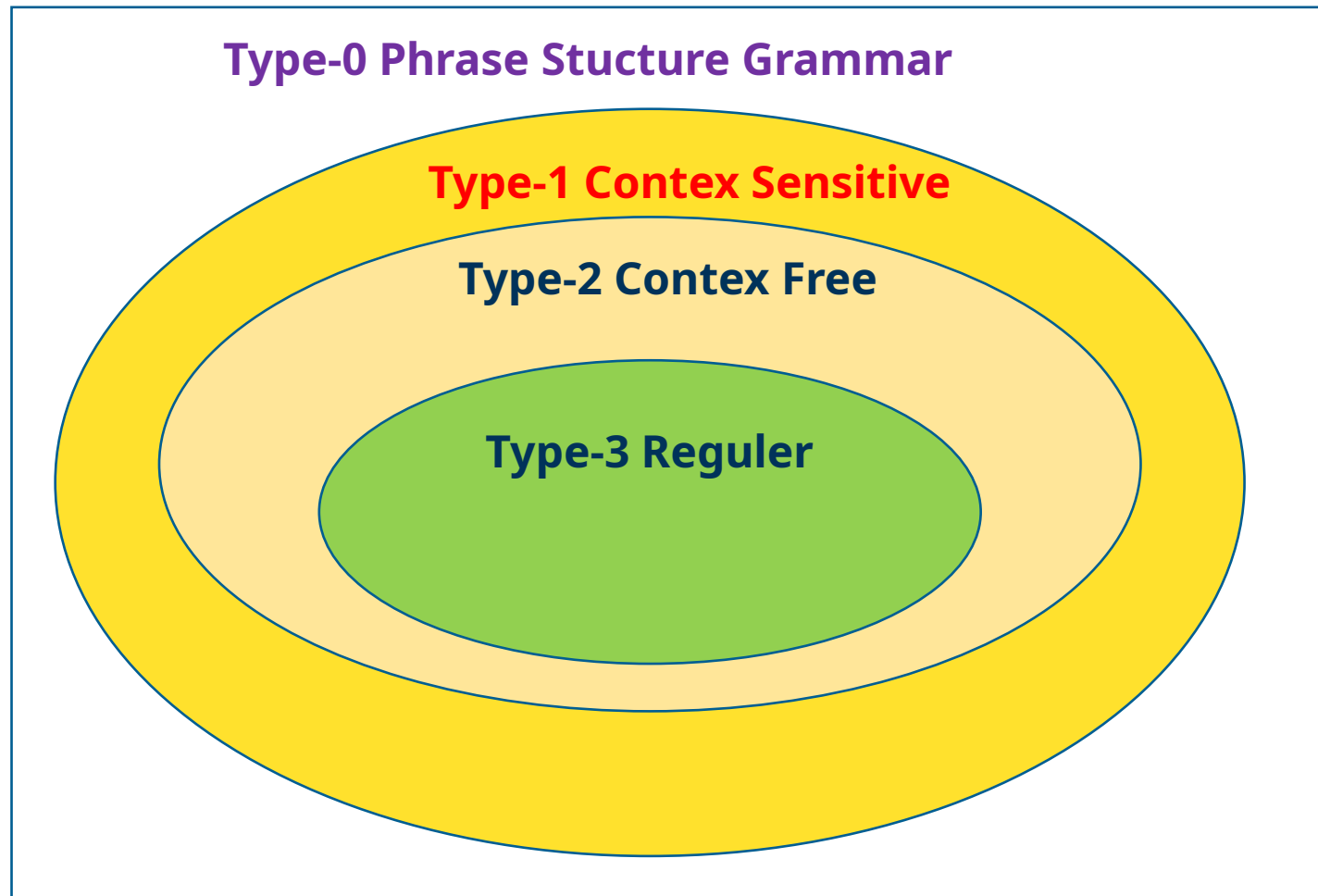


# Kuliah 2 TBO Tata Bahasa Elemen Bahasa Formal Tata Bahasa Reguler Aturan Produksi

[ukrida.ac.id](http://ukrida.ac.id)



**UKRIDA**  
Universitas Kristen Krida Wacana



**Hirarki Tata Bahasa (Grammar)**

- Tata bahasa (grammar) didefinisikan secara formal sebagai kumpulan himpunan-himpunan variabel, simbol-simbol terminal, simbol-simbol awal yang dibatasi oleh aturan tata bahasa.
- Suatu tata bahasa dapat menghasilkan sejumlah string dengan menerapkan aturan tata bahasa.

