## Hogyan készítsünk Greibach normál alakú (GNF) nyelvtant környezetfüggetlenből?

Legelőször tennék a címhez még egy kiegészítést. Ugyanis jól fésült nyelvtant fogunk csak átalakítani. Ha a nyelvtanunk nem jól fésült, vagyis tartalmaz egyszeres szabályt, $\varepsilon$-szabályt, vagy felesleges szimbólumot, akkor a már jól ismert fésülős eljárásainkat kell először alkalmaznunk.

**Definíció**
*A nyelvtan* Greibach normál alakú (GNF)*, ha minden szabálya*

$$A \to \mathtt{a}W$$

*alakú, azaz a szabály bal oldalán egy darab nemterminális szimbólum szerepel, míg a jobb oldala egy darab terminális szimbólummal kezdődik és utána tetszőlegesen sok (akár 0) nemterminális szimbólum található. A nyelvtanban az $S \to \varepsilon$ szabály is benne lehet ($S$ a mondatszimbólum), amennyiben a generált nyelv tartalmazza az üres jelsorozatot, de ekkor $S$ nem szerepelhet szabály jobboldalán.*

**Tétel**
*Minden jól fésült környezetfüggetlen nyelvtan Greibach normál alakúvá alakítható.*

**Bizonyítás**
Bizonyításunk szokás szerint konstruktív lesz, azaz fogunk adni egy módszert, amely egy jól fésült környezetfüggetlen nyelvtanból indul ki, és egy olyan nyelvtant eredményez, melynek szabályai $A \to \mathtt{a}W$ alakúak.

1. Rögzítsünk le egy rendezést a nemterminális szimbólumokon! A választott rendezés elvileg tetszőleges lehet, de algoritmusunk sokat gyorsulhat, ha jól választunk. Legyen ez a rendezés:

   $$A_1 < A_2 < \ldots < A_n$$

   ahol $n$ a nyelvtanban szereplő nemterminálisok száma.

2. Ha van olyan szabályunk, melynek jobb oldalán nemcsak az első szimbólum terminális, akkor a hátrébb szereplő terminálisok helyére helyettesítsünk be (terminálisonként) egy-egy új nemterminálist, és adjunk nyelvtanunkhoz egy-egy új szabályt, melynek bal oldalán az új nemterminális szimbólum van, míg a jobboldalon az eredeti terminális.

3. Első lépésként azt a célt tűzzük ki magunk elé, hogy olyan nyelvtant készítsünk, amelyben csak

   *I.* $A_i \to \mathtt{a}W$
   *II.* $A_i \to A_jW \qquad i < j$

   alakú szabályok vannak.

   Menjünk végig $i$-vel 1-től $n$-ig a számokon növekvő sorrendben! ($n$ a nyelvtanban szereplő nemterminálisok száma.) Célunk, hogy az adott iterációs körben elérjük, hogy az adott $i$-re az $A_i$ baloldalú szabályok alakja a fenti formájú legyen.

(a) Menjünk végig $j$-vel 1-től $i$–1-ig a számokon növekvő sorrendben!
- Ha van a nyelvtanunkban

$$A_i \rightarrow A_j W \qquad i > j$$

alakú szabály, akkor helyettesítsük az összes ilyen szabályban az $A_j$ helyére az összes olyan szabály jobboldalát, ahol a bal oldalon az $A_j$ nemterminális szerepel.
Mivel az $A_j$ baloldalú szabályokon már végigmentünk korábban, könnyen látható, hogy az átalakítás után kapott szabályok vagy *I.* típusúak lesznek, vagy pedig legalább $j + 1$-es indexű nemterminálissal kezdődő jobboldaluk lesz.

(b) Most már minden szabályra igaz, hogy vagy *I.* típusú, vagy pedig

$$A_i \rightarrow A_j W \qquad i \leq j$$

alakú.
- Ha van a nyelvtanunkban $A_i \rightarrow A_i W$ alakú szabály, akkor hajtsuk végre a közvetlen balrekurziót kiküszöbölő algoritmusunkat. Ezután minden $A_i$ baloldalú szabály *I.* vagy *II.* alakú lesz.

Miután $i$-vel elértük az $n$ értéket, minden szabályunk alakja megfelel a kitűzött célunknak. Felhívnám a figyelmet arra is, hogy mivel $n$-nél nagyobb indexünk nincsen, ezért az $A_n$ baloldalú szabályok csak *I.* típusúak lehetnek.

