

GRAMMAR DAN BAHASA

MATERI MINGGU KE-2

TATA BAHASA

- Dalam pembicaraan tata bahasa, anggota alfabet dinamakan simbol terminal atau token.
- Kalimat adalah deretan hingga simbo-lsimbol terminal.
- Bahasa adalah himpunan kalimat-kalimat. Anggota bahasa bisa tak hingga kalimat.

TATA BAHASA

- **Simbol-Simbol Terminal:**
 - huruf kecil awal alfabet, misalnya : a, b, c
 - simbol operator, misalnya : +, −, dan ×
 - simbol tanda baca, misalnya : () ,, dan ;
 - string yang tercetak tebal, misalnya : **if**, **then** dan **else**
- **Simbol-Simbol Non Terminal:**
 - huruf besar awal alfabet, misal: A, B, C
 - huruf S sebagai sebagai simbol awal
 - String yang tercetak miring, misalnya : *expr* dan *stmt*.

TATA BAHASA

- Huruf besar akhir alfabet melambangkan simbol terminal atau non terminal, misalnya : X, Y, Z.
- Huruf kecil akhir alfabet melambangkan string yang tersusun atas simbol-simbol terminal, misalnya : xyz
- Huruf yunani melambangkan string yang tersusun atas simbol-simbol terminal atau simbol-simbol non terminal atau campuran keduanya, misalnya : α , β , dan γ .

TATA BAHASA

Sebuah produksi dilambangkan sebagai $\alpha \rightarrow \beta$, artinya : dalam sebuah derivasi dapat dilakukan penggantian simbol α dengan simbol β .

$$\alpha \longrightarrow \beta$$

Simbol α dalam produksi berbentuk $\alpha \rightarrow \beta$. α disebut ruas kiri produksi sedangkan simbol β disebut ruas kanan produksi.

TATA BAHASA

- **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, sehingga kalimat merupakan kasus khusus dari sentensial.
- **Terminal** berasal dari kata *terminate* (berakhir), maksudnya derivasi berakhir jika sentensial yang dihasilkan adalah sebuah kalimat (yang tersusun atas simbol-simbol terminal itu).
- **Non Terminal** berasal dari kata *not terminate* (belum/tidak berakhir), maksudnya derivasi belum/tidak berakhir jika sentensial yang dihasilkan mengandung simbol non terminal.

TATA BAHASA

ATURAN PRODUKSI

Aturan produksi dinyatakan dalam bentuk $\alpha \rightarrow \beta$

- α menghasilkan atau menurunkan β
- α symbol-symbol untuk ruas kiri, β symbol-symbol untuk ruas kanan
- Symbol-symbol dapat berupa terminal dan non terminal dimana non terminal dapat diturunkan menjadi symbol yang lainnya.
- Umumnya symbol terminal disymbolkan dengan huruf kecil (a,b,c, dsb), sedangkan untuk symbol non terminal disymbolkan dengan huruf besar (A,B,C, dsb)

Contoh:

$T \rightarrow a$

“T menghasilkan a”

$T \rightarrow E \mid E + A$

“ T menghasilkan E atau T menghasilkan E + A

GRAMMAR

Grammar G didefinisikan sebagai pasangan 4 tuple : V_T , V_N , S , dan Q , dan dituliskan sebagai $\mathbf{G(V_T, V_N, S, Q)}$, dimana :

- V_T : himpunan simbol-simbol terminal (atau himpunan token - token, atau alfabet)
- V_N : himpunan simbol-simbol non terminal
- $S \in V_N$: simbol awal (atau simbol start)
- Q : himpunan produksi

Contoh :

