

Formális nyelvek szóbeli tételek kidolgozása

May 23, 2023

1 A környezetfüggetlen nyelvtan definíciója, a levezetés és a nyelvtan által generált nyelv fogalma. A reguláris nyelvtan és nyelv definíciója

1.1 környezetfüggetlen nyelvtan

- N egy ábécé, nemterminális ábécé
- Σ egy ábécé, a terminális (befejező, végső) ábécé, amire $N \cap \Sigma = \emptyset$
- $S \in N$ a kezdőszimbólum (vagy start szimbólum)
- P pedig $A \rightarrow \alpha$ alakú ún. átírási szabályok véges halmaza, ahol $A \in N$ és $\alpha \in (N \cup \Sigma)^*$

→ Példa:

$$G_1 = (\{S\}, \{a, b\}, \{S \rightarrow aSb, S \rightarrow \epsilon\}, S)$$

1.2 Közvetlen levezetés (deriváció)

Tetszőleges $\gamma, \beta \in (N \cup \Sigma)^*$ esetén $\gamma \Rightarrow_G \beta$, ha van olyan $A \rightarrow \alpha \in P$ szabály és vannak olyan $\alpha', \beta' \in (N \cup \Sigma)^*$ szavak, amelyekre fennállnak, hogy $\gamma = \alpha' A \beta'$, $\beta = \alpha' \alpha \beta'$

→ Példa:

$$G_1 = (\{S\}, \{a, b\}, \{S \rightarrow aSb, S \rightarrow \epsilon\}, S)$$

$$bSa \Rightarrow_{G_1} baSba \text{ az } S \rightarrow aSb \text{ szabállyal}$$

$$baaSa \Rightarrow_{G_1} baaa \text{ az } S \rightarrow \epsilon \text{ szabállyal}$$

1.3 Levezetések

- $\gamma \Rightarrow_G \beta$: egy lépés (= a közvetlen levezetés)
- $\gamma \Rightarrow_G^n \beta$: $n \geq 0$ lépés ($\gamma \Rightarrow_G^0 \beta \Leftrightarrow \gamma = \beta$)
- $\gamma \Rightarrow_G^+ \beta$: legalább egy lépés
- $\gamma \Rightarrow_G^* \beta$: valamennyi (esetleg 0) lépés
- A $G = (N, \Sigma, P, S)$ környezetfüggetlen nyelvtan által generált nyelv: $L(G) = \{ w \in \Sigma^* \mid S \Rightarrow_G^* w \}$
→ Az $w \in \Sigma^*$ feltétel miatt w -ben nincsenek nemterminálisok, tegát nem lehet belőle "tovább lépni".
→ pl. $L(G_1) = \{a^n b^n \mid n \geq 0\}$

*A nyelv és a nyelvtan két különböző fogalom. Egy nyelv szavak egy (véges vagy végtelen) halmaza, míg egy nyelvtan egy olyan végesen specifikált (adott) eszköz, amellyel nyelvet lehet generálni.
 $L(G) = L(G')$

1.4 környezetfüggetlen nyelvek

Egy L nyelvet környezetfüggetlen nyelvnek hívunk, ha van olyan G környezetfüggetlen nyelvtan, melyre $L = L(G)$. Az összes környezetfüggetlen nyelvek halmazát CF-fel jelöljük.

1.5 Reguláris nyelvtanok

Egy $G = (N, \Sigma, P, S)$ nyelvtan reguláris (vagy jobblinéáris), ha P -ben minden szabály $A \rightarrow xB$ vagy $A \rightarrow x$ alakú. Reguláris nyelvtan esetén minden levezetés

$$A_1 \Rightarrow x_1 x_2 A_3 \Rightarrow \dots \Rightarrow x_1 x_2 \dots x_n A_{n+1}$$

alakú, ahol az alkalmazott szabályok

$$A_1 \rightarrow x_1 A_2, A_2 \rightarrow x_2 A_3, \dots, A_n \rightarrow x_n A_{n+1}$$

Levezetést csak $A \rightarrow x$ alakú szabállyal fejezhetünk be. Ugyanez érvényes minden olyan szóra, amelyet az S kezdőszimbólumból vezetünk le.

Definíció szerint minden reguláris nyelvtan környezetfüggetlen, mert az $A \rightarrow xB$ szabályok esetén $xB \in (N \cup \Sigma)^*$ és az $A \rightarrow x$ szabályok esetén $x \in (N \cup \Sigma)^*$.

→ Példa: $G_3 \Rightarrow S \rightarrow aS|bS|\epsilon$

$$L(G_3) = \Sigma^* \text{ és } \Sigma = \{a, b\}$$

$$\text{abb levezetése: } S \Rightarrow aS \Rightarrow abS \Rightarrow abbs \Rightarrow abb$$

Derivációt csak az $S \rightarrow \epsilon$ szabállyal tudjuk befejezni.