- Contoh tata bahasa :

$\alpha \rightarrow \beta$

Dibaca  $\alpha$  menghasilkan  $\beta$

- Tata bahasa berfungsi untuk menentukan kebenaran dalam penulisan suatu statement, sesuai dengan aturan yang terdapat pada suatu program.

- Tata bahasa merupakan salah satu bagian penting dalam pembuatan implementasi bahasa formal.
- Masukan yang tidak sesuai dengan tata bahasa yang telah ditetapkan menyebabkan proses tidak dapat dilakukan.
- Tata bahasa bebas konteks (Context Free Grammar) adalah salah satu yang digunakan dalam implementasi bahasa formal selain tata bahasa regular.

- Tata bahasa bebas konteks yang digunakan menghasilkan aturan produksi.
- Aturan produksi ini merupakan pusat dari tata bahasa yang menspesifikasikan bagaimana suatu tata bahasa melakukan transformasi suatu string ke bentuk lain.

- Elemen-elemen bahasa adalah alphabet, tata bahasa dan semantik.
- Alphabet adalah himpunan terhingga dari token-token dimana kalimat dibentuk dalam suatu bahasa.
- Contoh :  
 $\{a,b\} \rightarrow$  himpunan yang terdiri dari simbol "a" dan "b".
- Tata bahasa adalah himpunan dari aturan-aturan struktural yang didefinisikan yang berlaku dalam suatu kalimat pada token-token.

- Grammar  $G$  didefinisikan sebagai pasangan 4 tuple:  $V_t$ ,  $V_n$ ,  $S$  dan  $P$ , dan dituliskan sebagai  $Q(V_t, V_n, S, P)$  dimana:
  - $V_t$  : himpunan simbol-simbol terminal (alphabet)
  - $V_n$  : himpunan simbol-simbol non terminal
  - $S \in V$ : simbol awal atau start
  - $P$  : himpunan produksi



- Contoh:

$G1 : V_t = \{I, \text{want}, \text{need}, \text{you}\}, V = \{S, A, B, C\},$   
 $P = \{S \rightarrow ABC, A \rightarrow I, B \rightarrow \text{want} \mid \text{need}, C \rightarrow \text{you}\},$   
 $S \rightarrow ABC$

Iwantyou

$L(G1) = \{\text{Iwantyou}, \text{Ineedyou}\}$

- Semantik adalah himpunan aturan-aturan yang didefinisikan dan mempunyai efek operasional pada setiap program yang ditulis dalam bahasa apabila ditranslasi atau dieksekusi pada suatu mesin.
- Perbedaan bahasa inggris dengan bahasa pemrograman komputer adalah aturan-aturan ejaan dan tata bahasa dalam bahasa inggris sangat kompleks dan banyak pengecualian dan keragu-raguan, sementara dalam bahasa pemrograman harus mempunyai struktur yang tepat dan pasti

- a. Anggota alphabet dinamakan simbol terminal
- b. Kalimat adalah deretan hingga simbol-simbol terminal.
- c. Bahasa adalah himpunan kalimat-kalimat. Anggota bahasa tak hingga kalimat.
- d. Simbol-simbol berikut adalah simbol terminal:
  - Huruf kecil, misalnya : a, b, c,
  - + dan \* simbol operator, misalnya : +, -
  - Simbol tanda baca, misalnya: (,), dan ;
  - String yang tercetak tebal, misalnya **if**, **then**, dan **else**.

e. Simbol-simbol berikut adalah simbol non terminal/variabel:

- Huruf besar, misalnya : A, B, C
- Huruf S sebagai simbol awal
- String yang tercetak miring, misalnya : *expr*
- Huruf yunani melambangkan string yang tersusun atas simbol non terminal atau campuran keduanya, misalnya:  $\alpha$  ,  $\beta$  dan  $\epsilon$
- Sebuah produksi dilambangkan sebagai  $\alpha \rightarrow \beta$  , artinya: dalam sebuah derivasi dapat dilakukan penggantian simbol  $\alpha$  dengan simbol  $\beta$  .

- Derivasi adalah proses pembentukan sebuah kalimat atau sentensial. Sebuah derivasi dilambangkan sebagai :  $\alpha \rightarrow \beta$
- Sentensial adalah string yang tersusun atas simbol-simbol terminal atau simbol-simbol non terminal atau campuran keduanya.
- Kalimat adalah string yang tersusun atas simbol-simbol terminal. Kalimat adalah merupakan sentensial, sebaliknya belum tentu.