$$G_1 : V_T = \{a\}, \quad V_N = \{S\}, \quad Q = \{S \rightarrow aS \mid a\}$$

$$S \Rightarrow aS$$

$$\Rightarrow aaS$$

$$\Rightarrow aaa$$

$$L(G_1) = \{a, aa, aaa, aaaa, \dots\}$$

$$\mathbf{L(G_1) = \{a^n \mid n \geq 1\}}$$

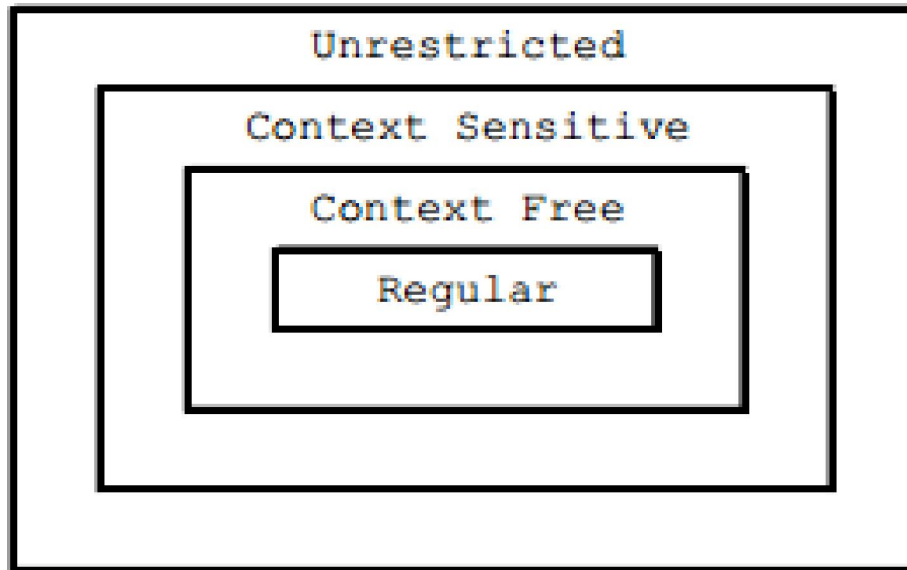
GRAMMAR

Tipe sebuah grammar (atau bahasa) ditentukan dengan aturan sebagai berikut :

A language is said to be type- i ($i = 0, 1, 2, 3$) language if it can be specified by a type- i grammar but can't be specified any type- $(i+1)$ grammar.

HIRARKI CHOMSKY

Ada 4(empat) kelas pengelompokan suatu bahasa, yang dikenal dengan “**Chomsky Hierarchy**”. Hirarki atau tingkatan bahasa ini dikembangkan oleh Noam Chomsky pada tahun 1959.



Berdasarkan komposisi bentuk ruas kiri dan ruas kanan produksinya ($\alpha \rightarrow \beta$), Noam Chomsky mengklasifikasikan 4 tipe grammar :

HIRARKI CHOMSKY

1. Grammar tipe ke-0 : Unrestricted Grammar (UG)

$$\alpha, \beta \in (V_T \mid V_N)^*, |\alpha| > 0$$

Ciri : Tidak ada batasan pada aturan produksi

Mesin pengenalan bahasa disebut : **Mesin Turing**

Contoh :

- $Abc \rightarrow De$
- $G = (V, T, P, S)$
 $V = \{S, A, B, G\}$
 $T = \{a, b, d\}$
 $Q : S \rightarrow aSa$
 $A \rightarrow bdG$
 $AB \rightarrow a$

HIRARKI CHOMSKY

2. Grammar tipe ke-1 : Context Sensitive Grammar (CSG)

$$\alpha, \beta \in (V_T \mid V_N)^*, 0 < |\alpha| \leq |\beta|$$

Ciri : Panjang string ruas kiri harus $<$ (lebih kecil) atau $=$ (sama dengan) ruas kanan.

Mesin pengenalan bahasa disebut : **Linear Bounded Automata (LBA)**

Contoh :

$$G = \{V, T, P, S\}$$

$$V = \{S, B, C\}$$

$$T = \{a, b, c\}$$

$$Q : S \rightarrow aSBC \mid aBC \mid \varepsilon$$

$$CB \rightarrow BC$$

$$aB \rightarrow ab$$

$$bB \rightarrow bb$$

$$bC \rightarrow bc$$

$$cC \rightarrow cc$$

Keterangan : $S \rightarrow \varepsilon$, karena S adalah simbol awal, maka ini juga memenuhi Walau panjang $S = 1$ dan panjang $\varepsilon = 0$

HIRARKI CHOMSKY

3. Grammar tipe ke-2 : Context Free Grammar (CFG)

$$\alpha \in V_N, \beta \in (V_T | V_N)^*$$

Ciri : Ruas kiri haruslah tepat satu symbol variabel, yaitu simbol non terminal

Mesin pengenalan bahasa disebut : **Push Down Automata (PDA)**

Contoh :

$$G = \{V, T, P, S\}$$

$$V = \{S, A, B\}$$

$$T = \{a, b\}$$

$$Q : S \rightarrow aB \mid bA$$

$$A \rightarrow a \mid aS \mid bAA$$

$$B \rightarrow b \mid bS \mid aBB$$

HIRARKI CHOMSKY

4. Grammar tipe ke-3 : Regular Grammar (RG)

$$\alpha \in V_N, \beta \in \{V_T, V_T V_N\} \text{ atau } \alpha \in V_N, \beta \in \{V_T, V_N V_T\}$$

Ciri :

- Ruas kiri hanya memiliki maksimal satu symbol non terminal
- α adalah simbol nonterminal tunggal
- β maksimal memiliki maksimal satu simbol non terminal tunggal dan ditempatkan pada posisi paling kanan.

