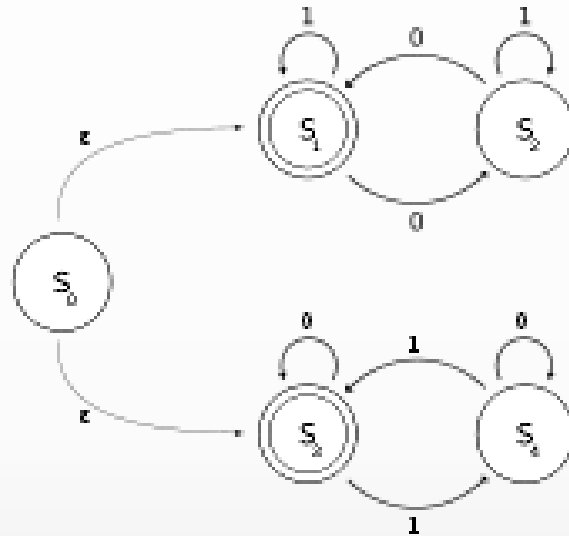


Otomata dan Teori Bahasa Formal



Dr. Hanny Hikmayanti, M.Kom

Hilda Yulia Novita, M.Kom

✉ : hanny.hikmayanti@ubpkarawang.ac.id

✉ : hilda.yulia@ubpkarawang.ac.id

☎ : +62 857 1414 4026



**PROGRAM STUDI TEKNIK INFORMATIKA
UNIVERSITAS BUANA PERJUANGAN
KARAWANG**

Pertemuan ke-3

Hirarki Chomsky

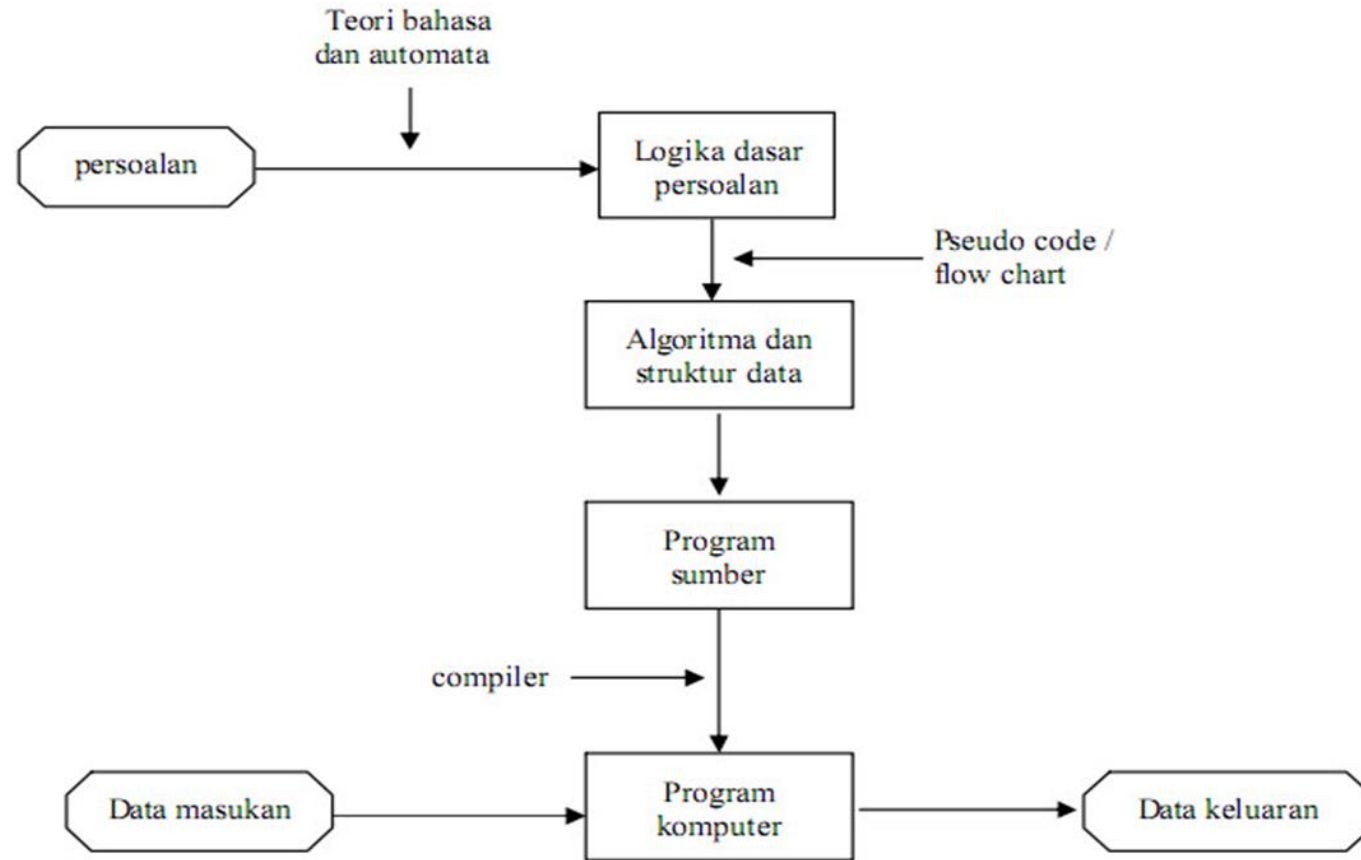
Deskripsi OTB

- Otamata dan Teori Bahasa (OTB) matakuliah wajib
- Komponen utama dari Ilmu Komputer yaitu model dan gagasan mendasar mengenai komputasi.
- Suatu bentuk yang memiliki fungsi-fungsi dari komputer digital, menerima input, menghasilkan output, bisa memiliki penyimpanan sementara, dan mampu membuat keputusan dalam menstransformasikan dari input ke output

Otomata & Teori Bahasa?

- Bagian dari teori komputasi pada ilmu komputer
 - model matematika yang memiliki fungsi dari komputer digital yaitu menerima input, menghasilkan output, bisa memiliki penyimpanan sementara dan mampu membuat keputusan dalam transformasi input ke output
- Penekanannya adalah pada **pemecahan masalah**

Otomata & Teori Bahasa?



Kedudukan OTB

Ilmu komputer memiliki dua komponen utama:

