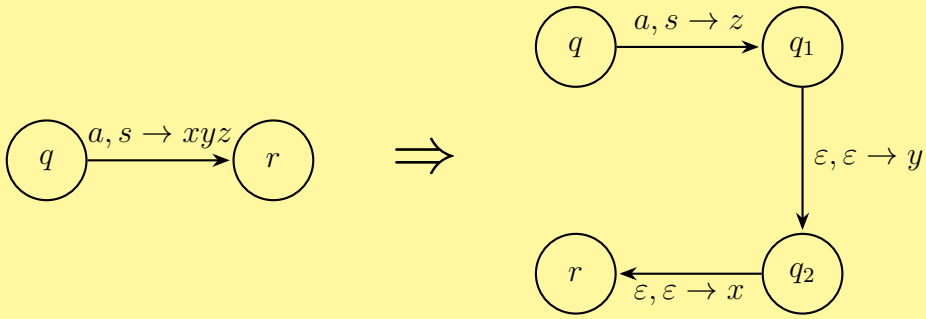
שלב 2) לכל כלל $A \to w$ של הדקדוק G, מוסיפים מעבר על המצב q_{loop}:



שלב 3) לכל אות $a \in \Sigma$ באלפבית של הדקדוק G, מוסיפים מעבר על המצב q_{loop}:



שלב 4) על כל לולאה על המצב q_{loop} מפעילים את ההמרה:



משפט 25:

יהי P אוטומט מחסנית. קיים דקדוק חסר הקשר G שקול היוצר את השפה של P.

בפרט, אם $P = (Q, \Sigma, \Gamma, \Delta, q_1, \{q_{\text{acc}}\})$ אוטומט מחסנית, אזי נבנה דקדוק שקול $G = (V, \Sigma, R, S)$ היוצר את השפה $L(P)$ על פי האלגוריתם הבא.

1) הקבוצת המשתנים היא:

$$V = \{X_{pq} \mid p, q \in Q\}$$

2) המשתנה ההתחלתי הוא: $S = X_{q_s, q_f}$.

3) נבנה קבוצת הכללים R על פי האלגוריתם הבא:

א) לכל 4 מצבים $p, q, r, s \in Q$, לכל $t \in \Gamma$ ולכל $a, b \in \Sigma_\varepsilon$:
אם $(r, t) \in \Delta(p, a, \varepsilon)$ וגם $(q, \varepsilon) \in \Delta(s, b, t)$ אז מוסיפים את הכלל $X_{pq} \to aX_{rs}b$ לקבוצת הכללים R.

ב) לכל 3 מצבים $p, q, r \in Q$, מוסיפים את הכלל $X_{pq} \to X_{pr}X_{rq}$ לקבוצת הכללים R.

ג) לכל מצב יחיד $p \in Q$, מוסיפים את הכלל $X_{pp} \to \varepsilon$ לקבוצת הכללים R.