- Tata bahasa digunakan dalam implementasi bahasa formal. Tata bahasa regular yang digunakan menghasilkan aturan produksi.
- Aturan produksi merupakan pusat dari tata bahasa yang menspesifikasikan bagaimana suatu tata bahasa melakukan transformasi suatu string ke bentuk lain.
  - a. Dalam pembicaraan grammar, anggota alphabet dinamakan simbol terminal atau token.
  - b. Kalimat adalah deretan simbol.
  - c. Bahasa adalah himpunan kalimat-kalimat. Anggota bahasa tak hingga kalimat.
  - d. Simbol-simbol berikut adalah:
    - Huruf kecil, misalnya : a, b, c,
    - Simbol operator, misalnya : +, - , x
    - Simbol tanda baca, misalnya: (,), dan ;
    - String yang tercetak tebal, misalnya **if**, **then**, dan **else**.

- e. Simbol-simbol berikut adalah simbol non terminal:
  - Huruf besar, misalnya : A, B, C
  - Huruf S sebagai simbol awal
  - String yang tercetak miring, misalnya : *expr* dan *stmt*
- f. Huruf besar akhir alphabet melambangkan simbol terminal atau non terminal, misalnya: X, Y, Z
- g. Huruf kecil akhir alphabet melambangkan string yang tersusun atas simbol-simbol terminal, misalnya: x, y, z.
- h. Huruf yunani melambangkan string yang tersusun atas simbol-simbol terminal atau simbol-simbol non terminal atau campuran keduanya, misalnya:  $\alpha$  ,  $\beta$  dan  $\gamma$

- i. Sebuah produksi dilambangkan sebagai  $\alpha \rightarrow \beta$  artinya dalam sebuah derivasi dapat dilakukan penggantian simbol  $\alpha$  dengan simbol  $\beta$ .
- j. Simbol  $\alpha$  dalam aturan produksi berbentuk  $\alpha \rightarrow \beta$  disebut ruas kiri produksi sedangkan simbol  $\beta$  disebut ruas kanan produksi.
- k. Derivasi adalah proses pembentukan sebuah kalimat atau sentensial. Sebuah derivasi dilambangkan sebagai :  $\alpha \rightarrow \beta$
- l. Sentensial adalah string yang tersusun atas simbol-simbol terminal atau simbol-simbol non terminal atau campuran keduanya.
- m. Kalimat adalah string yang tersusun atas simbol-simbol terminal. Jelaslah bahwa kalimat adalah kasus khusus dari sentensial.



- n. Pengertian terminal berasal dari kata terminate (berakhir), maksudnya derivasi berakhir jika sentensial yang dihasilkan adalah sebuah kalimat (yang tersusun atas simbol-simbol terminal itu).
- o. Pengertian non terminal berasal dari kata not terminate (belum berakhir/tidak berakhir), maksudnya derivasi belum/tidak berakhir jika sentensial yang dihasilkan mengandung simbol non terminal.

- Aturan produksi  $\alpha \rightarrow \beta$  yang diterapkan pada suatu string  $w = a\alpha c$  mengganti kemunculan  $\alpha$  menjadi  $\beta$ , sehingga string tersebut berubah menjadi  $w = a\beta c$ , sehingga dituliskan  $a\alpha c \rightarrow a\beta c$  ( $a\alpha c$  memproduksi  $a\beta c$ ).
- Produksi tersebut dapat diterapkan berkali-kali,  $w_1 \rightarrow w_2 \rightarrow w_3 \dots \rightarrow w_n$  atau dapat ditulis  $w_1 \rightarrow^* w_n$  jika minimal harus ada aturan produksi yang diterapkan :  $w_1 \rightarrow^+ w_n$

- Contoh

Tata bahasa  $G = \{\{S,A\}, \{a,b\}, S, P\}$  dengan aturan produksi  $P$  adalah:

$$S \rightarrow Ab$$

$$A \rightarrow aAb$$

$$A \rightarrow \epsilon$$

Maka dapat dihasilkan suatu string

$$S \rightarrow Ab \rightarrow b$$

$$S \rightarrow Ab \rightarrow aAbb \rightarrow abb$$

$$S \rightarrow Ab \rightarrow aAbb \rightarrow aaAbbb \rightarrow aabbbb$$

Bahasa yang dihasilkan dari tata bahasa tersebut adalah:

$L(G) = \{b, abb, aabbb, aaabbbb, aaaabbbbb, \dots\}$   
atau dapat dituliskan  $L(G) = \{a^n b^{n+1} \mid n \geq 0\}$