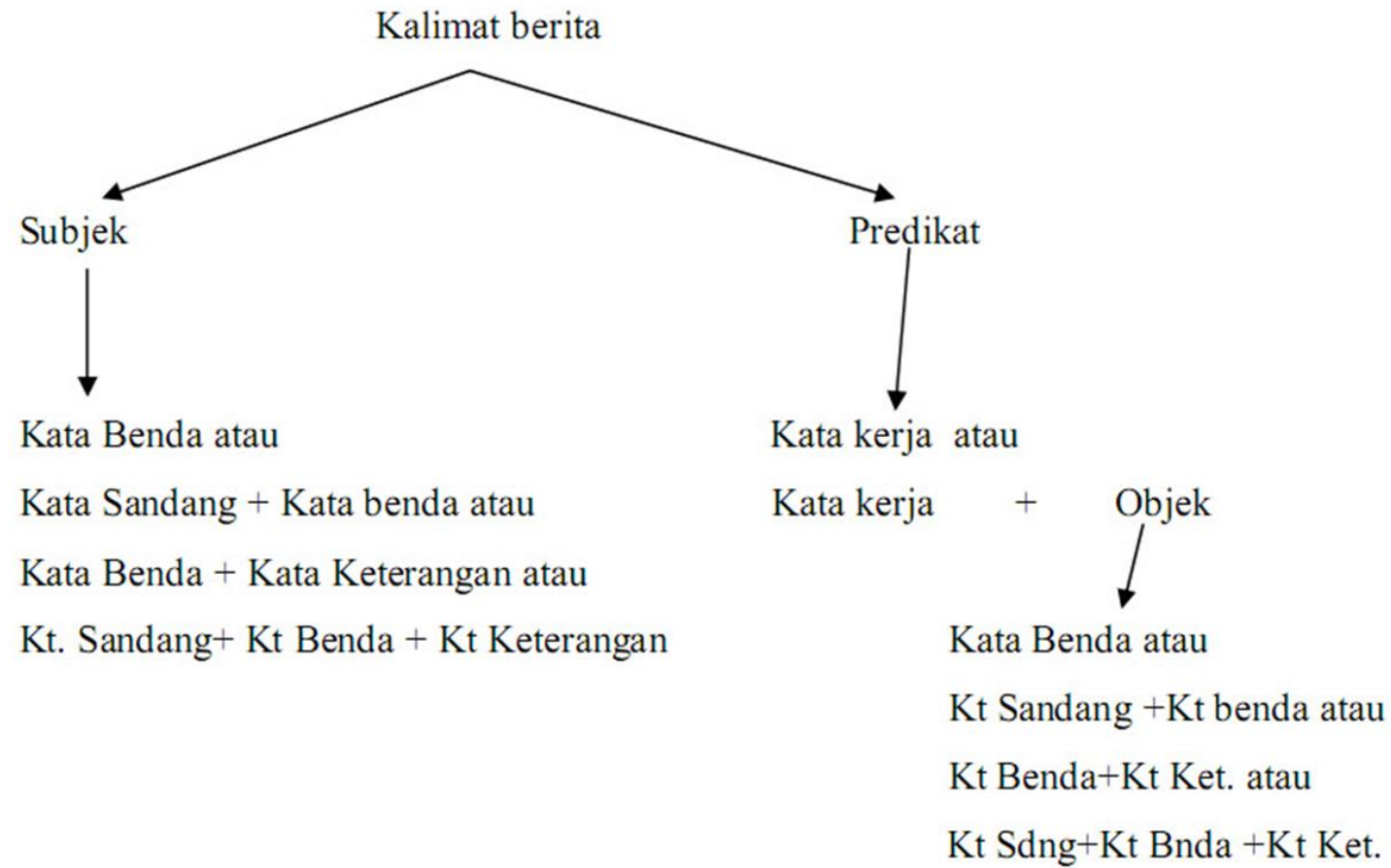
- Pertama: model dan gagasan mendasar mengenai komputasi
- Kedua: teknik rekayasa untuk perancangan sistem komputasi, meliputi perangkat keras dan perangkat lunak, khususnya penerapan rancangan dari teori.

Otomata dan Teori Bahasa merupakan bagian pertama

Tata Bahasa

- Penulisan suatu kalimat dalam sebuah bahasa, akan mengikuti suatu aturan tertentu yang berlaku pada bahasa tersebut
- Aturan tersebut dikenal sebagai Tata Bahasa (Grammar)

