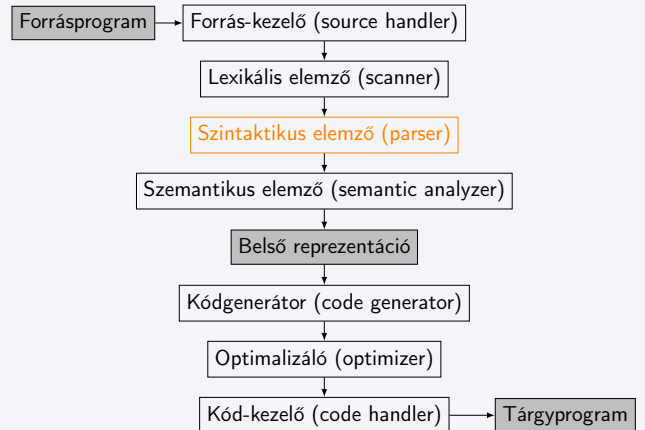


A szintaktikus elemzés

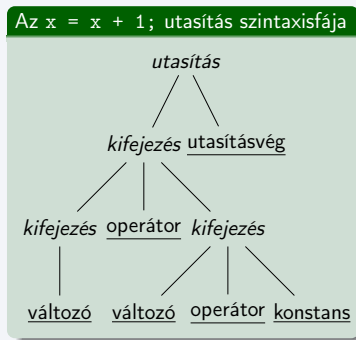
Fordítóprogramok előadás (A,C,T szakirány)

A szintaktikus elemzés helye



Szintaktikus elemzés

- bemenet: szimbólumsorozat
- kimenet: szintaxisfa, szintaktikus hibák
- feladata:
 - a program szerkezetének felismerése
 - a szerkezet ellenőrzése: megfelel-e a nyelv definíciójának?



Környezetfüggetlen nyelvtanok

- a szintaktikus elemzés: környezetfüggetlen nyelvtannal (Chomsky 2)
- $A \rightarrow \alpha$ alakú szabályok

„az A szimbólum a környezetétől függetlenül helyettesíthető α -ra”

Környezetfüggetlen nyelvtanok

- a szintaktikus elemzés: környezetfüggetlen nyelvtannal (Chomsky 2)
- $A \rightarrow \alpha$ alakú szabályok

„az A szimbólum a környezetétől függetlenül helyettesíthető α -ra”

Jelölés	Jelentés
a, b, c, \dots	terminális szimbólum (ezek a lexikális elemek, tokenek)
A, B, C, \dots	nemterminális szimbólum
X, Y, Z, \dots	terminális vagy nemterminális
x, y, z, \dots	terminális szimbólumok sorozata
$\alpha, \beta, \gamma, \dots$	terminális vagy nemterminális szimbólumok sorozata

Levezetések és kapcsolódó fogalmak

Jelölés	Jelentés
$A \rightarrow \alpha$	szabály
$A \Rightarrow \alpha$	egy lépéses levezetés (1 szabály alk.)
$A \Rightarrow^* \alpha$	nulla, egy vagy több lépéses levezetés
$A \Rightarrow^+ \alpha$	legalább egy lépésből álló levezetés

Levezetések és kapcsolódó fogalmak

Jelölés	Jelentés
$A \rightarrow \alpha$	szabály
$A \Rightarrow \alpha$	egy lépéses levezetés (1 szabály alk.)
$A \Rightarrow^* \alpha$	nulla, egy vagy több lépéses levezetés
$A \Rightarrow^+ \alpha$	legalább egy lépésből álló levezetés

- **mondat:** a startszimbólumból levezethető terminális sorozat ($S \Rightarrow^* x$)
- **mondatforma:** a startszimbólumból jeltől levezethető bármilyen sorozat ($S \Rightarrow^* \alpha$)
- **részmondat:** β az $\alpha_1\beta\alpha_2$ mondatforma részmondata, ha $S \Rightarrow^* \alpha_1A\alpha_2 \Rightarrow^+ \alpha_1\beta\alpha_2$.
- **egyszerű részmondat:** β az $\alpha_1\beta\alpha_2$ mondatforma egyszerű részmondata, ha $S \Rightarrow^* \alpha_1A\alpha_2 \Rightarrow \alpha_1\beta\alpha_2$.

