Szintaktikai Elemzések:
általános felülről lefelé és alulról felfelé haladó elemzés

Az elemzés alapfeladata:

Adjunk olyan algoritmust, amely tetszőleges $G=(V_N, V_T, S, H)$ környezetfüggetlen nyelvtan és $w \in \Sigma^*$ szó esetén eldönti, hogy $w \in L(G)$ teljesül-e!

A felülről-lefelé haladó elemzések (top-down algoritmusok):

Az S kezdőszimbólumból kiindulva megpróbálunk felépíteni egy olyan derivációs fát, amelynek a határa w.

Az alulról-felfelé haladó elemzések (bottom-up algoritmusok):

A w-ből kiindulva megpróbálunk felépíteni egy olyan derivációs fát, amelynek a gyökere S és a határa w.

Felülről-lefelé haladó elemzések

Definíció: Alternatívák

Egy adott $A \in V_N$ nemterminális lehetséges behelyettesítési szabályainak a jobbolalai.

$A \rightarrow \gamma_1 \mid \gamma_2 \mid \ldots \mid \gamma_k$

Definíció: Kiterjesztés

Egy nemterminálisnak valamely alternatívájával való helyettesítése a derivációs fában.

Definíció: Illesztés

Annak ellenőrzése, hogy a kiterjesztésnél alkalmazott alternatívában szereplő terminálisok illeszkednek-e az elemzendő szó megfelelő részéhez.

Definíció: Felülről-lefelé haladó elemzés

Minden nemterminálisra lerögzítjük az alternatíváinak egy sorrendjét. Egy nemterminális kiterjesztése esetén az alternatívákat ebben a lerögzített sorrendben vizsgáljuk meg, hogy alkalmasak-e a kiterjesztésre. Ha nem találunk megfelelő alternatívát akkor egy backtrack-et (egy szinttel feljebb történő visszalépést) hajtunk végre.

Példa. Legyen $G=(V_N, V_T, S, H)$, ahol $H=\{ S \to T + S , S \to T, T \to a , T \to b\}$.

Feladat: $b+a \in L(G)$?

Alternatívák:

$S \to T + S \mid T$

$T \to a \mid b$

azaz S alternatívái : $S0 = T+S$, $S1 = T$, T alternatívái : $T0 = a$, $T1=b$.

- Vizsgáljuk meg az előző algoritmus segítségével, hogy b+a benne van-e a nyelvben!

Definíció: Alternatívák

Egy adott $A \in V_N$ nemterminális lehetséges behelyettesítési szabályainak a jobbolalai.

$A \to \gamma_1 \mid \gamma_2 \mid \dots \mid \gamma_k$

Definíció: Kiterjesztés

Egy nemterminálisnak valamely alternatívájával való helyettesítése a derivációs fában.

Definíció: Illesztés

Annak ellenőrzése, hogy a kiterjesztésnél alkalmazott alternatívában szereplő terminálisok illeszkednek-e az elemzendő szó megfelelő részéhez.

Definíció: Felülről-lefelé haladó elemzés

Minden nemterminálisra lerögzítjük az alternatíváinak egy sorrendjét. Egy nemterminális kiterjesztése esetén az alternatívákat ebben a lerögzített sorrendben vizsgáljuk meg, hogy alkalmasak-e a kiterjesztésre. Ha nem találunk megfelelő alternatívát akkor egy backtrack-et (egy szinttel feljebb történő visszalépést) hajtunk végre.

Algoritmus inputja:

Egy nem balrekurzív $G=(V_N , V_T , S, H)$ környezetfüggetlen nyelvtan és egy $w=a_0a_1a_2\ldots a_{n-1}$, $n \geq 1$ input szó.

A w szót n+1. szimbólumként egy # jel zárja le. A # nem tartozik sem V_N -hez, sem V_T-hez.

Algoritmus outputja:

Igen jelzés, és a w szónak egy baloldali levezetése, ha $w \in L(G)$.

Nem jelzés egyébként.

Módszer:

1. Minden $A \in V_N$ esetén rögzítsük le az A alternatíváit $A \rightarrow \gamma_0 \mid \gamma_1 \mid \ldots \mid \gamma_k$ alakban. Az A i-dik alternatíváját A_i jelöli. (Implementáláskor az (A, i) párt alkalmazzuk A_i jelölésére.)
2. Az elemzés (s, i, α, β) alakú konfigurációk sorozata.
3. A konfigurációk halmazán megadunk egy $\vdash$ átmeneti relációt. A rákövetkező konfiguráció meghatározása az alábbiakban megadott felsorolásból történik.
4. A kezdő konfiguráció $(q,0, \lambda, S)$.
   A befejező konfiguráció: (t, n,α, λ)
   $w \in L(G)$ akkor és csak akkor, ha $(q,0, \lambda, S) \vdash^* (t, n, \alpha, \lambda)$

a konfiguráció

(s, i, α, β) **értelmezése:**

s az elemzés állapota.

q - normál

t -elfogadó

b - backtrack

i pointer az input szóban ($0 \leq i \leq n$)

α jobbvégtetejű verem, az elemzés története backtrack-hez és a baloldali levezetéshez. (Passzív verem)

β balvégtetejű verem, a még levezetendő baloldali mondatforma. (Aktív verem)