Tata Bahasa



Klasifikasi Tata Bahasa

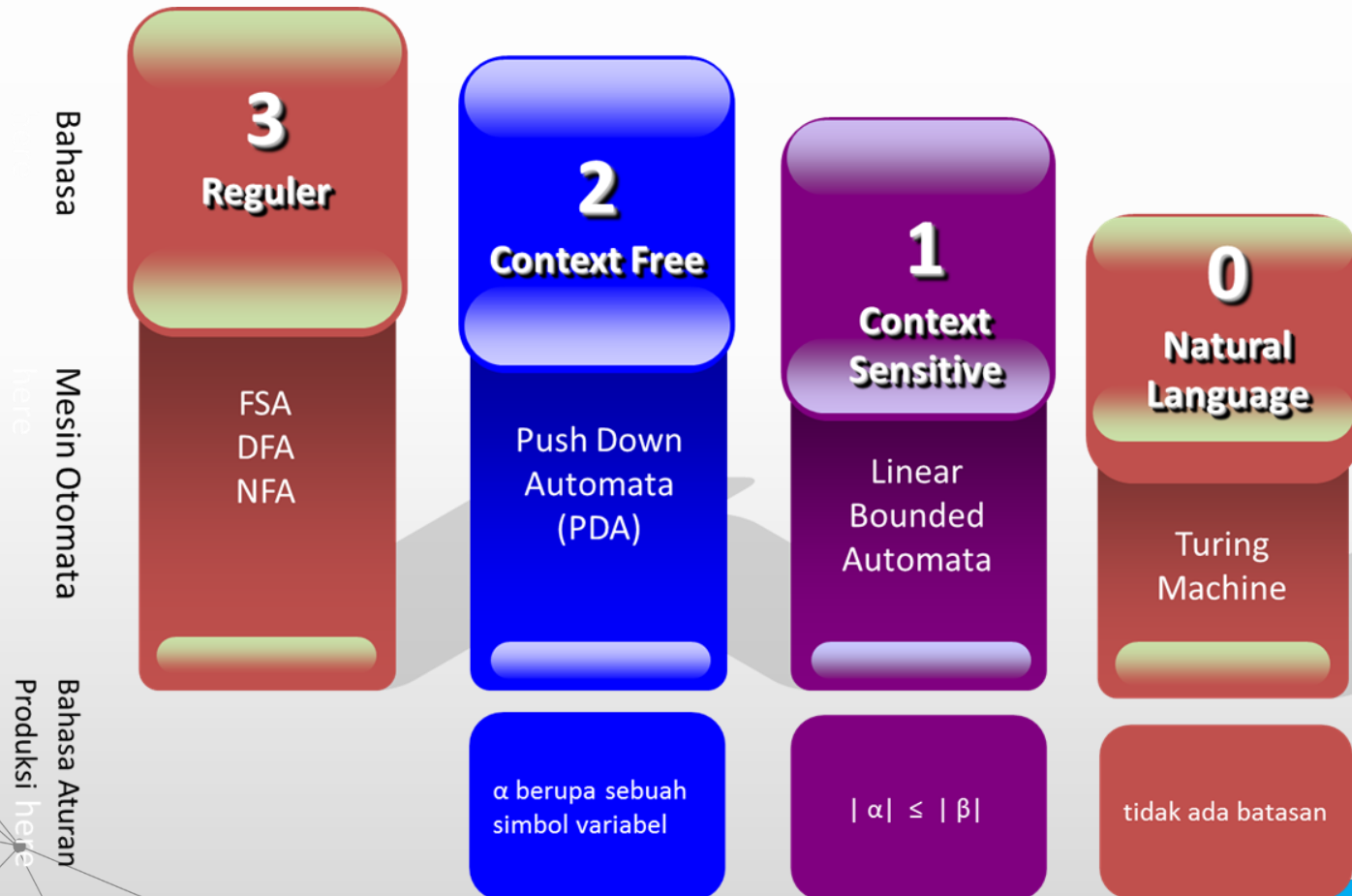
- Tata Bahasa/Grammar adalah sebagai kumpulan dari himpunan-himpunan variabel, simbol-simbol terminal, simbol awal, yang dibatasi oleh aturan-aturan produksi.
- Aturan produksi merupakan pusat dari grammar yang menspesifikasikan bagaimana suatu tata bahasa melakukan transformasi suatu string atau karakter ke bentuk lainnya.

Klasifikasi Tata Bahasa

- Pada tahun 1959 seorang ahli bernama Noam Chomsky melakukan penggolongan tingkatan bahasa menjadi empat, yang disebut dengan Hirarki Chomsky



Klasifikasi Tata Bahasa



Konsep Bahasa

- Abjad
- String
- Bahasa (Language)
- Bahasa Kosong
- Bahasa Universal dari Σ

Konsep Bahasa

Abjad

- Sebuah himpunan berhingga tak kosong dari simbol – simbol.
- Notasi : Σ
 - $\Sigma = \{a, b, c, \dots, x, y, z\}$ (Roman alphabet)
 - $\Sigma = \{0, 1, \dots, 9\}$
 - $\Sigma = \{0, 1\}$ (binary alphabet)

