Sifat Context Free Language

Kuliah Teori Bahasa dan Automata
Program Studi Ilmu Komputer
Fasilkom UI

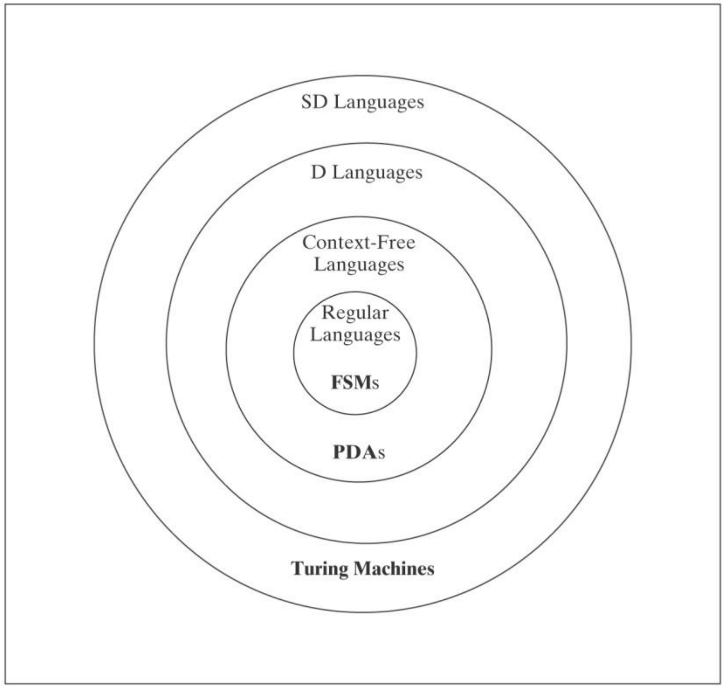
Prepared by:
Suryana Setiawan

CFL atau Bukan?

a^*b^* → regular.

$A^nB^n = \{a^nb^n : n \geq 0\}$ → CFL tapi tidak reguler

$A^nB^nC^n = \{a^nb^nc^n : n \geq 0\}$ → bukan CFL



Teorema-teorema

- Kelas CFL meliputi juga bahasa-bahasa reguler
  - Buktikan untuk suatu FSM dapat dibentuk PDA-nya
  - Tidak berlaku sebaliknya, karena ada *language* yang reguler tapi tidak CFL
- Banyaknya CFL adalah countably infinite
- Banyaknya bahasa yang dapat dibentuk oleh suatu alfabet, uncountably infinite. Berarti lebih banyak lagi bahasa yang tidak termasuk dalam CFL.
- Setiap bahasa CFL memenuhi teorema pumping