4. Most pedig $i$-vel végig kell lépegetnünk $n$–1-től 1-ig.
   - Minden *II.* típusú szabályban a jobboldalon álló első, nemterminális szimbólum helyére helyettesítsük be az olyan szabályjobboldalakat, ahol a szabály bal oldalán a helyettesíteni kívánt nemterminális áll.

   Az adott lépés végén az olyan szabályok, melyeknek bal oldalán $i$-nél nem kisebb indexű nemterminális van, *I.* típusúak lesznek.

Amennyiben a 2. pontban újonnan bevezetett nemterminálisokat és a közvetlen balrekurzió kiküszöbölésekor keletkező új nemterminálisokat mindig a legelső helyre állítjuk be a nemterminálisok éppen aktuális rendezésében, és a 4. pontban leírt dolgokat rájuk is kiterjesztjük, akkor a végén minden szabály meg fog felelni a Greibach normál alak előírásainak.

## Az átalakítás szemléltetése egy példán

A feladat legyen a következő nyelvtan GNF-be transzformálása!

$$\begin{aligned} A &\rightarrow B\mathtt{c} \\ B &\rightarrow CA \mid \mathtt{b} \\ C &\rightarrow AB \mid \mathtt{a} \end{aligned}$$

1. Legyen a nemterminálisok sorrendje: $A < B < C$. Vagy másképpen leírva, legyen $A_1 = A$, $A_2 = B$ és $A_3 = C$.

2. Itt az $A \to B\mathtt{c}$ szabály helyett a továbbiakban az $A \to B\bar{C}$ illetve a $\bar{C} \to \mathtt{c}$ szabályokat használjuk. $\bar{C}$-t $\bar{C} < A_1 = A$ módon soroljuk be a nemterminálisok közé. A $\bar{C}$ baloldalú szabállyal a továbbiakban semmi dolgunk nincs.

3. Most az $i$ szerinti iterációt hajtjuk végre.

$i = 1$ (a) A $j$ szerinti iterációt úgy, ahogy van, ki is hagyhatjuk, mivel 1-nél kisebb sorszámú eredeti nemterminális nem lehet szabály jobb oldalának első pozíciójában.

(b) Jelen esetben nincs $A$ baloldalú, közvetlen balrekurziót okozó szabály, tehát itt sem kell tennünk semmit.

$i = 2$ Most sincs semmi teendőnk, mert $B \to CA$ *II.*, illetve $B \to \mathtt{b}$ *I.* típusú volt már kezdetben is.

$i = 3$ A $C \to \mathtt{a}$ szabállyal minden rendben is van (*I.* típusú), de a másik szabály megsérti a részcélként kitűzött *II.* második szabályalakot.

(a) Jöjjön akkor a $j$ szerinti iteráció!

$j = 1$ $C \to AB \quad \rightsquigarrow \quad C \to B\bar{C}B$, ahol $\rightsquigarrow$ jelzi azt, hogy a tőle balra levő szabály helyett a továbbiakban a tőle jobbra levő szabály(oka)t használjuk.

$j = 2$

$$C \to B\bar{C}B \quad \rightsquigarrow \quad \left\{ \begin{array}{l} C \to \mathtt{b}\bar{C}B \\ C \to CA\bar{C}B \end{array} \right.$$

(b) Ismétlésként nézzük meg, hogy milyen szabályaink vannak most!

$$\begin{array}{lcl} \bar{C} & \to & \mathtt{c} \\ A & \to & B\bar{C} \\ B & \to & CA \mid \mathtt{b} \\ C & \to & CA\bar{C}B \\ C & \to & \mathtt{a} \mid \mathtt{b}\bar{C}B \end{array}$$

A közvetlen balrekurziót szűrő algoritmust a negyedik és az ötödik sorral indítva kell végrehajtani. Tehát

$$\left. \begin{array}{l} C \to CA\bar{C}B \\ C \to \mathtt{a} \mid \mathtt{b}\bar{C}B \end{array} \right\} \quad \rightsquigarrow \quad \left\{ \begin{array}{l} C \to \mathtt{a} \mid \mathtt{b}\bar{C}B \mid \mathtt{a}\hat{C} \mid \mathtt{b}\bar{C}B\hat{C} \\ \hat{C} \to A\bar{C}B \mid A\bar{C}B\hat{C} \end{array} \right.$$