Konsep Bahasa

Bahasa (Language)

- Kumpulan dari string. Notasi : L
 - $\Sigma = \{1, 2, 3, 4, 5\}$,
 - Kumpulan $\{1, 12, 123, 1234, 12345, 112\}$ adalah bahasa dari abjad tersebut
 - $\Sigma = \{1\}$,
 - Kumpulan $\{1, 11, 111, 1111, \dots\}$ adalah sebuah bahasa.

Bahasa Kosong

- Suatu bahasa yang tidak terdiri dari string. Notasi : $\{\}$ atau \emptyset

Konsep Bahasa

Bahasa universal dari Σ

- Bahasa yang terdiri dari semua string berdasarkan suatu abjad Σ
- Notasi Σ^*
- $\Sigma = \{1\}$, maka $\Sigma^* = \{\epsilon, 1, 11, 111, 1111, \dots\}$

Konsep Bahasa

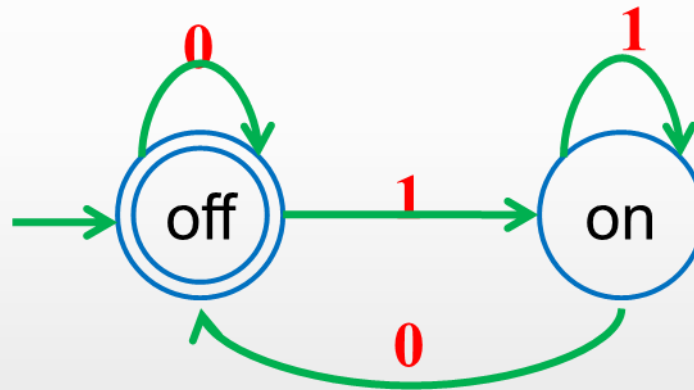
Operasi-operasi pada Bahasa

- Concatenation
- Eksponensial
- Union (Gabungan)
- Intersection (Irisan)
- Sub Set (Sub Bahasa)
- Equal
- Closure Language

Otomata Sederhana

■ Contoh mesin otomata sederhana :

■ Otomata pada saklar listrik

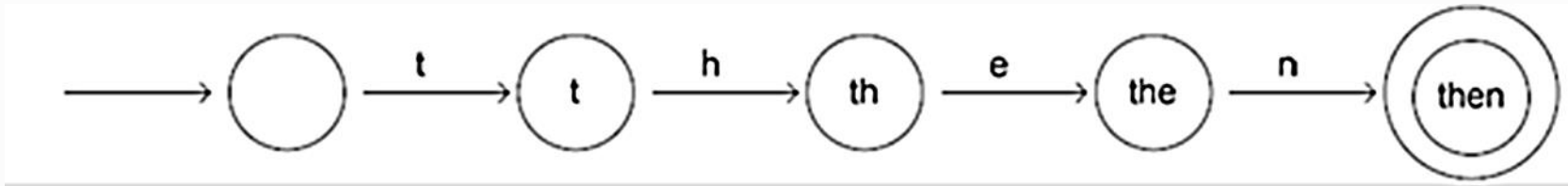


- Bila mesin mendapat string input :
- 0 → diterima
- 010 → diterima
-???