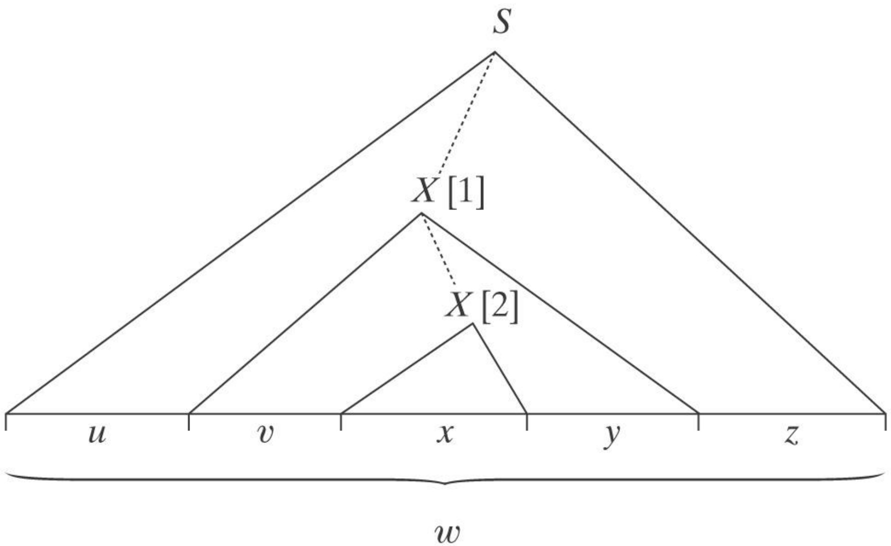
Pembuktian suatu Bahasa adalah CFL

- Menunjukkan suatu CFG untuk bahasa tsb
- Menunjukkan suatu PDA untuk bahasa tsb
- Pertanyaan: jika untuk sembarang bahasa L, apakah bisa dibuktikan ada tidaknya CFG/PDA?

Review Parse Tree

- Suatu parse tree dalam derivasi menurut grammar $G = (V, \Sigma, R, S)$, adalah *rooted, ordered tree* yang mana:
  - Setiap *leaf node* berlabelkan suatu elemen $(\Sigma \cup \{\varepsilon\})$,
  - *Root node* berlabel *S*,
  - Setiap node yang lain berlabel elemen-elemen $(V - \Sigma)$, dan
  - Jika *m* adalah *nonleaf node* berlabel *X*, dan anak-anak *m* berlabel $x_1, x_2, \ldots, x_n$, maka R berisi rule $X \rightarrow x_1, x_2, \ldots, x_n$.

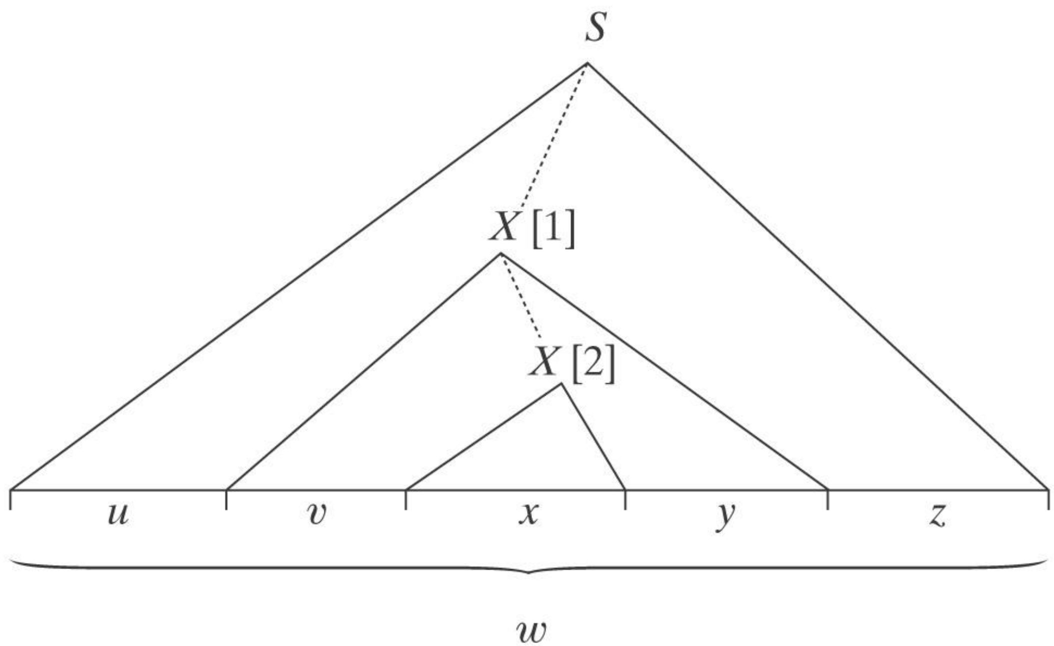
Teorema Pumping Untuk CFL



$$S \Rightarrow^* uXz \Rightarrow^* \mathrm{uvXyz} \Rightarrow^* uvxyz$$

[1] → rule 1
[2] → rule 2 diterapkan pada rule [1] sehingga menghasilkan *uvxyz* yang merupakan elemen dari L(G)

Teorema Pumping Untuk CFL



Beberapa derivasi lain dari G:

$$S \Rightarrow^* uXz \Rightarrow^* uvXyz \Rightarrow^* uvvXyyz \Rightarrow^* uvvxyyz$$

Derivasi tersebut menghasilkan string:

$uv^2xy^2z, uv^3xy^3z, \ldots$

Yang seluruhnya juga merupakan anggota dari $L(G)$.