Követelmények

- **ciklusmentesség:** nincs $A \Rightarrow^+ A$ levezetés
 - ellenpélda:

$$A \rightarrow B$$

$$B \rightarrow A$$

Követelmények

- **ciklusmentesség:** nincs $A \Rightarrow^+ A$ levezetés
 - ellenpélda:

$$A \rightarrow B$$

$$B \rightarrow A$$
- **redukáltság:** „nincsenek felesleges nemterminálisok”
 - minden nemterminális szimbólum előfordul valamelyik mondatformában
 - mindegyikből levezethető valamely terminális sorozat
 - ellenpélda: $A \rightarrow aA$, ha ez az egyetlen szabály A -ra

Követelmények

- **ciklusmentesség:** nincs $A \Rightarrow^+ A$ levezetés
 - ellenpélda:

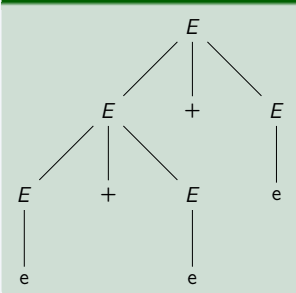
$$A \rightarrow B$$

$$B \rightarrow A$$
- **redukáltság:** „nincsenek felesleges nemterminálisok”
 - minden nemterminális szimbólum előfordul valamelyik mondatformában
 - mindegyikből levezethető valamely terminális sorozat
 - ellenpélda: $A \rightarrow aA$, ha ez az egyetlen szabály A -ra
- **egyértelműség:** minden mondatához pontosan egy szintaxisfa tartozik

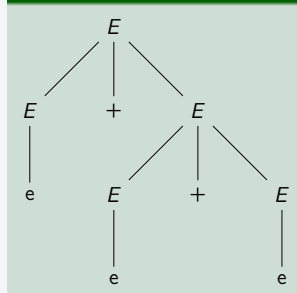
Példa nem egyértelmű nyelvtanra

$$E \rightarrow e \mid E + E$$

e+e+e (I.)



e+e+e (II.)



Legbal és legjobb levezetések

- **Legbal:** mindig a *legbaloldalibb* nemterminálíst helyettesítjük
- **Legjobb:** mindig a *legjobboldalibb* nemterminálíst helyettesítjük

Legbal és legjobb levezetések

- *Legbal*: mindig a *legbaloldalibb* nemterminálist helyettesítjük

Legbal levezetés

$$S \Rightarrow AB \Rightarrow aaB \Rightarrow aab$$

- *Legjobb*: mindig a *legjobboldalibb* nemterminálist helyettesítjük

Legjobb levezetés

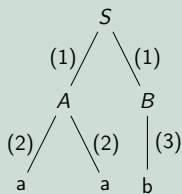
$$S \Rightarrow AB \Rightarrow Ab \Rightarrow aab$$

Elemzési irányok

- *Felülről lefelé*: A startszimbólumból indulva, felülről lefelé építjük a szintaxisfát. A mondatforma baloldalán megjelenő terminálisokat illesztjük az elemzendő szövegre.
- *Alulról felfelé*: Az elemzendő szöveg összetartozó részeit helyettesítjük nemterminális szimbólumokkal (redukció) és így alulról, a startszimbólum felé építjük a fát.

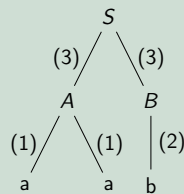
Elemzési irányok

Felülről lefelé



$$S \xrightarrow{(1)} AB \xrightarrow{(2)} aaB \xrightarrow{(3)} aab$$

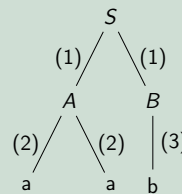
Alulról felfelé



$$aab \xleftarrow{(1)} Ab \xleftarrow{(2)} AB \xleftarrow{(3)} S$$

Elemzési irányok

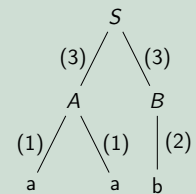
Felülről lefelé



$$S \xrightarrow{(1)} AB \xrightarrow{(2)} aaB \xrightarrow{(3)} aab$$

- Ez egy *legbal* levezetés!

