Sifat Context Free Language (Bagian 1)

Kuliah Teori Bahasa dan Automata
Program Studi Ilmu Komputer
Fasilkom UI

Prepared by:
Suryana Setiawan

Sifat-sifat Yang Sudah Dipelajari

- Sifat-sifat CFL yang sudah dipelajari:
  - CFL adalah kelas dari Bahasa-bahasa yang dapat diderivasi oleh CFG.
  - Ekivalensi CFG-PDA.
- Suatu Bahasa CFL adalah CFL dengan:
  - Menunjukkan suatu CFG untuk bahasa tsb.
  - Menunjukkan suatu PDA untuk bahasa tsb.
- Sembarang bahasa L, apakah itu CFL?
  - Jika kita tidak mengetahui adanya CFG/PDA untuk Bahasa L,
  - bukan berarti L bukan CFL!

CFL dan Non-CFL

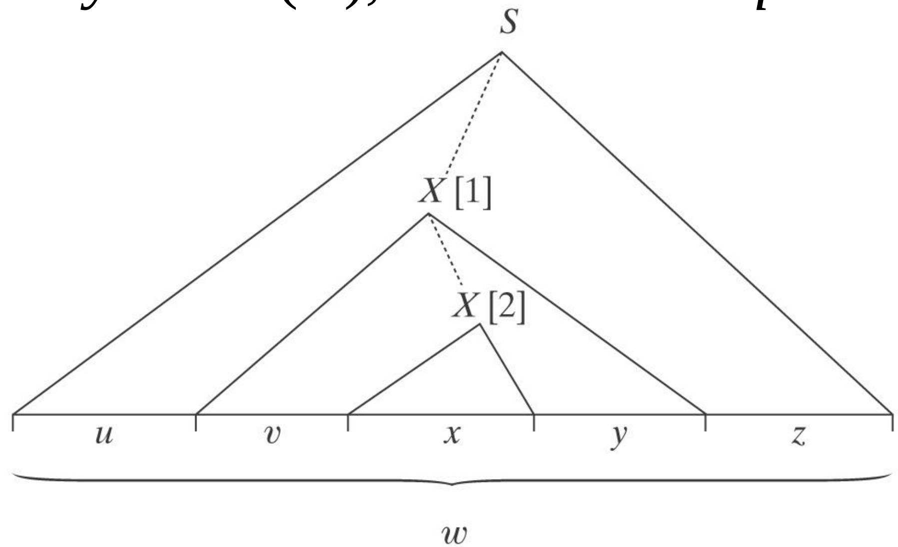
- Pada Bahasa regular, sifat Teorema Pumping berguna untuk memastikan suatu Bahasa nonregular adalah nonregular.
- Untuk Bahasa CF ada dua sifat:
  - Teorema pumping (versi CFL)
  - Ogden's Lemma
- Pada beberapa kasus bahasa, teorema pumping tidak dapat/mudah digunakan untul pembuktiannya.
- Ogden's lemma "versi lebih lemah" dari Teorema Pumping mungkin bisa digunakan.
  - Ogden's Lema bisa disebut sebagai generalisasi dari Teorema Pumpung.

Mengingat Kembali Parse Tree

- Suatu parse tree dalam derivasi menurut grammar $G = (V, \Sigma, R, S)$, adalah *rooted*, *ordered tree* yang mana:
  - *Root node* berlabel S,
  - Setiap *internal node* berlabelkan nonterminal $(V - \Sigma)$,
  - Setiap *leaf node* berlabelkan terminal $(\Sigma \cup \{\varepsilon\})$,
  - Jika X suatu *internal node* dengan cabang-cabangnya $x_1, x_2, \ldots., x_n$, hanya jika R berisi rule $X \rightarrow x_1\, x_2 \ldots\, x_n$.
- Jika grammar memenuhi CNF, maka parse tree berbentuk *binary-tree*.

