Formális nyelvek szóbeli tételek kidolgozása

May 22, 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 $\bigcap \Sigma = \emptyset$
- 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 $\bigcup \Sigma$)*

$\rightarrow$ Példa:

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

1.2 Közvetlen levezetés (deriváció)

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

$\rightarrow$ Példa:

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

$bSa \Rightarrow_{G_1} baSba$ az $S \rightarrow aSb$ szabállyal
$baaSa \Rightarrow_{G_1} baaa$ az $S \rightarrow \epsilon$ 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, Σ , P, S) környezetfüggetlen nyelvtan által generált nyelv: L(G) = $\{ w \in \Sigma^* | S \Rightarrow_G^* w\}$
  $\rightarrow$ Az $w \in \Sigma^*$ feltétel miatt w-ben nincsenek nemterminálisok, tegát nem lehet belőle ”tovább lépni”.
  $\rightarrow$ pl. $L(G_1) = \{a^n b^n | n \geq 0\}$

*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, Σ, P, S) nyelvtan reguláris (vagy jobblineáris), ha P-ben minden szabály A → xB vagy A → x alakú. Reguláris nyelvtan esetén minden levezetés

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

alakú, ahol az alkalmazott szabályok

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

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

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

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

L(G_3) = Σ^* és Σ = {a, b}
abb levezetése: $S \Rightarrow aS \Rightarrow abS \Rightarrow abbs \Rightarrow abb$
Derivációt csak az $S \to \epsilon$ 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 ∅, az $\{\ a^n b^m | n, m \geq 0\}$ nyelv, és minden véges nyelv reguláris.
REQ ⊆ CF

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

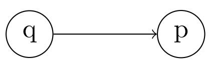
2.1 Véges automaták fogalma

2.1.1 Determinisztikus automata

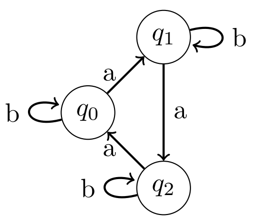
Az M = (Q, Σ, β, q_0, F) rendszert determinisztikus automatának nevezzük, ahol:

- 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
- $\beta : Q\ x\ \Sigma \to Q$ - Egy leképezés, az átmenetfüggvény

Gráfként:



$\to$ Példa: M$_3$ = (Q, Σ, β, 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 = QxΣ^*
A (q, $a_1 ... a_n$) konfiguráció azt jelenti, hogy M a q állapotbanvan és az $a_1 ... a_n$ szót kapja inputként.

Átmeneti reláció:

- (q, w) , (q', w') ∈ 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

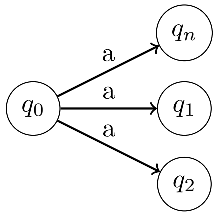
Az M = (Q, Σ, β, q_0, F) automata által felismert nyelven az
L(M) = $\{w \in \Sigma^* | (q_0, w) \vdash_M^* (q, \epsilon)$ és $q \in F\}$

2.1.2 Nemdeterminisztikus automata

Az M = (Q, Σ, β, 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
- $\beta : Qx\Sigma \to \mathcal{P}(Q)$ egy leképezés, az átmenetfüggvény

Nemdeterminisztikus: egy input szimbólum hatására egy állapotból több állapotba is átmehet.
$\to$ Példa: $\beta(q, a) = \{q_1, ..., q_n\}$



Az átmeneti reláció és a felismert nyelv nemdeterminisztikus autmoatákra:
(q, w), ($q^{'}, w^{'}$) $\in$ C esetén (q, w) $\vdash_M$ ($q^{'}, w^{'}$), ha w = a$w^{'}$, valamely $a \in \Sigma$-ra és $q^{'} \in \beta(q, a)$

Az M = (Q, Σ, β, q_0, F) automata által felismert nyelven az
L(M) = $\{w \in \Sigma^* | (q_0, w) \vdash_M^* (q, \epsilon)$ valamely $q \in F$ -re $\}$ nyelvet értjük.

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

Igen