Alulról felfelé



$$aab \xleftarrow{(1)} Ab \xleftarrow{(2)} AB \xleftarrow{(3)} S$$

- Ez egy *legjobb* levezetés inverze!

Felülről lefelé elemzések

- Probléma: „Melyik helyettesítési szabályt kell alkalmazni?”

Felülről lefelé elemzések

- Probléma: „Melyik helyettesítési szabályt kell alkalmazni?”
- Stratégiák:
 - visszalépéses keresés (backtrack): ha nem illeszkednek a szövegre a mondatforma baloldalán megjelenő terminálisok, lépünk vissza, és válasszunk másik szabályt
⇒ visszalépéses felülről lefelé elemzések (nem tananyag)
 - lassú
 - ha hibás a szöveg, az csak túl későn derül ki

Felülről lefelé elemzések

- Probléma: „Melyik helyettesítési szabályt kell alkalmazni?”
- Stratégiák:
 - visszalépéses keresés (backtrack): ha nem illeszkednek a szövegre a mondatforma baloldalán megjelenő terminálisok, lépünk vissza, és válasszunk másik szabályt
⇒visszalépéses felülről lefelé elemzések (nem tananyag)
 - lassú
 - ha hibás a szöveg, az csak túl későn derül ki
 - előreolvasás: olvassunk előre a szövegben valahány szimbólumot, és az alapján döntünk az alkalmazandó szabályról
⇒LL elemzések
 - csak szűk nyelvosztályra alkalmazható

Alulról felfelé elemzések: „Mit redukáljunk?”

Alulról felfelé elemzések: „Mit redukáljunk?”

- visszalépéses keresés (backtrack): ha nem sikerül eljutni a startszimbólumig, lépünk vissza, és válasszunk másik redukciót
⇒visszalépéses alulról felfelé elemzések (nem tananyag)
 - lassú
 - ha hibás a szöveg, az csak túl későn derül ki

Alulról felfelé elemzések: „Mit redukáljunk?”

- visszalépéses keresés (backtrack): ha nem sikerül eljutni a startszimbólumig, lépünk vissza, és válasszunk másik redukciót
⇒visszalépéses alulról felfelé elemzések (nem tananyag)
 - lassú
 - ha hibás a szöveg, az csak túl későn derül ki
- precedenciák használata: az egyes szimbólumok között adjunk meg precedenciarelációkat és ennek segítségével határozzuk meg a megfelelő redukciót
⇒precedencia elemzések (nem tananyag)
 - ma már kevésbé használt
 - operátorokkal felépített kifejezések esetén természetes a használata

Alulról felfelé elemzések: „Mit redukáljunk?”

- visszalépéses keresés (backtrack): ha nem sikerül eljutni a startszimbólumig, lépünk vissza, és válasszunk másik redukciót
⇒visszalépéses alulról felfelé elemzések (nem tananyag)
 - lassú
 - ha hibás a szöveg, az csak túl későn derül ki
- precedenciák használata: az egyes szimbólumok között adjunk meg precedenciarelációkat és ennek segítségével határozzuk meg a megfelelő redukciót
⇒precedencia elemzések (nem tananyag)
 - ma már kevésbé használt
 - operátorokkal felépített kifejezések esetén természetes a használata
- előreolvasás: olvassunk előre a szövegben valahány szimbólumot, és az alapján döntünk a redukcióról
⇒LR elemzések
 - minden programozási nyelvhez lehet (LR) elemzőt készíteni
 - majdnem mindegyikhez lehet gyors (LALR) elemzőt készíteni