átmeneti reláció

1. Kiterjesztés:

$(q, i, \alpha, A\beta) \vdash (q, i, \alpha A_0, \gamma_0\beta)$: az aktív szimbólum (az A) egy nemterminális és γ_0 az első alternatívája

2. Input illesztés sikeres: $a=a_i$ mellett $(q, i, \alpha, a\beta) \vdash (q, i+1, \alpha a, \beta)$: az aktív szimbólum egy olyan terminális, mely pont az i-edik betű

3. Sikeres elemzés

$(q, n, \alpha, \lambda) \vdash (t, n, \alpha, \lambda)$ elértük a befejező konfigurációt

4. Input illesztés sikertelen: $a \neq a_i$: az aktív szimbólum olyan terminális, mely nem illeszkedik az inputra : $(q, i, \alpha, a\beta) \vdash (b, i, \alpha, a\beta)$

5. Backtrack az inputban: b állapotban a passzív verem tetején terminális van.

$(b, i, \alpha a, \beta) \vdash (b, i\text{-}1, \alpha, a\beta)$

6. Backtrack a kiterjesztésben $(b, i, \alpha A_j, \gamma_j\beta)$ esetén a $\vdash$ jelet követi

I. A-nak van j+1. alternatívája

$(b, i, \alpha A_j, \gamma_j\beta) \vdash (q, i, \alpha A_{j+1}, \gamma_{j+1}\beta)$ vesszük a következő alternatívát

II. i=0, A=S, és S-nek csak j alternatívája van: nincs átmenet semelyik konfigurációba, az elemzett sztring nem eleme a nyelvnek

III Egyébként $(b, i, \alpha A_j, \gamma_j\beta) \vdash (b, i, \alpha, A\beta)$ nincs több alternatívája A-nak, visszatérünk az előző szintre.

Példa. Legyen $G=(V_N, V_T, S, H)$, ahol $H=\{ S \to T + S , S \to T, T \to a , T \to b\}$.

Feladat: $b+a \in L(G)$?

Alternatívák:

$S \to T + S \mid T$

$T \to a \mid b$

azaz S alternatívái : $S0 = T+S$, $S1 = T$, T alternatívái : $T0 = a$, $T1=b$.

Levezetés:

$(q, 0, \lambda, S) \vdash (q, 0, S0, T + S)$ (S kiterjesztése)

- Fejezzük be a konfigurációsorozat felírását

Példa. Legyen G=(V_N , V_T , S, H), ahol H={ S → T + S , S → T, T → a , T → b}.

Feladat: b+a ϵ L(G)?

Alternatívák:

S → T + S | T

T → a | b

azaz S alternatívái : S0 = T+S, S1 = T , T alternatívái : T0 = a, T1=b.

Levezetés:

(q, 0, λ, S) ⊦ (q, 1,S0, T + S) (S kiterjesztése)

⊦ (q, 0,S0T0, a + S) (T kiterjesztése)

⊦ (b, 0,S0T0, a + S) (sikertelen input illesztés)

⊦ (q, 0,S0T1, b + S) (backtrack a kiterjesztésben I.: T következő alternatíváját vesszük)

⊦ (q, 1,S0T1b,+S) (sikeres input illeszkedés: az első betű b)

⊦ (q, 2,S0T1b+,S) (sikeres input illeszkedés: a 2. betű +)

⊦ (q, 2,S0T1b + S0, T + S) (S kiterjesztése)

⊦ (q, 2,S0T0b + S0T0, a + S) (T kiterjesztése)

⊦ (q, 3,S0T1b + S0T0a,+S) (sikeres input illeszkedés: a 3. betű a)

⊦ (b, 3,S0T1b + S0T0a,+S) (sikertelen input illesztés)

⊦ (b, 2,S0T1b + S0T0, a + S) (backtrack az inputban)

⊦ (q, 2,S0T1b + S0T1, b + S) (backtrack a kiterjesztésben I.: T következő alternatíváját vesszük)

⊦ (b, 2,S0T1b + S0, T + S) (backtrack a kiterjesztésben III.: visszatérünk az előző szintre)

⊦ (q, 2,S0T1b + S1, T) (backtrack a kiterjesztésben I.: S következő alternatíváját vesszük)

⊦ (q, 2,S0T1b + S1T0, a) (T kiterjesztése)

⊦ (q, 3,S0T1b + S0T0a,λ) (sikeres input illesztés: a 3. betű a)

⊦ (t, 3,S0T1b + S1T10,λ). (sikeres elemzés, elértünk egy végkonfigurációt)

Következésképpen b + a ϵ L(G).
Baloldali levezetés (mindig balról az első nemterminálist helyettesítjük):
A baloldali levezetésben egymásután alkalmazandó alternatívák
(az αverem S0T1b + S1T0a tartalma alapján a szereplő terminálisok elhagyásával):
S0T1S1T0

Baloldali levezetés:
(S0) S → T + S alkalmazásával: S => T+S, (T1) T → b alkalmazásával T+S => b+S,
(S1) S → T alkalmazásával : b+S => b+T, (T0) T → a alkalmazásával b+T => b+a,

azaz a következő baloldali levezetést kapjuk:

S => T+S => b+S => b+T => b+a.