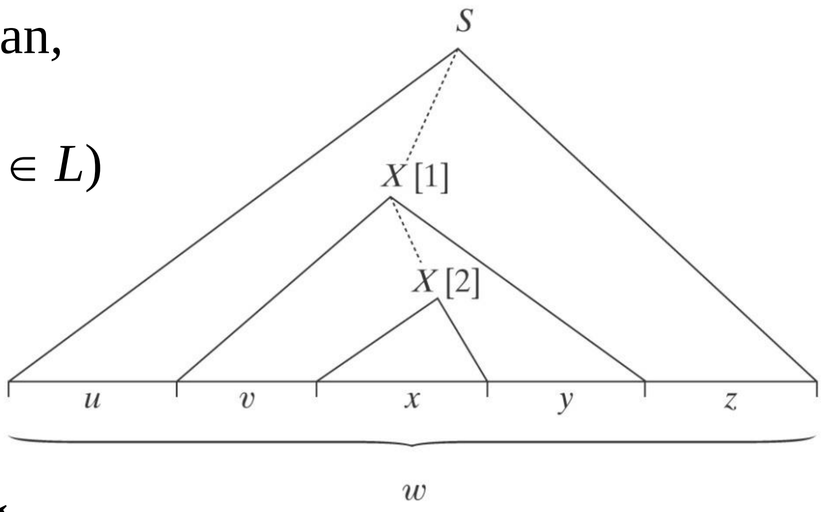
Observasi Parse Tree

- Misalkan dalam $\underline{G}$ terdapat self-embedding rule dari X sehingga dapat terjadi $X \Rightarrow^* x$ atau $X \Rightarrow^* vXy$.
- Maka dipastikan pula dapat terjadi $X \Rightarrow^* v^q x y^q$ untuk semua $q \geq 0$.
- Jika juga diketahui start symbol $S \Rightarrow^* uXz$, maka semua string $uv^q xy^q z \in L(G)$, untuk semua $q \geq 0$.



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,
    $\forall q \geq 0$
    $(uv^q xy^q z \in L)$
)
)
)



- $uv^q xy^q z$ kita sebut hasil pumping.

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 $uv^q xy^q z \in L$.
- Kemungkinan-kemungkinan pemecahan berfokus pada *v* dan *y*:
  - Semuanya a; berisi a dan b; semuanya b; berisi b dan c; atau semuanya c.
  - Tidakpernah berisi ketiga simbol karena $|vxy| \leq k$.
- Semua kemungkinan menyebabkan hasil pumping berisi deretan yang berbeda dari lainnya.

Contoh 1 (bahasa $A^nB^nC^n$, cont'd)

- Misalnya:
  - jika $k = 3$, w = aaabbbccc , sehingga $1 \leq |vy| \leq 3$.
  - Jika v dan y keduanya terdiri dari a, maka hasil pumping berisi deretan a yang berbeda dari lainya.
  - Jika v adalah a^p dan y adalah b^r maka hasil pumping berisi deretan dengan panjang berbeda-beda.
  - Jika salah satu dari v dan y berisi ab, maka terdapat$(ab)^{q-1}$ dalam hasil pumping.
- Semua kemungkinan mengarah pada hasil pumping yang bukan anggota L.

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.
  - $|w| = k^4 = (k^2)^2$, string berikutnya w'' (proper ordering) memiliki panjang $|w''| = (k^2+1)^2 = k^4+2k^2 + 1$.
  - Di lain pihak, karena $uv = a^p$, maka juga $|w'| = k^4 + p$.
  - Di atas $p \leq k$, dan karena $k < 2k^2 + 1$, string w' hanyalah string dengan panjang antara $|w|$ dan $|w''|$, dan berarti $w' \notin L$

Contoh 3

- Untuk memeriksa apakah $L = \{a^n b^m a^n : n, m \geq 0 \text{ dan } m \geq n\}$ 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 $q = 0$, deretan b lebih pendek dari deretan a.
  - (3,3): dengan $q = 2$, argumen sama dengan (1,1)
  - (1,2): dengan $q = 2$ maka argumen sama dengan (1,1)
  - (2,3): dengan $q = 2$, menghasilkan deretan a yang kedua lebih Panjang dari yang pertama
  - (1,3): tidak mungkin karena $|vxy| \leq k$.