■ Sebuah string input diterima bila mencapai state akhir (final state) yang digambarkan dengan lingkaran ganda

Otomata Sederhana

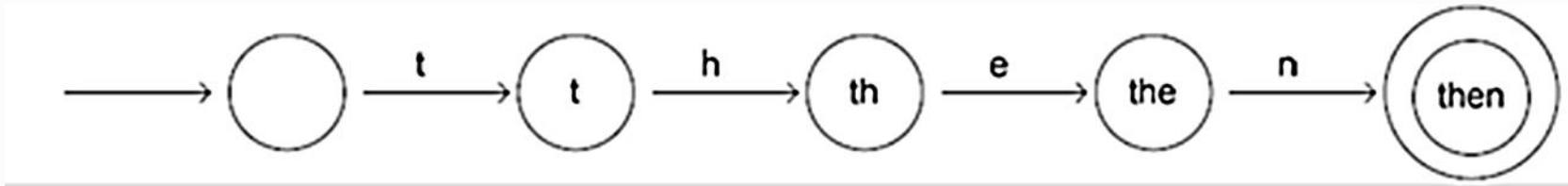
- Finite automaton berikut dapat dinyatakan sebagai bagian dari lexical analyzer



- Vending Machine

Otomata Sederhana

- Finite automaton berikut dapat dinyatakan sebagai bagian dari lexical analyzer



- Vending Machine

ATURAN PRODUKSI

- Semua aturan produksi dinyatakan dalam bentuk “ $\alpha \rightarrow \beta$ ” (bisa dibaca α menghasilkan β , atau dibaca α menurunkan β)
- α merupakan simbol-simbol pada ruas kiri aturan produksi, sedangkan β merupakan simbol-simbol ruas kanan aturan produksi

ATURAN PRODUKSI

- Simbol-simbol tersebut dapat berupa simbol terminal (V_t) dan simbol NON-Terminal (V_n)/Variabel.
- Simbol V_n adalah simbol yang masih dapat diturunkan, biasanya identik dengan huruf besar ('A','B','C')
- Simbol V_t adalah simbol yang sudah tidak dapat diturunkan lagi, biasanya identik dengan huruf kecil ('a','b','c')

ATURAN PRODUKSI

- Dengan menerapkan aturan produksi, suatu tata bahasa bisa menghasilkan sejumlah string.
- Contoh aturan produksi (Video 2.1)

$$E \rightarrow T \mid T+E \mid T * E$$
$$T \rightarrow a$$

ATURAN PRODUKSI

- Dari aturan produksi di atas, menghasilkan suatu variabel a atau variabel ekspresi $a+a$ atau $a*a$
- $E \rightarrow T$
 $T \rightarrow a$
- $E \rightarrow T+E$
 $E \rightarrow a+T$
 $E \rightarrow a+a$
- $E \rightarrow T * E$
 $E \rightarrow a * T$
 $E \rightarrow a * a$

Tata Bahasa dan Hirarki Chomsky

Tata bahasa/Grammar G didefinisikan sebagai pasangan 4 tuple : VT , VN , S , dan Q , dan dituliskan sebagai $G(VT, VN, S, Q)$, dimana :

VT : himpunan simbol-simbol terminal (atau himpunan token-token, atau alfabet)

VN : himpunan simbol-simbol non terminal

$S \in VN$: simbol awal (atau simbol start)

Q : himpunan produksi

Hierarki Chomsky (Video 2.2)

Bahasa	Mesin Otomata	Batasan Aturan Produksi
Reguler/RG/Tipe 3	Finite State Automata (FSA) meliputi Deterministic Finite Automata (DFA), & Non-deterministic Finite Automata (NFA).	α adalah sebuah simbol variable Nonterminal (VN) β <u>maks</u> memiliki sebuah simbol variabel Nonterminal (VN) yg bila ada terletak di posisi paling kanan
Bebas Konteks/CFG/Context Free/Tipe 2	Push Down Automata (PDA)	α berupa sebuah simbol variabel Non Terminal
Context Sensitive/CSG/Tipe 1	Linier Bounded Automata	$ \alpha \leq \beta $
Unrestricted/ Phase Structure/ Natural Language/Tipe 0	Mesin Turing	Tidak ada batasan

Hierarki Chomsky (Video 2.2)

Berdasarkan komposisi bentuk ruas kiri dan ruas kanan produksinya ($\alpha \rightarrow \beta$),

Noam Chomsky mengklasifikasikan 4 tipe tata bahasa :

