Nyelvtani transzformációk

Formális nyelvek, 6. gyakorlat

Célja: A nyelvtani transzformációk bemutatása

Fogalmak: Megszorított típusok, normálformák, 0. típusú epszilon-mentesítés, 2. típusú epszilon-mentesítés, láncmentesítés, Chomsky-féle normálformává alakítás, reguláris műveletekre való zártságot bizonyító konstrukciók.

Feladatok jellege: Konkrét nyelvtanokból kiindulva a konstrukciók tényleges, mechanikus elvégzése.

2005/06 II. félév

Házi feladatok megoldása

1. feladat

Melyik nyelvet generálja a következő nyelvtan? $T = \{(,)\}$

a. $S \to (S) \mid SS \mid \varepsilon$ **b.** $S \to XS \mid \varepsilon$ *és* $X \to (S)$ **c.** $S \to (SS \mid)$

Megoldás:

a. és **b.**: **HE**, **c.**: **HE**). Például **a**:

Jelölje L a generált nyelvet.

"$L \subseteq$ **HE**":

Minden újonnan behozott jobbzárójel elé valahova kerül ugyanakkor egy új balzárójel is. Így az aktuális jelsorozat minden prefixében legalább annyi balzárójel van mint jobb. Formálisan:

Ha $\alpha \in \{S, (,)\}^*$, legyen $\phi(\alpha)$ a következő tulajdonság:

$\forall u \in \mathrm{Pre}(\alpha): \ \ell_{(}(u) \geq \ell_{)}(u)$ és $\ell_{(}(\alpha) = \ell_{)}(\alpha)$.

Házi feladatok megoldása

1. feladat

Melyik nyelvet generálja a következő nyelvtan? $T = \{(,)\}$

a. $S \to (S) \mid SS \mid \varepsilon$ **b.** $S \to XS \mid \varepsilon$ *és* $X \to (S)$ **c.** $S \to (SS \mid)$

Állítás: Ha $S \xrightarrow{*} \alpha$, akkor $\phi(\alpha)$.

A levezetés hosszára vonatkozó teljes indukcióval bizonyítunk. Ha $\alpha = S$, akkor igaz az állítás.

Tegyük fel, hogy $S \xrightarrow{n} \alpha = \alpha_1 S \alpha_2$.

Három eset lehetséges aszerint, hogy melyik szabályt alkalmazzuk. A kapott szavak legyenek rendre $\beta_1 = \alpha_1(S)\alpha_2$, $\beta_2 = \alpha_1 SS \alpha_2$, $\beta_3 = \alpha_1\alpha_2$.

Legyen $u \in \mathrm{Pre}(\beta_i)$, azaz $uv = \beta_i$ ($i = 1, 2, 3$).

Házi feladatok megoldása

1. feladat

Melyik nyelvet generálja a következő nyelvtan? $T = \{(,)\}$

a. $S \to (S) \mid SS \mid \varepsilon$ **b.** $S \to XS \mid \varepsilon$ *és* $X \to (S)$ **c.** $S \to (SS \mid)$

Ha $u \in \mathrm{Pre}(\alpha_1)$, akkor az indukció alapján $\ell_{(}(u) \geq \ell_{)}(u)$.

Ha $v \in \mathrm{Suf}(\alpha_2)$, akkor $\exists u', \ u'v = \alpha$. Az indukció alapján $\ell_{(}(u') \geq \ell_{)}(u')$ és akármelyik szabályt is alkalmaztuk, ugyanannyival ($i = 2, 3$ esetén 0-val, $i = 1$ esetén 1-gyel) nőtt a bal- és jobbzárójelek száma u-ban u'-höz képest. Tehát $\ell_{(}(u) \geq \ell_{)}(u)$.

Ha $u = \alpha_1\gamma_1$ és $v = \gamma_2\alpha_2$, (azaz γ_1 prefixe valamelyik levezetési szabály jobboldalának) akkor indukció alapján $\ell_{(}(\alpha_1) \geq \ell_{)}(\alpha_1)$ és könnyen ellenőrizhető, hogy $\ell_{(}(\gamma_1) \geq \ell_{)}(\gamma_1)$, tehát $\ell_{(}(u) = \ell_{(}(\alpha_1) + \ell_{(}(\gamma_1) \geq \ell_{)}(\alpha_1) + \ell_{)}(\gamma_1) = \ell_{)}(u)$.