Mesin pengenalan bahasa disebut : **Finite State Automata (FSA)**

Contoh :

$$G = (V, T, P, S)$$

$$V = \{S, A, B\}$$

$$T = \{0, 1\}$$

$$Q : S \rightarrow 0A \mid 1B \mid 0$$

$$A \rightarrow 0A \mid 0S \mid 1B$$

$$B \rightarrow 1B \mid 1 \mid 0 \mid \varepsilon$$

Contoh Analisa Penentuan Type Grammar

1. Grammar G_1 dengan $Q_1 = \{S \rightarrow aB, B \rightarrow bB, B \rightarrow b\}$.

- Ruas kiri semua produksinya terdiri dari sebuah V_N maka G_1 kemungkinan tipe CFG atau RG.
- Karena semua ruas kanannya terdiri dari sebuah V_T atau string $V_T V_N$ maka G_1 adalah RG.

2. Grammar G_2 dengan $Q_2 = \{S \rightarrow Ba, B \rightarrow Bb, B \rightarrow b\}$.

- Ruas kiri semua produksinya terdiri dari sebuah V_N maka G_2 kemungkinan tipe CFG atau RG.
- Selanjutnya karena semua ruas kanannya terdiri dari sebuah V_T atau string $V_N V_T$ maka G_2 adalah RG

Contoh Analisa Penentuan Type Grammar

3. Grammar G_3 dengan $Q_3 = \{S \rightarrow aA, S \rightarrow aB, aAb \rightarrow aBCb\}$.

- Ruas kirinya mengandung string yang panjangnya lebih dari 1 (yaitu aAb) maka G_3 kemungkinan tipe CSG atau UG
- karena semua ruas kirinya lebih pendek atau sama dengan ruas kananya maka G_3 adalah CSG

4. Grammar G_4 dengan $Q_4 = \{aS \rightarrow ab, SAc \rightarrow bc\}$

- Ruas kirinya mengandung string yang panjangnya lebih dari 1 maka G_4 kemungkinan tipe CSG atau UG
- Karena terdapat ruas kirinya yang lebih panjang daripada ruas kananya (yaitu SAc) maka G_4 adalah UG.

DERIVASI KALIMAT DAN PENENTUAN BAHASA

Tentukan bahasa dari masing-masing gramat berikut :

1. G_1 dengan $Q_1 = \{1. S \rightarrow aAa, 2. A \rightarrow aAa, 3. A \rightarrow b\}$.

Jawab :

Derivasi kalimat terpendek :

$$S \Rightarrow aAa \quad (1)$$

$$\Rightarrow aba \quad (3)$$

Derivasi kalimat umum :

$$S \Rightarrow aAa \quad (1)$$

$$\Rightarrow aaAaa \quad (2)$$

...

$$\Rightarrow a^n A a^n \quad (2)$$

$$\Rightarrow a^n b a^n \quad (3)$$

Dari pola kedua kalimat disimpulkan : $L_1(G_1) = \{a^n b a^n \mid n \geq 1\}$

DERIVASI KALIMAT DAN PENENTUAN BAHASA

2. G_2 dengan $Q_2 = \{1. S \rightarrow aS, 2. S \rightarrow aB, 3. B \rightarrow bC, 4. C \rightarrow aC, 5. C \rightarrow a\}$.

Jawab :

Derivasi kalimat terpendek :

$$S \Rightarrow aB \quad (2)$$

$$\Rightarrow abC \quad (3)$$

$$\Rightarrow aba \quad (5)$$

Derivasi kalimat umum :

$$S \Rightarrow aS \quad (1)$$

...

$$\Rightarrow a^{n-1}S \quad (1)$$

$$\Rightarrow a^nB \quad (2)$$

$$\Rightarrow a^nbC \quad (3)$$

$$\Rightarrow a^nbaC \quad (4)$$

...

$$\Rightarrow a^nba^{m-1}C \quad (4)$$

$$\Rightarrow a^nba^m \quad (5)$$

Dari pola kedua kalimat disimpulkan : $L(G_2) = \{ a^nba^m \mid n \geq 1, m \geq 1 \}$

DERIVASI KALIMAT DAN PENENTUAN BAHASA

3. G_3 dengan $Q_3 = \{1. S \rightarrow aSBC, 2. S \rightarrow abC, 3. bB \rightarrow bb, 4. bC \rightarrow bc, 5. CB \rightarrow BC, 6. cC \rightarrow cc\}$.