- Tata bahasa tipe ke-0 : Unrestricted Grammar (UG)
- Ciri : $\alpha, \beta \in (VT \mid VN)^*$, $|\alpha| > 0$ *atau* $|\alpha| > |\beta|$
- Tata Bahasa tipe ke-1 : Context Sensitive Grammar (CSG)
- Ciri : $\alpha, \beta \in (VT \mid VN)^*$, $0 < |\alpha| \leq |\beta|$

Hierarki Chomsky (Video 2.2)

Berdasarkan komposisi bentuk ruas kiri dan ruas kanan produksinya ($\alpha \rightarrow \beta$),

Noam Chomsky mengklasifikasikan 4 tipe tata bahasa :

- Tata bahasa tipe ke-0 : Unrestricted Grammar (UG)
- Ciri : $\alpha, \beta \in (VT \mid VN)^*$, $|\alpha| > 0$ *atau* $|\alpha| > |\beta|$
- Tata Bahasa tipe ke-1 : Context Sensitive Grammar (CSG)
- Ciri : $\alpha, \beta \in (VT \mid VN)^*$, $0 < |\alpha| \leq |\beta|$

Hierarki Chomsky (Video 2.2)

- Tata Bahasa tipe ke-2 : Context Free Grammar (CFG)
Ciri : $\alpha \in VN$, $\beta \in (VT \mid VN)^*$
- Tata Bahasa tipe ke-3 : Regular Grammar (RG)
Ciri : $\alpha \in VN$, $\beta \in \{VT, VT VN\}$ atau $\alpha \in VN$, $\beta \in \{VT, VN VT\}$

Contoh Analisa Penentuan Type Tata Bahasa

- Grammar $G1$ dengan $Q1 = \{S \rightarrow aB, B \rightarrow bB, B \rightarrow b\}$. Ruas kiri semua produksinya terdiri dari sebuah VN maka $G1$ kemungkinan tipe CFG (Context Free Grammar) atau RG (Regular Grammar).
- Selanjutnya karena semua ruas kanannya terdiri dari sebuah VT atau string VT VN maka $G1$ adalah RG (Regular Grammar)

- Grammar G_2 dengan $Q_2 = \{S \rightarrow Ba, B \rightarrow Bb, B \rightarrow b\}$.
- Ruas kiri semua produksinya terdiri dari sebuah VN maka G_2 kemungkinan tipe CFG atau RG. Selanjutnya karena semua ruas kanannya terdiri dari sebuah VT atau string VN VT maka G_2 adalah CFG

- Grammar G_3 dengan $Q_3 = \{S \rightarrow Ba, B \rightarrow bB, B \rightarrow b\}$. Ruas kiri semua produksinya terdiri dari sebuah VN maka G_3 kemungkinan tipe CFG atau RG.
- Selanjutnya karena ruas kanannya mengandung string VT VN (yaitu bB) dan juga string VN VT (Ba) maka G_3 bukan RG, dengan kata lain G_3 adalah CFG

- Grammar G_4 dengan $Q_4 = \{S \rightarrow aAb, B \rightarrow aB\}$.
Ruas kiri semua produksinya terdiri dari sebuah VN maka G_4 kemungkinan tipe CFG atau RG.
- Selanjutnya karena ruas kanannya mengandung string yang panjangnya lebih dari 2 (yaitu aAb) maka G_4 bukan RG, dengan kata lain G_4 adalah CFG

- Grammar G_5 dengan $Q_5 = \{S \rightarrow aA, S \rightarrow aB, aAb \rightarrow aBCb\}$. Ruas kirinya mengandung string yang panjangnya lebih dari 1 (yaitu aAb) maka G_5 kemungkinan tipe CSG atau UG.
- Selanjutnya karena semua ruas kirinya lebih pendek atau sama dengan ruas kananya maka G_5 adalah CSG