Teorema Pumping Untuk CFL

- Jika L adalah CFL, maka:
  - $\exists k \geq 1$ (

    $\forall\, w \in L$, dimana $|w| \geq k$ (

    $\exists u, v, x, y, z$ (

    $w = uvxyz$,

    $|vxy| \leq k$,

    $vy \neq \varepsilon$, dan,

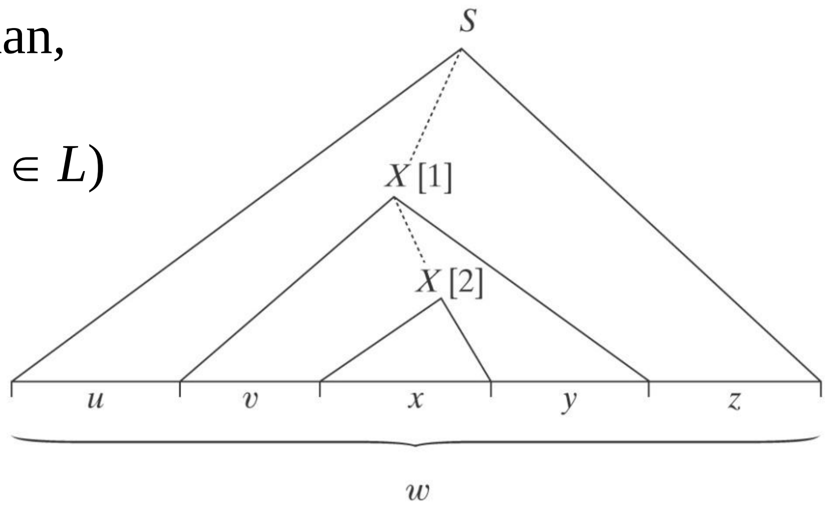
    $\forall q \geq 0$

    $(uv^q xy^q z \in L)$

)

)

)



Perbedaan dengan PL u/ Bahasa Reguler

- Adanya dua region *v* dan *y* yang dipompa bersamaan (sementara untuk Bhs Reguler hanya *y*)
- Kita tidak tahu mana yang menjadi *v* dan *y*, yang kita ketahui posisinya berdekatan akibat batasan $|vxy| \leq k$. (Untuk Bhs Reguler, kita tidak tahu juga mana yang *y*)
- Salah satu dari *v* dan *y* boleh kosong, tapi tidak keduanya. (Untuk Bhs Reguler, *x* minimal satu simbol)

Contoh 1 (bahasa $A^nB^nC^n$)

- $L = \{a^n b^n c^n : n \geq 0\}$
- Diberikan suatu harga k.
- Jika $w = a^k b^k c^k$ (misalnya, jika $k = 3$, w = aaabbbccc).
- Maka bisa ditunjukkan, tidak ada suatu cara pemecahan w ke dalam u,v,x,y, dan z yang bisa memompa dengan setiap harga q selalu $uv^q xy^q z \in L$.
- Misalnya jika $k = 3$,
  - w = aaabbbccc , sehingga $|vy|$ berharga 1 , 2 atu 3.
  - Jika $|vy| = 1$, dimana salah satu v atau y adalah ε, maka salah satu sequence simbol memiliki panjang lebih panjang/pendek dari dua yang lain untuk $q > 1$.
  - Demikian halnya untuk vy yang lain (bisa dicoba!).