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

Kendala Teorema Pumping

- Teorema Pumping untuk bahasa reguler memiliki kepastian posisi *y*, yaitu dalam *k* simbol awal dari *w*.
  - Jadi w bisa lebih mudah "diatur" untuk mempersempit ruang kemungkinan *y*.
- Teorema Pumping untuk CFL, *v* dan *y* bisa dimana saja dalam w asalkan keduanya berada dalam substring sepanjang *w*.
  - Beberapa bahasa non CFL sulit dibuktikan bukan CFL jika dengan Teorema Pumping.

Contoh: $L = \{a^ib^ic^j : i, j \geq 0, i \neq j\}$

- Bahasa ini secara intuitif pasti bukan CFL karena stack hanya bisa memeriksa pasangan a dan b saja.
- Dengan teorema Pumping jika digunakan, untuk k > 1
  - $w = a^kb^kc^{k+1}$ maka sifat pumping terpenuhi dengan memilih v=aa dan y=bb
  - $w = a^kb^kc^{2k}$ maka sifat pumping terpenuhi dengan memilih v=c y= ε
  - *Untuk* $w = a^kb^kc^{k+k!}$ juga sifat pumping terpenuhi dengan memilih v=c dan y=c
- ➔ kegagalan membuktikan "**L bukan CFL**".