$\ell_{(}(\beta_i) = \ell_{)}(\beta_i)$, hiszen minden szabály jobboldala ugyanannyi bal- és jobbzárójelet tartalmaz (0-t vagy 1-et), tehát $\phi(\beta_i)$, ($i = 1, 2, 3$).

Házi feladatok megoldása

1. feladat

Melyik nyelvet generálja a következő nyelvtan? $T = \{(,)\}$
a. $S \to (S)|SS|\varepsilon$ **b.** $S \to XS|\varepsilon$ *és* $X \to (S)$ **c.** $S \to (SS|)$

"$L \supseteq$ **HE**" :
A zárójelek számára vonatkozó indukcóval belátjuk, hogy minden helyes zárójelezés levezethető.
Az üres zárójelezés az $S \to \varepsilon$ szabállyal levezethető.
Tekintsünk egy w helyes zárójelezést. Ekkor vagy $w = (w_1)$, vagy $w = w_1 w_2$, ahol w_1, w_2 helyes zárójelezések. (Attól függően, hogy van-e w-nek valódi prefixe, mely ugyanannyi bal- és jobbzárójelet tartalmaz.)
Indukció alapján w_1 és w_2 levezethető.

Az első esetben az $S \to (S) \xrightarrow{*} (w_1)$, a másodikban az $S \to SS \xrightarrow{*} w_1 S \xrightarrow{*} w_1 w_2$ levezetés w-nek egy jó levezetését adja.

Házi feladatok megoldása

2. feladat

Adjunk az $L = \{v; v = uu\}$ ($T = \{a, b\}$) nyelvet generáló nyelvtant a "vv^{-1}" alakú szavak nyelvénél látott módszerre való visszavezetéssel!

Megoldás:

$S' \to \varepsilon \mid S$
$S \to tSX_t \mid tY_t \qquad \forall t \in T$
$Y_t X_{t'} \to Y_{t'} t \qquad \forall t, t' \in T$
$tX_{t'} \to X_{t'} t \qquad \forall t, t' \in T$
$Y_t \to t$

Házi feladatok megoldása

3. feladat

Adjunk nyelvtant! $T = \{a\}$
$L = \{a^{2^n}; n \geq 0\}$.

Megoldás:

$S \to LDaR$
$LD \to \varepsilon$
$R \to \varepsilon$
$Da \to aaD$
$DR \to ER$
$ER \to \varepsilon$
$L \to \varepsilon$
$aE \to Eaa$
$LE \to LD$

0. típusú nyelvtanok

ε-mentesítés

$G = \langle \{a, b\}, \{S, X, Y\}, \mathcal{P}, S \rangle$
$S \to aXSbY \mid ab \mid \varepsilon$
$Xb \to \varepsilon$
$Xa \to aaX$
Végezzük el a 0. típusú ε-mentesítést!

Megoldás:
$S' \to \varepsilon \mid S$
$S \to aXSbY \mid ab$

$aS \to a$	$Sa \to a$	$aXb \to a$	$Xba \to a$
$bS \to b$	$Sb \to b$	$bXb \to b$	$Xbb \to b$
$XS \to X$	$SX \to X$	$XXb \to X$	$XbX \to X$
$YS \to Y$	$SY \to Y$	$YXb \to Y$	$XbY \to Y$

$Xa \to aaX$

Kuroda normálforma (1. típusú nyelvtan)

A normálformára alakítás lépései

Alakítsuk át a négyzetszám hosszúságú szavakat generáló nyelvtan szabályait környezetfüggővé! Hozzuk a szabályokat Kuroda normálformára!

Példa: $XY \to YaX$ szabály:

1. lépés: új, lokális változók (álterminálisok) bevezetése:
$XY \to YQ_aX$ $\quad Q_a \to a$

2. lépés: új nyelvtani jelekkel egyesével átírjuk a kívánt sorozatra (már csak "$AB \to CD$" alakú rossz szabályok maradnak):
$XY \to YZ_1$ $\quad Z_1 \to Q_aX$

3.lépés: Az "$AB \to CD$" alakú szabályok eliminálása:
$XY \to XW$ $\quad XW \to YW$ $\quad YW \to YZ_1$

Chomsky normálforma (2. típus)

1. lépés: ε-mentesítés

$G = < \{a, b, c\}, \{S, A, B, C\}, \mathcal{P}, S >$
$S \to AB \mid \varepsilon$
$A \to aAa \mid C$
$B \to bBb \mid C$
$C \to Cccc \mid \varepsilon$

ε-mentesítés:

Konstruálunk egy $H \subseteq \{S, A, B, C\}$ segédhalmazt, melynek pontosan azok a nyelvtani jelek lesznek az elemei, melyekből levezethető ε. Ehhez segítségünkre lesznek a rekurzívan definiált H_i halmazok. A H_i halmaz a H_{i-1} halmaz bővítése azon nyelvtani jelekkel, amelyekből közvetlenül levezethető H_{i-1}-beli nyelvtani jel. A kiindulási halmaz H_1, azon nyelvtani jelek halmaza, melyekből közvetlenül levezethető ε.