1.6 Reguláris nyelvek

Egy L nyelvet reguláris nyelvnek hívunk, ha van olyan G reguláris nyelvtan, melyre $L = L(G)$. Az összes reguláris nyelvek halmazát REQ-gel jelöljük. Például az \emptyset , az $\{a^n b^m | n, m \geq 0\}$ nyelv, és minden véges nyelv reguláris. $\text{REQ} \subseteq \text{CF}$

2 Véges automata fogalma, felismert nyelv, kiterjesztések és ezek ekvivalenciája

2.1 Véges automaták fogalma

2.1.1 Determinisztikus automata

Az $M = (Q, \Sigma, \beta, q_0, F)$ rendszert determinisztikus automatának nevezzük, ahol:

- Q - Egy nem üres, véges halmaz, az állapotok halmaza
- Σ - Egy ábécé, az input ábécé
- $q_0 \in Q$ - Kezdőállapot
- $F \subseteq Q$ - A végállapotok halmaza
- $\beta : Q \times \Sigma \rightarrow Q$ - Egy leképezés, az átmenetfüggvény

Gráfként:

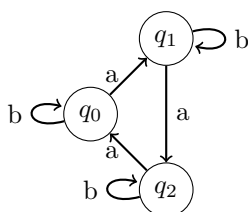


→ Példa: $M_3 = (Q, \Sigma, \beta, q_0, F)$

$Q = (q_0, q_1, q_2)$ $\beta(q_0, a) = q_1, \beta(q_0, b) = q_0$

$\Sigma = a, b$ $\beta(q_1, a) = q_2, \beta(q_1, b) = q_1$

$F = q_0$ $\beta(q_2, a) = q_0, \beta(q_2, b) = q_2$



M konfigurációinak halmaza: $C = Q \times \Sigma^*$

A $(q, a_1 \dots a_n)$ konfiguráció azt jelenti, hogy M a q állapotban van és az $a_1 \dots a_n$ szót kapja inputként.

Átmeneti reláció:

- $(q, w), (q', w') \in C$ esetén $(q, w) \vdash_M (q', w')$, ha $w = aw'$, valamely $a \in \Sigma$ -ra és $\beta(q, a) = q'$
- $(q, w) \vdash_M (q', w')$, egy lépés
- $(q, w) \vdash_M^n (q', w')$, $n \geq 0$ lépés
- $(q, w) \vdash_M^+ (q', w')$, legalább 1 lépés
- $(q, w) \vdash_M^* (q', w')$, valamennyi (esetleg 0) lépés

Az $M = (Q, \Sigma, \beta, q_0, F)$ automata által felismert nyelven az

$L(M) = \{w \in \Sigma^* \mid (q_0, w) \vdash_M^* (q, \epsilon) \text{ és } q \in F\}$

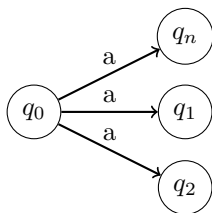
2.1.2 Nemdeterminisztikus automata

Az $M = (Q, \Sigma, \beta, q_0, F)$ rendszert Nemdeterminisztikus automatának nevezzük, ahol:

- Q - egy nem üres, véges halmaz, az állapotok halmaza
- Σ - egy ábécé, az input ábécé
- $q_0 \in Q$ - a kezdőállapot
- $F \subseteq Q$ - a végállapotok halmaza
- $\beta : Q \times \Sigma \rightarrow \mathcal{P}(Q)$ egy leképezés, az átmenetfüggvény

Nemdeterminisztikus: egy input szimbólum hatására egy állapotból több állapotba is átmehet.

→ Példa: $\beta(q, a) = \{q_1, \dots, q_n\}$



Az átmeneti reláció és a felismert nyelv nemdeterminisztikus automataakra:

$(q, w), (q', w') \in C$ esetén $(q, w) \vdash_M (q', w')$, ha $w = aw'$, valamely $a \in \Sigma$ -ra és $q' \in \beta(q, a)$

Az $M = (Q, \Sigma, \beta, q_0, F)$ automata által felismert nyelven az

$L(M) = \{w \in \Sigma^* \mid (q_0, w) \vdash_M^* (q, \epsilon) \text{ valamely } q \in F \text{-re}\}$ nyelvet értjük.

Az $M = (Q, \Sigma, \beta, q_0, F)$ nemdeterminisztikus automata teljesen definiált, ha minden $q \in Q$ és $a \in \Sigma$ esetén $\beta(q, a)$ legalább egy elemű

□ TÉTEL:

Testszöveges $M = (Q, \Sigma, \beta, q_0, F)$ nemdeterminisztikus automatához megadható olyan $M' = (Q', \Sigma, \beta', q_0, F)$ teljesen definiált automata, melyre $L(M) = L(M')$.

◦ BIZONYÍTÁS:

Ha M teljesen definiált, akkor legyen $M' = M$. Különben legyen $Q' = Q \cup \{q_c\}$, ahol $q_c \notin Q$, vagyis egy új állapot \Rightarrow csapda állapot.

Továbbá, minden $q \in Q$ és $a \in \Sigma$ esetén legyen

$$\beta(q, a) = \begin{cases} \beta(q, a), & \text{ha } \beta(q, a) \neq \emptyset \\ \{q_c\}, & \text{ha } \beta(q, a) = \emptyset \end{cases} \quad (1)$$

Végül, minden $a \in \Sigma$ -ra, legyen $\beta'(q_c, a) = \{q_c\}$

□ TÉTEL:

Egy nyelv akkor és csak akkor ismerhető fel nemdeterminisztikus automatával, ha felismerhető determinisztikus automatával.

◦ BIZONYÍTÁS:

- Ha egy nyelv felismerhető determinisztikus automatával, akkor felismerhető nemdeterminisztikus automatával is.
- Fordítva: Legyen $M = (Q, \Sigma, \beta, q_0, F)$ egy nemdeterminisztikus automata. Megadunk egy $M' = (Q', \Sigma, \beta', q'_0, F')$ determinisztikus automatát, amelyre $L(M') = L(M)$.
A konstrukció neve: hatványhalmaz konstrukció