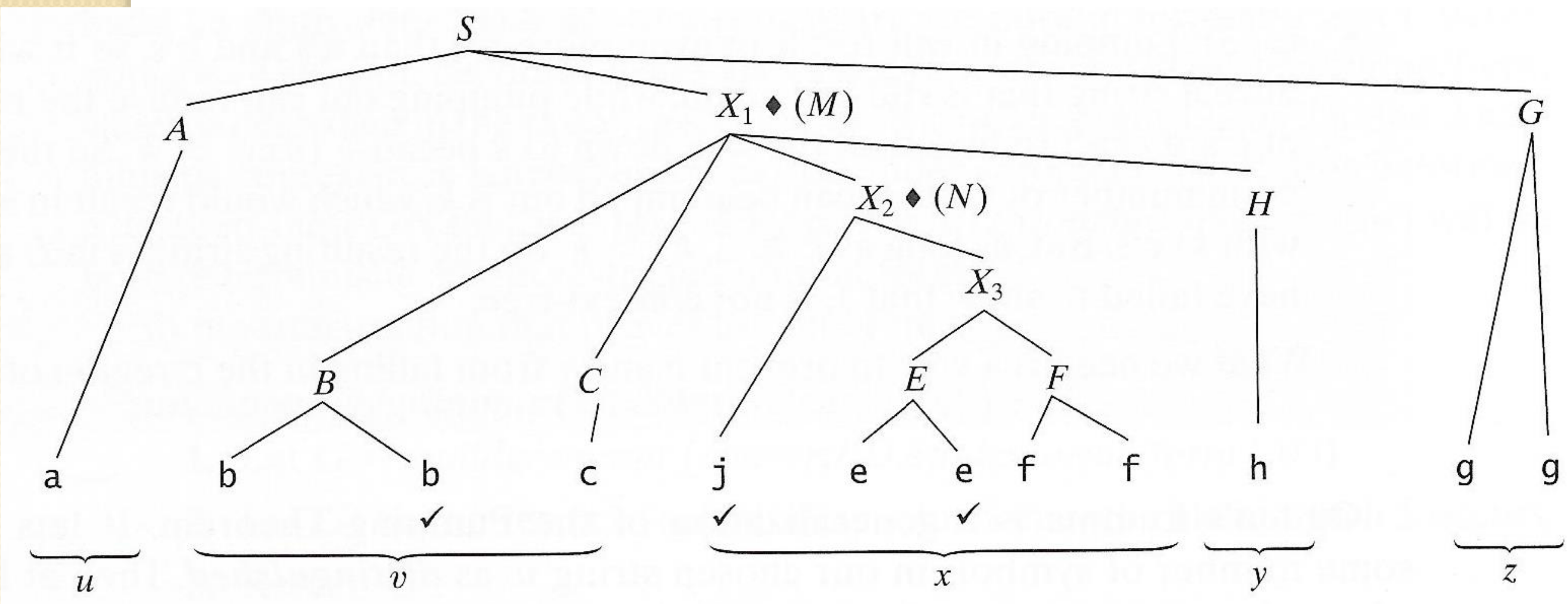
Ogden’s Lemma

- Ogden Lemma lebih powerful dari Teorema Pumping
- Ogden Lemma adalah generalisasi dari Teorema Pumping
- Menggunakan terminologi **distinguished symbol** (yaitu simbol-simbol “tertentu” dalam string *w*)

Ogden's Lemma

- Jika L adalah context-free, maka
- $\exists k \geq 1(\forall w \in L$, dengan $|w| \geq k$), bila kita menandai k buah symbol dalam w sebagai **distinguished**, maka:
  - ($\exists u, v, x, y, z$ (
    - $w = uvxyz$,
    - vy berisi setidaknya satu simbol **distinguished**,
    - vxy berisi paling banyak k simbol **distinguished**
    - $\forall q \geq 0$ ($uv^q wxy^q z \in L$)))

Ilustrasi



Contoh: $L = \{a^i b^i c^j : i, j \geq 0, i \neq j\}$

- Dengan Ogden's Lemma, kembali $w = a^k b^k c^{k+k!}$.
- Setiap a ditandai sebagai **distinguished.**
- Jika v atau y berisi dua atau lebih simbol berbeda, ambil q = 2, langsung terbukti bukan CFL.
- Untuk kemungkinan lain:
- (1, 1) dan (1, 3): ambil q= 2, panjang deretan a akan bebeda dari deretan b.
- (1, 2): Jika $v \neq y$ maka ambil q = 2, segera berbeda; jika $v = y$, pilih $q = (k!/|v|) + 1$, mengakibatkan panjang deretan a sama dengan deretan c.
- (2, 2), (2, 3), (3, 3) tidak bisa dipilih karena menyalahi - batasan teorema (segmen-2 dan segmen-3 tidak berisi distinguished)!
- Note: Ogden's lemma menghindarkan pemilihan v=c dan y=c yang tidak bisa dihindari oleh Pumping Theorem

Ogden’s Lemma sebagai Generalisasi Terorema Pumping

- Teorema Pumping adalah Ogden’s Lemma dengan menandai semua simbol.
  - Maka pembuktiannya menjadi identik dengan Teorema Pumping
  - Contoh pada pembuktian $L = \{a^i b^i c^j : i, j \geq 0, i \neq j\}$ semua simbol a,b dan c ditandai maka efektif menjadi Teorema Pmping

Potensi Masalah Ogden's Lemma

- Dalam theorem pumping segmen *vxy* berukuran maksimum *k* simbol, tetapi dalam Ogden's Lemma ukurannya tidak dibatasi selama berisi *k* distinguished symbol.
  - Keterbatasan tertentu (pilihan segmen *vxy*) berakibat ketidak terbatasan yang lain (ukuran *vxy*).
- Penting untuk menghindari adanya kemungkinan pemilihan *v* atau *y* yang memungkinkan sifat pumping dipenuhi.
  - Contoh: $L = \{a^i b^j c^i : j \geq 2i\}$