## Hierarki Bahasa

Kelas	Mesin Pengenal
Regular language	Finite State Automata
Context free language	Push Down Automata
Context sensitive language	Linear Bounded Automata
Unrestricted language	Turing Machine

Kelas	Ruas kiri	Ruas kanan	Contoh
Regular	$\alpha \in N$	$\leq 1$ non terminal (paling kiri/kanan)	$P \rightarrow abR$ $Q \rightarrow abc$ $R \rightarrow Scac$
Context free	$\alpha \in N$	-	$P \rightarrow aQb$ $Q \rightarrow abPRS$
Context sensitive	$\alpha \in (T \cup N)^+$	$ \alpha  \leq  \beta $	$aD \rightarrow Da$ $AD \rightarrow aCD$
Unrestricted	$\alpha \in (T \cup N)^+$	-	$CB \rightarrow DB$ $Adc \rightarrow \varepsilon$

Ruas kiri harus memuat simbol non terminal

# Derivasi Kalimat dan Penentuan Bahasa



**UKRIDA**

- Tentukan bahasa dari masing-masing grammar berikut:

1. G1 dengan  $Q_1 = \{1. S \rightarrow aAa, 2. A \rightarrow aAa, 3. A \rightarrow b\}$ .

**Jawab:**

Derivasi kalimat terpendek: derivasi kalimat umum:

$$\begin{array}{ll} S \rightarrow aAa & (1) \\ \rightarrow aba & (3) \end{array} \quad \begin{array}{ll} S \rightarrow aAa & (1) \\ \rightarrow aaAaa & (2) \\ \dots & \\ \rightarrow a^n A a^n & (2) \\ \rightarrow a^n b a^n & (3) \end{array}$$

Dari pola kedua kalimat disimpulkan :

$$L^1(G^1) = \{a^n b a^n \mid n \geq 1\}$$

# Derivasi Kalimat dan Penentuan Bahasa



**UKRIDA**

## 2. G3 dengan

$Q3 = \{1. S \rightarrow aSBC, 2. S \rightarrow abC, 3. bB \rightarrow bb, 4. bC \rightarrow bc, 5. CB \rightarrow BC, 6. cC \rightarrow cc\}.$

**Jawab:**

Derivasi kalimat terpendek 1: Derivasi kalimat terpendek 3:

$S \rightarrow abC$	(2)	$S \rightarrow aSBC$	(1)
$\rightarrow abc$	(4)	$\rightarrow aaSBCBC$	(1)
Derivasi kalimat terpendek 2:		$\rightarrow aaabCBCBC$	(2)
$S \rightarrow aSBC$	(1)	$\rightarrow aaabBCCBC$	(5)
$\rightarrow aabCBC$	(2)	$\rightarrow aaabBCBCC$	(5)
$\rightarrow aabBCC$	(5)	$\rightarrow aaabBBCCC$	(5)
$\rightarrow aabbCC$	(3)	$\rightarrow aaabbBCCC$	(3)
$\rightarrow aabbcC$	(4)	$\rightarrow aaabbbCCC$	(3)
$\rightarrow aabbcc$	(6)	$\rightarrow aaabbbbcCC$	(4)
		$\rightarrow aaabbbccC$	(6)
		$\rightarrow aaabbbccc$	(6)

Dari pola ketiga kalimat dapat disimpulkan:  $L^3(G^3) = \{a^n b^n c^n | n \geq 1\}$

# Derivasi Kalimat dan Penentuan Bahasa



**UKRIDA**

- Aturan yang disebutkan pada proses pengenalan dan pembangkitan kalimat. Secara formal, tata bahasa terdiri dari 4 komponen yaitu:
  1. Himpunan berhingga, tidak kosong dari simbol-simbol dan non terminal  $T^1$ .
  2. Himpunan berhingga, dari simbol-simbol non terminal  $N$ .
  3. Simbol awal  $S \in N$ , yang merupakan salah satu anggota dari himpunan simbol non terminal.
  4. Himpunan berhingga aturan produksi  $P$  yang setiap elemennya dituliskan dalam bentuk:  $\alpha \rightarrow \beta$ , dimana  $\alpha$  dan  $\beta$  adalah string yang dibentuk dari himpunan  $T \cup N$  dan  $\alpha$  harus berisi paling sedikit satu simbol non terminal.



# Derivasi Kalimat dan Penentuan Bahasa UKRIDA

- Keempat komponen tersebut sering dituliskan sebagai berikut:  $G = (T, N, S, P)$ .
- Bahasa yang dihasilkan oleh  $G$  ditulis  $L(G)$ , yaitu himpunan string yang dapat diturunkan dari simbol awal  $S$  dengan menerapkan aturan aturan produksi yang terdapat pada  $P$ .

