

Formális és természetes nyelvek nyelvtanai

Csima Judit

BME, VIK,
Számítástudományi és Információelméleti Tanszék

2015. szeptember 24.

- adott programozási nyelven írt szintaktikailag helyes kódok halmaza
= nyelv
- (potenciálisan) végtelen halmaz

- adott programozási nyelven írt szintaktikailag helyes kódok halmaza = nyelv
- (potenciálisan) végtelen halmaz
- fordító célja: egy adott kódról eldönteni, hogy helyes-e (szintaktikailag), azaz bele tartozik-e a nyelvbe

- adott programozási nyelven írt szintaktikailag helyes kódok halmaza = nyelv
- (potenciálisan) végtelen halmaz
- fordító célja: egy adott kódról eldönteni, hogy helyes-e (szintaktikailag), azaz bele tartozik-e a nyelvbe
- ehhez van egy véges definíció a programozási nyelvhez: nyelvtan
- ez egy generáló eszköz, az számít helyes kódnak, amit ezzel a nyelvtannal lehet generálni

single_input: NEWLINE | simple_stmt | compound_stmt NEWLINE

file_input: (NEWLINE | stmt)* ENDMARKER

...

stmt: simple_stmt | compound_stmt

...

compound_stmt: if_stmt | while_stmt | for_stmt | try_stmt | with_stmt |
funcdef | classdef | decorated

...

if_stmt: 'if' test ':' suite ('elif' test ':' suite)* ['else' ':' suite]

while_stmt: 'while' test ':' suite ['else' ':' suite]

for_stmt: 'for' exprlist 'in' testlist ':' suite ['else' ':' suite]

- van (két) kezdőszimbólum, ezek egyikéből indul minden helyes kód generálása

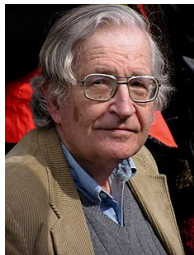
- van (két) kezdőszimbólum, ezek egyikéből indul minden helyes kód generálása
- a generálás azt jelenti, hogy ami a szabály bal oldalán áll (a : előtt) azt helyettesítjük a szabály jobb oldalával

- van (két) kezdőszimbólum, ezek egyikéből indul minden helyes kód generálása
- a generálás azt jelenti, hogy ami a szabály bal oldalán áll (a : előtt) azt helyettesítjük a szabály jobb oldalával
- addig csináljuk, amíg olyan szimbólumok lesznek csak, amik nem írhatók tovább:
atom: ... | NAME | NUMBER | STRING+

- van (két) kezdőszimbólum, ezek egyikéből indul minden helyes kód generálása
- a generálás azt jelenti, hogy ami a szabály bal oldalán áll (a : előtt) azt helyettesítjük a szabály jobb oldalával
- addig csináljuk, amíg olyan szimbólumok lesznek csak, amik nem írhatók tovább:
atom: ... | NAME | NUMBER | STRING+
- ha egy kód helyes, akkor a fordító/interpreter nem csak ezt veszi észre, hanem azt a struktúrát is felismeri, ahogy a kód előállt: levezetési fa

Noam Chomsky





- 1928-ban született
- nyelvész, filozófus, ...
- a modern nyelvészet atyja



- 1928-ban született
- nyelvész, filozófus, ...
- a modern nyelvészet atyja
- 1957: Syntactic Structures
- a transformational grammar ötlete

- nyelv megtanulása = szabályok megtanulása

Transformational grammar

- nyelv megtanulása = szabályok megtanulása
- szabályok = mint amilyenek a Python-nál voltak
- beszéd = S (sentence, mondat) szimbólumból szabályokkal levezetünk egy struktúrát
- ennek a struktúrának a felszíni megjelenése a mondat, ami elhangzik

Transformational grammar

- nyelv megtanulása = szabályok megtanulása
- szabályok = mint amilyenek a Python-nál voltak
- beszéd = S (sentence, mondat) szimbólumból szabályokkal levezetünk egy struktúrát
- ennek a struktúrának a felszíni megjelenése a mondat, ami elhangzik
- szövegértés = parser fut a fejünkben a megtanult nyelvtan és a hallott mondat alapján:

Transformational grammar

- nyelv megtanulása = szabályok megtanulása
- szabályok = mint amilyenek a Python-nál voltak
- beszéd = S (sentence, mondat) szimbólumból szabályokkal levezetünk egy struktúrát
- ennek a struktúrának a felszíni megjelenése a mondat, ami elhangzik
- szövegértés = parser fut a fejünkben a megtanult nyelvtan és a hallott mondat alapján: mi a mondat struktúrája, ez alapján mi a jelentése

Transformational grammar

- nyelv megtanulása = szabályok megtanulása
- szabályok = mint amilyenek a Python-nál voltak
- beszéd = S (sentence, mondat) szimbólumból szabályokkal levezetünk egy struktúrát
- ennek a struktúrának a felszíni megjelenése a mondat, ami elhangzik
- szövegértés = parser fut a fejünkben a megtanult nyelvtan és a hallott mondat alapján: mi a mondat struktúrája, ez alapján mi a jelentése
- az anyanyelvi beszélő kompetenciája: új, sose hallott mondatokat tud létrehozni és megérteni

Transformational grammar

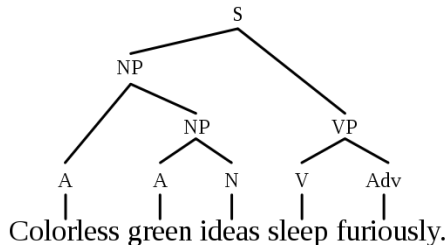
- nyelv megtanulása = szabályok megtanulása
- szabályok = mint amilyenek a Python-nál voltak
- beszéd = S (sentence, mondat) szimbólumból szabályokkal levezetünk egy struktúrát
- ennek a struktúrának a felszíni megjelenése a mondat, ami elhangzik
- szövegértés = parser fut a fejünkben a megtanult nyelvtan és a hallott mondat alapján: mi a mondat struktúrája, ez alapján mi a jelentése
- az anyanyelvi beszélő kompetenciája: új, sose hallott mondatokat tud létrehozni és megérteni
- "Colorless green ideas sleep furiously."

Az angol nyelv szabályai, részletek

$S \rightarrow NP VP$

$NP \rightarrow (DET) (ADJ)^* N$

$VP \rightarrow V (ADV)^*$



Universal grammar

- a különböző nyelvek nyelvtanai hasonlóak
- mintha lenne egy véges készlet és ennek valami variációja minden nyelv nyelvtana

- a különböző nyelvek nyelvtanai hasonlóak
- mintha lenne egy véges készlet és ennek valami variációja minden nyelv nyelvtana
- minden ember úgy születik, hogy a fejében van ez az univerzális nyelvtan
- nyelvtanulás: kigyomlálni a felesleges szabályokat