Most itt tartunk ($\hat{C}$ a legkisebb nemterminális):

$$\begin{array}{lcl} \hat{C} & \to & A\bar{C}B \mid A\bar{C}B\hat{C} \\ \bar{C} & \to & \mathtt{c} \\ A & \to & B\bar{C} \\ B & \to & CA \mid \mathtt{b} \\ C & \to & \mathtt{a} \mid \mathtt{b}\bar{C}B \mid \mathtt{a}\hat{C} \mid \mathtt{b}\bar{C}B\hat{C} \end{array}$$

4. Eztán jöhet a visszafele lépegetés.

$i = 2$ $B \to CA \quad \rightsquigarrow \quad B \to \mathtt{a}A \mid \mathtt{b}\bar{C}BA \mid \mathtt{a}\hat{C}A \mid \mathtt{b}\bar{C}B\hat{C}A$

$i = 1$ $A \to B\bar{C} \quad \rightsquigarrow \quad A \to \mathtt{a}A\bar{C} \mid \mathtt{b}\bar{C}BA\bar{C} \mid \mathtt{a}\hat{C}A\bar{C} \mid \mathtt{b}\bar{C}B\hat{C}A\bar{C} \mid \mathtt{b}\bar{C}$

$i < 1$ Az algoritmus kiterjesztése alapján:

$$\left.\begin{array}{lcl} \hat{C} & \to & A\bar{C}B \\ \hat{C} & \to & A\bar{C}B\hat{C} \end{array}\right\} \rightsquigarrow \left\{\begin{array}{lcl} \hat{C} & \to & \mathtt{a}A\bar{C}\bar{C}B \mid \mathtt{b}\bar{C}BA\bar{C}\bar{C}B \\ & & \mathtt{a}\hat{C}A\bar{C}\bar{C}B \mid \mathtt{b}\bar{C}B\hat{C}A\bar{C}\bar{C}B \\ & & \mathtt{b}\bar{C}\bar{C}B \\ \hat{C} & \to & \mathtt{a}A\bar{C}\bar{C}B\hat{C} \mid \mathtt{b}\bar{C}BA\bar{C}\bar{C}B\hat{C} \\ & & \mathtt{a}\hat{C}A\bar{C}\bar{C}B\hat{C} \mid \mathtt{b}\bar{C}B\hat{C}A\bar{C}\bar{C}B\hat{C} \\ & & \mathtt{b}\bar{C}\bar{C}B\hat{C} \end{array}\right.$$

Az eredményképpen kapott GNF alakú nyelvtant tehát a

$$\begin{array}{lcl} \hat{C} & \to & \mathtt{a}A\bar{C}\bar{C}B \mid \mathtt{b}\bar{C}BA\bar{C}\bar{C}B \\ & & \mathtt{a}\hat{C}A\bar{C}\bar{C}B \mid \mathtt{b}\bar{C}B\hat{C}A\bar{C}\bar{C}B \\ & & \mathtt{b}\bar{C}\bar{C}B \\ \hat{C} & \to & \mathtt{a}A\bar{C}\bar{C}B\hat{C} \mid \mathtt{b}\bar{C}BA\bar{C}\bar{C}B\hat{C} \\ & & \mathtt{a}\hat{C}A\bar{C}\bar{C}B\hat{C} \mid \mathtt{b}\bar{C}B\hat{C}A\bar{C}\bar{C}B\hat{C} \\ & & \mathtt{b}\bar{C}\bar{C}B\hat{C} \\ \bar{C} & \to & \mathtt{c} \\ A & \to & \mathtt{a}A\bar{C} \mid \mathtt{b}\bar{C}BA\bar{C} \mid \mathtt{a}\hat{C}A\bar{C} \mid \mathtt{b}\bar{C}B\hat{C}A\bar{C} \mid \mathtt{b}\bar{C} \\ B & \to & \mathtt{a}A \mid \mathtt{b}\bar{C}BA \mid \mathtt{a}\hat{C}A \mid \mathtt{b}\bar{C}B\hat{C}A \mid \mathtt{b} \\ C & \to & \mathtt{a} \mid \mathtt{b}\bar{C}B \mid \mathtt{a}\hat{C} \mid \mathtt{b}\bar{C}B\hat{C} \end{array}$$

szabályok alkotják.