1. Tentukan sebuah grammar regular untuk bahasa  $L^1 = \{a^n | n \geq 1\}$
2. Tentukan sebuah grammar bebas konteks untuk bahasa  $L^2$ : himpunan bilangan bulat non negatif ganjil.
3. Tentukan sebuah grammar bebas konteks untuk bahasa  $L^4(G^4) = \{a^n b^m | n, m \geq 1, n \neq m\}$
4. Tentukan sebuah grammar bebas konteks untuk bahasa  $L^5$ : bilangan bulat non negatif genap. Jika bilangan tersebut terdiri dari dua digit atau lebih maka nol tidak boleh muncul sebagai digit pertama.
5.  $G = (\{S, A, B, C\}, [a, b], S, Q)$   
Aturan produksi:

$$S \rightarrow aS | aB$$

$$B \rightarrow bC$$

$$C \rightarrow aC | a$$

# Latihan soal



**UKRIDA**

1. Tentukan sebuah grammar regular untuk bahasa  $L^1 = \{a^n | n \geq 1\}$   
 $Q^1(L^1) = \{S \rightarrow aS | a\}$
2. Tentukan sebuah grammar bebas konteks untuk bahasa  $L^2$ : himpunan bilangan bulat non negatif ganjil.

Langkah : digit terakhir bilangan ganjil

Buat dua buah himpunan bilangan terpisah: genap (G) dan ganjil (J)

$$Q^2(L^2) = \{S \rightarrow J | GS | JS, G \rightarrow 0 | 2 | 4 | 6, J \rightarrow 1 | 3 | 5 | 7 | 9\}$$

3. Tentukan sebuah grammar bebas konteks untuk bahasa  $L^4(G^4) = \{a^n b^m | n, m \geq 1, n \neq m\}$

Langkah kunci: sulit mendefinisikan  $L^4(G^4)$  secara langsung. Langkah penyelesaiannya adalah dengan mengingat bahwa  $x \neq y$  berarti  $x > y$  atau  $x < y$



$$\begin{aligned} L^4 &= L^A \cup L^B, & L^A &= \{a^n b^m \mid n > m \geq 1\}, & L^B &= \{a^n b^m \mid 1 \leq n < m\} \\ Q^A(L^A) &= \{A \rightarrow aA \mid aC, C \rightarrow aCb \mid ab\}, & Q(L^B) &= \{B \rightarrow Bb \mid Db, D \rightarrow aDb \mid ab\} \\ Q^4(L^4) &= B\{S \rightarrow A \mid B, A \rightarrow aA \mid aC, C \rightarrow aCb \mid ab, B \rightarrow Bb \mid Db, D \rightarrow aDb \mid ab\} \end{aligned}$$

4. Tentukan sebuah grammar bebas konteks untuk bahasa  $L^5$ : bilangan bulat non negatif genap. Jika bilangan tersebut terdiri dari dua digit atau lebih maka nol tidak boleh muncul sebagai digit pertama.

Langkah kunci: digit terakhir bilangan harus genap. Digit pertama tidak boleh nol.

Buat tiga himpunan terpisah: bilangan genap tanpa nol  $\{G\}$ , bilangan genap dengan nol  $\{N\}$ , serta bilangan ganjil  $\{J\}$ .

$$Q^5(L^5) = \{S \rightarrow N \mid GA \mid JA, A \rightarrow N \mid NA \mid JA, G \rightarrow 2 \mid 4 \mid 6 \mid 8, N \rightarrow 0 \mid 2 \mid 4 \mid 6 \mid 8, J \rightarrow 1 \mid 3 \mid 5 \mid 7 \mid 9\}$$

5.  $G = (\{S, A, B, C\}, [a, b], S, Q)$

Aturan produksi:

$$S \rightarrow aS \mid aB$$

$$B \rightarrow bC$$

$$C \rightarrow aC \mid a$$

Kalimat  $a^3ba^2$  atau  $aaabaa$  merupakan derivasi berikut:

$$\begin{aligned} S &\rightarrow aS \\ &\rightarrow aaS \\ &\rightarrow aaaB \\ &\rightarrow aaabC \\ &\rightarrow aaabaa \end{aligned}$$

# Thank You

[ukrida.ac.id](http://ukrida.ac.id)



**UKRIDA**  
Universitas Kristen Krida Wacana