Jawab :

Derivasi kalimat terpendek 1:

$S \Rightarrow abC$ (2)

$\Rightarrow abc$ (4)

Derivasi kalimat terpendek 2 :

$S \Rightarrow aSBC$ (1)

$\Rightarrow aabCBC$ (2)

$\Rightarrow aabBCC$ (5)

$\Rightarrow aabbCC$ (3)

$\Rightarrow aabbcC$ (4)

$\Rightarrow aabbcc$ (6)

Derivasi kalimat terpendek 3 :

$S \Rightarrow aSBC$ (1)

$\Rightarrow aaSBCBC$ (1)

$\Rightarrow aaabCBCBC$ (2)

$\Rightarrow aaabBCCBC$ (5)

$\Rightarrow aaabBCBCC$ (5)

$\Rightarrow aaabBBCCC$ (5)

$\Rightarrow aaabbBCCC$ (3)

$\Rightarrow aaabbbCCC$ (3)

$\Rightarrow aaabbbbcCC$ (4)

$\Rightarrow aaabbbccC$ (6)

$\Rightarrow aaabbbccc$ (6)

Dari pola ketiga kalimat disimpulkan : $L_3 (G_3) = \{ a^n b^n c^n \mid n \geq 1 \}$

Menentukan Grammar Sebuah Bahasa

1. Tentukan sebuah grammar regular untuk bahasa $L_1 = \{ a^n \mid n \geq 1 \}$

Jawab :

$$Q_1(L_1) = \{ S \rightarrow aS \mid a \}$$

2. Tentukan sebuah grammar bebas konteks untuk bahasa L_2 : himpunan bilangan bulat non negatif ganjil

Jawab :

Langkah kunci : digit terakhir bilangan harus ganjil.

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

$$Q_2(L_2) = \{ S \rightarrow J \mid GS \mid JS, G \rightarrow 0 \mid 2 \mid 4 \mid 6 \mid 8, J \rightarrow 1 \mid 3 \mid 5 \mid 7 \mid 9 \}$$

Menentukan Grammar Sebuah Bahasa

3. Tentukan sebuah grammar bebas konteks untuk bahasa :

L_3 = himpunan semua identifier yang sah menurut bahasa pemrograman Pascal dengan batasan : terdiri dari simbol huruf kecil dan angka, panjang identifier boleh lebih dari 8 karakter

Jawab :

Langkah kunci : karakter pertama identifier harus huruf.

Buat dua himpunan bilangan terpisah : huruf (H) dan angka (A)

$$Q_3(L_3) = \{S \rightarrow H \mid HT, T \rightarrow AT \mid HT \mid H \mid A, \\ H \rightarrow a \mid b \mid c \mid \dots, A \rightarrow 0 \mid 1 \mid 2 \mid \dots\}$$

Menentukan Grammar Sebuah Bahasa

4. Tentukan gramat bebas konteks untuk bahasa

$$Q_4(L_4) = \{a^n b^m \mid n, m \geq 1, n \neq m\}$$

Jawab :

Langkah kunci : sulit untuk mendefinisikan L_4 (G_4) secara langsung. Jalan keluarnya adalah dengan mengingat bahwa $x \neq y$ berarti $x > y$ atau $x < y$.

$$L_4 = L_A \cup L_B, L_A = \{a^n b^m \mid n > m \geq 1\}, L_B = \{a^n b^m \mid 1 \leq n < m\}.$$

$$Q_A(L_A) = \{A \rightarrow aA \mid aC, C \rightarrow aCb \mid ab\}, Q(L_B) = \{B \rightarrow Bb \mid Db, D \rightarrow aDb \mid ab\}$$

$$Q_4(L_4) = \{S \rightarrow A \mid B, A \rightarrow aA \mid aC, C \rightarrow aCb \mid ab, B \rightarrow Bb \mid Db, D \rightarrow aDb \mid ab\}$$

Menentukan Grammar Sebuah Bahasa

5. 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.

Jawab :

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).

$$Q_5(L_5) = \{S \rightarrow N \mid GA \mid JA, A \rightarrow N \mid NA \mid JA, \\ G \rightarrow 2 \mid 4 \mid 6 \mid 8, N \rightarrow 0 \mid 2 \mid 4 \mid 6 \mid 8, J \rightarrow 1 \mid 3 \mid 5 \mid 7 \mid 9\}$$



TERIMA KASIH

Lilis Setyowati