Panduan Praktis

- Pilih w yang menangkap inti dari *L* yang bersifat context free.
  - Yang menyebabkan setiap kemungkinan pemecahan *w* menjadi *u*, *v*, *x*, *y*, dan *z* tidak memenuhi teorema pumping.
  - Se-homogen mungkin sehingga banyaknya kemungkinan pemecahan menjadi lebih sedikit (dari panduan untuk bhs reguler)
- Mencari harga q sehingga w dengan pemecahan yang diberikan (given) tidak dapat dipompa.
- Bisa menerapkan sifat closure dan pembuktian dilakukan pada bahasa hasil operasi closurenya
  - Sifat *closure* akan dibahas kemudian

Contoh 2

- $L = \{a^m : m = n^2$, dengan $n \geq 0\}$
- Diberikan suatu harga k.
- Jika $|w| = k^4$
- Selanjutnya, jelas $vy = a^p$ dengan $1 \leq p \leq k$ pada semua kemungkinan pemecahan u,v,x,y, dan z dari w.
- Dengan harga $q = 2$ maka $w' = uv^qxy^qz \notin L$ karena sbb.
  - Sementara $|w| = k^4 = (k^2)^2$, string berikutnya w'' (proper ordering) memiliki panjang $(k^2+1)^2 = k^4+2k^2 + 1$.
  - Karena $uv = a^p$, maka $|w'| = k^4 + p$ dan $p \leq k < 2k^2 + 1$, sehingga w' hanyalah string dengan panjang antara $|w|$ dan $|w''|$, dan $w' \notin L$

Contoh 3

- Untuk memeriksa apakah $L = \{a^n b^m a^n : n, m \geq 0 \text{ dan } n \geq m\}$ context free dengan suatu k, kita gunakan $w = a^k b^k a^k$ dan kita sebut a^k pertama sbg region 1, b^k sbg region 2 dan a^k terakhir sbg region 3.
- Jika salah satu dari v atau y melintasi region, dengan $q = 2$ menghasilkan string di luar L.
- Untuk kemungkinan lainnya ($(i, j) = v$ di region i dan y di region j):
  - (1,1): dengan $q = 2$, menghasilkan deretan a pertama lebih panjang dari deretan a kedua.
  - (2,2): dengan $q = 2$, deretan b lebih panjang dari satu deretan a.
  - (3,3): dengan $q = 0$, argumen sama dengan (1,1)
  - (1,2): dengan $q = 2$ maka argumen sama dengan (1,1) atau (2,2)
  - (2,3): dengan $q = 2$, menghasilkan deretan b atau deretan a yang kedua lebih panjang
  - (1,3): tidak mungkin karena $|vxy| \leq k$.

Sifat-sifat Closure pada CFL

- CFL Closure dalam operasi union, konkatenasi, Kleene Star, Reverse, dan letter substitution
  - Sifat closure **op** adalah “Jika L_1 **op** L_2 adalah CFL jika L_1 dan L_2 keduanya CFL”

Sifat-sifat Nonclosure pada CFL

- CFL tidak closure dalam operasi irisan, komplemen, dan different
  - Catatan: perbedaan penting dibanding bahasa reguler (semua operasi di atas closure untuk bahasa reguler)
- CFL closure dalam <u>operasi irisan/different dengan bahasa reguler</u>.

Penggunaan Teorema Pumping dalam konjungsi dengan Sifat Closure

- Pembuktian teorema pumping terhadap L_1 dapat dilakukan pembuktian pada L_2 jika
- L_1 **op** $L_3 = L_2$
- Jika diketahui L_3 adalah CFL dan **op** adalah operasi yang bersifat closure dalam CFL
- Jika L_1 **op** L_3 adalah CFL jika L_1 CFL, tetapi jika terbukti L_2 bukan CFL maka L_1 bukan CFL.

Contoh 4

- Bahasa $L = \{w \in \{a,b,c\}^*: \#_a(w) = \#_a(w) = \#_c(w)\}$ dapat lebih mudah dibuktikan dengan memeriksa $L' = L \cap a^*b^*c^* = A^nB^nC^n$ apakah juga context free. Karena sudah diperiksa sebelumnya bahwa L' bukan context free, maka L juga bukan context free.