Chomsky normálforma (2. típus)

1. lépés: ε-mentesítés

$H_1 = \{S, C\}$,
$H_2 = \{S, C\} \cup \{A, B\}$,
$H_3 = H_2 \Rightarrow H = \{S, A, B, C\}$.

Képezzük az összes olyan szabályt, mely az eredetiből a jobboldalakon néhány H-beli elhagyásával kapható, de marad legalább egy terminális vagy nem terminális jel.

$S \to AB \mid A \mid B$
$A \to aAa \mid aa \mid C$
$B \to bBb \mid bb \mid C$
$C \to Cccc \mid ccc$

Mivel $S \in H$ hozzá kell adni még a következőt:
$S' \to S \mid \varepsilon$

Chomsky normálforma (2. típus)

2. lépés: Álterminálisok bevezetése

Álterminálisok bevezetése:

Minden terminális jel helyett behozunk egy új nyelvtani jelet, a szabályokban a terminális jeleket ezekre cseréljük, és az álterminálisokat terminálsokra cserélő szabályokat hozzáadjuk.

$S' \to S \mid \varepsilon$
$S \to AB \mid A \mid B$
$A \to Q_aAQ_a \mid Q_aQ_a \mid C$
$B \to Q_bBQ_b \mid Q_bQ_b \mid C$
$C \to CQ_cQ_cQ_c \mid Q_cQ_cQ_c$
$Q_a \to a$
$Q_b \to b$
$Q_c \to c$

Chomsky normálforma (2. típus)

3. lépés: Láncmentesítés

Láncmentesítés:

Meghatározzuk minden nyelvtani jelhez azon nyelvtani jelek halmazát, melyek levezethetők belőle.

$H(S') = \{S', S, A, B, C\}, \quad H(S) = \{S, A, B, C\},$
$H(A) = \{A, C\}, \quad H(B) = \{B, C\}, \quad H(C) = \{C\}.$

Minden $Y \in H(X)$ nyelvtani jelhez vesszük azon szabályokat, amelyeknek baloldalán X, jobboldalán pedig egy Y-ra vonatkozó eredeti szabály jobboldala áll, kivéve ha ez a jobboldal egyetlen nyelvtani jel.

Chomsky normálforma (2. típus)

4. lépés: Hosszredukció

Tehát eddig kaptuk:

$S' \to \varepsilon \mid AB \mid Q_aAQ_a \mid Q_aQ_a \mid Q_bBQ_b \mid Q_bQ_b \mid CQ_cQ_cQ_c \mid Q_cQ_cQ_c$
$S \to AB \mid Q_aAQ_a \mid Q_aQ_a \mid Q_bBQ_b \mid Q_bQ_b \mid CQ_cQ_cQ_c \mid Q_cQ_cQ_c$
$A \to Q_aAQ_a \mid Q_aQ_a \mid CQ_cQ_cQ_c \mid Q_cQ_cQ_c$
$B \to Q_bBQ_b \mid Q_bQ_b \mid CQ_cQ_cQ_c \mid Q_cQ_cQ_c$
$C \to CQ_cQ_cQ_c \mid Q_cQ_cQ_c$
$Q_a \to a \quad Q_b \to b \quad Q_c \to c$

Hosszredukció:

Például: $C \to CQ_cQ_cQ_c$ szabály:

$C \to CZ_1 \quad Z_1 \to Q_cZ_2 \quad Z_2 \to Q_cQ_c$

3. típusú normálforma

Kiterjesztett 3. típusú nyelvtan normálformára hozása

$S \to \varepsilon \mid bA$
$A \to aaA \mid S \mid b$

Láncmentesítés és hosszredukció után:

$S \to \varepsilon \mid bA$
$A \to aZ_1 \mid \varepsilon \mid bA \mid b$
$Z_1 \to aA$

A nyelvtani jelből terminális alakú szabályok átalakítása:

$S \to \varepsilon \mid bA$
$A \to aZ_1 \mid \varepsilon \mid bA \mid bF_1$
$Z_1 \to aA$
$F_1 \to \varepsilon$