- Grammar G_6 dengan $Q_6 = \{aS \rightarrow ab, SAc \rightarrow bc\}$. Ruas kirinya mengandung string yang panjangnya lebih dari 1 maka G_6 kemungkinan tipe CSG atau UG.
- Selanjutnya karena terdapat ruas kirinya yang lebih panjang daripada ruas kananya (yaitu SAc) maka G_6 adalah UG

Grammar dalam Bahasa (Video 2.3)

Contoh :

$G1 : VT = \{I, \text{Love}, \text{Miss}, \text{You}\}, V = \{S, A, B, C\}, P = \{S \rightarrow ABC, A \rightarrow I, B \rightarrow \text{Love} \mid \text{Miss}, C \rightarrow \text{You}\}$

$S \Rightarrow ABC$

$\Rightarrow I \text{ loveYou}$

$L(G1) = \{I \text{ love You}, I \text{ Miss You}\}$

TUGAS

(Penjelasan Cara Mengerjakan bisa dilihat pada Video 2.2 Tata Bahasa & Hirarki Chomsky)

LATIHAN

1. Tentukan apakah produksi-produksi berikut memenuhi aturan tata bahasa Regular.

- (a) $A \rightarrow b$
- (b) $B \rightarrow bdB$
- (c) $B \rightarrow C$
- (d) $B \rightarrow bC$
- (e) $B \rightarrow Ad$
- (f) $B \rightarrow bcdef$
- (g) $B \rightarrow bcdefG$
- (h) $A \rightarrow aSa$
- (i) $A \rightarrow aSS$
- (j) $A \rightarrow \varepsilon$
- (k) $Ad \rightarrow dB$

2. Tentukan apakah aturan produksi-produksi berikut memenuhi aturan tata bahasa bebas konteks.

- (a) $A \rightarrow aSa$
- (b) $A \rightarrow Ace$
- (c) $A \rightarrow ab$
- (d) $A \rightarrow \epsilon$
- (e) $B \rightarrow bcdef$
- (f) $B \rightarrow bcdefG$
- (g) $A \rightarrow aSa$
- (h) $A \rightarrow aSS$
- (i) $A \rightarrow BCDEF$
- (j) $Ad \rightarrow dB$
- (k) $A \rightarrow AAAAA$
- (l) $d \rightarrow A$

3. Tentukan apakah produksi-produksi berikut memenuhi aturan tata bahasa *context sensitive*.

- (a) $B \rightarrow bcdefG$
- (b) $A \rightarrow aSa$
- (c) $A \rightarrow aSS$
- (d) $A \rightarrow BCDEF$
- (e) $Ad \rightarrow dB$
- (f) $A \rightarrow \epsilon$
- (g) $AB \rightarrow \epsilon$
- (h) $ad \rightarrow b$
- (i) $ad \rightarrow \epsilon$
- (j) $abC \rightarrow DE$
- (k) $abcDef \rightarrow ghijkl$
- (l) $AB \rightarrow cde$
- (m) $AAA \rightarrow BBB$

4. Tentukan apakah produksi-produksi berikut memenuhi aturan tata bahasa *unrestricted*.

- (a) $A \rightarrow \epsilon$
- (b) $AB \rightarrow \epsilon$
- (c) $ad \rightarrow b$
- (d) $ad \rightarrow \epsilon$
- (e) $abC \rightarrow DE$
- (f) $AB \rightarrow cde$

TUGAS

- Kerjakan dengan tulis tangan masing-masing pada kertas polio bergaris
- Beri Nama, NIM, dan Kelas masing-masingdan beri tanda tangan masing-masing
- Tugas di kumpulkan dengan format FOTO
- Kumpulkan tugas melalui BOC mahasiwa
- Di BOC kumpulkan paling lambat Minggu tanggal 24 Maret 2024 pukul 23.00 WIB
- Tugas yang sudah dikerjakan dikertas polio dikumpulkan pada Dosen saat perkuliahan.



PROGRAM STUDI TEKNIK INFORMATIKA
UNIVERSITAS BUANA PERJUANGAN
KARAWANG

THANK YOU