



PROGRAM STUDI
TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS DIAN NUSWANTORO

MATA KULIAH
Otomata dan teori Bahasa

Pertemuan ke_6

HIRARKI CHOMSKY

Tim pengampu

2022

HIRARKI CHOMSKY

- Tata bahasa (*grammar*) bisa didefinisikan secara formal sebagai kumpulan dari himpunan-himpunan variabel, simbol-simbol terminal, simbol awal, yang dibatasi oleh aturan-aturan produksi.
- Pada tahun 1959, seorang ahli bernama Noam Chomsky melakukan penggolongan tingkatan bahasa menjadi empat, yang disebut dengan hirarki Chomsky.
- Penggolongan tersebut bisa dilihat pada tabel berikut.

HIRARKI CHOMSKY

Bahasa	Mesin Otomata	Batasan Aturan Produksi
Regular	<i>Finite State Automata (FSA)</i> meliputi <i>Deterministic Finite Automata (DFA)</i> & <i>Non Deterministic Finite Automata (NFA)</i>	α adalah sebuah simbol variabel. β maksimal memiliki sebuah simbol variabel yang bila ada terletak di posisi paling kanan
Bebas Konteks / <i>Context Free</i>	<i>Push Down Automata (PDA)</i>	α berupa sebuah simbol variabel
<i>Context Sensitive</i>	<i>Linier Bounded Automata</i>	$ \alpha \leq \beta $
<i>Unrestricted / Phase Structure / Natural Language</i>	Mesin Turing	Tidak ada batasan

Secara umum tata bahasa dirumuskan sebagai :

$\alpha \rightarrow \beta$, yang berarti α menghasilkan β atau α menurunkan β .

Di mana α menyatakan simbol-simbol pada ruas kiri aturan produksi (sebelah kiri tanda ' \rightarrow ') dan β menyatakan simbol-simbol pada ruas kanan aturan produksi (sebelah kanan tanda ' \rightarrow ')

HIRARKI CHOMSKY

- Aturan Produksi merupakan pusat dari tata bahasa yang menspesifikasikan bagaimana suatu tata bahasa melakukan transformasi suatu string ke bentuk lainnya, dan melalui aturan produksi didefinisikan suatu bahasa yang berhubungan dengan tata bahasa tersebut.
- Bentuk umum : $\alpha \rightarrow \beta$
- Dibaca : α memproduksi/menurunkan β

HIRARKI CHOMSKY

Simbol-simbol pada Aturan Produksi :

1. Simbol variabel / non terminal adalah simbol yang masih bisa diturunkan atau memproduksi, dinotasikan dengan huruf besar seperti A, B, C, dst.
2. Simbol terminal adalah simbol yang sudah tidak bisa diturunkan, dinotasikan dengan huruf kecil seperti a, b, c, dst atau angka (0,1, 2,..., 9)

Contoh Aturan Produksi (AP) :

Contoh :

- $A \rightarrow b$ (AP)
- $a \rightarrow B$ (bukan termasuk AP)
- $A \rightarrow B$ (AP)
- $B \rightarrow bC$ (AP)
- $b \rightarrow Bc$ (bukan termasuk AP)
- $A \rightarrow bcD$ (AP)
- $A \rightarrow bCD$ (AP)
- $Ab \rightarrow c$ (AP)

TATA BAHASA REGULER

Simbol pada Sebelah kiri harus berupa sebuah simbol variabel

Simbol pada sebelah kanan maksimal hanya memiliki sebuah simbol variabel dan bila ada terletak di posisi paling kanan.

Contoh :

- $A \rightarrow b$ (reguler)
- $a \rightarrow B$ (bukan reguler, karena simbol pada sebelah kiri harus berupa sebuah simbol variabel)
- $A \rightarrow B$ (reguler)
- $A \rightarrow bC$ (reguler)
- $A \rightarrow Bc$ (bukan reguler, karena simbol variabel pada sebelah kanan harus berada pada posisi paling kanan)
- $A \rightarrow bcD$ (reguler)
- $A \rightarrow bCD$ (bukan reguler, karena simbol pada sebelah kanan maksimal hanya memiliki sebuah simbol variabel)
- $Ab \rightarrow c$ (bukan reguler, karena simbol pada sebelah kiri harus berupa sebuah simbol variabel)

TATA BAHASA BEBAS KONTEKS

Simbol pada Sebelah kiri harus berupa sebuah simbol variabel

Contoh : $A \rightarrow b$ (reguler)

- $A \rightarrow B$ (reguler)
- $A \rightarrow bC$ (reguler)
- $A \rightarrow Bc$ (bebas konteks)
- $A \rightarrow BcD$ (bebas konteks)
- $A \rightarrow AAA$ (bebas konteks)
- $a \rightarrow b$ (Ditolak, karena simbol pada sebelah kiri harus berupa sebuah simbol variabel)
- $Ab \rightarrow c$ (bukan bebas konteks, karena simbol pada sebelah kiri harus berupa sebuah simbol variabel)
- $AB \rightarrow c$ (bukan bebas konteks, karena simbol pada sebelah kiri harus berupa sebuah simbol variabel)

TATA BAHASA CONTEX SENSITIVE

Simbol pada Sebelah kiri harus minimal ada sebuah simbol variabel
Jumlah simbol pada ruas sebelah kiri harus lebih kecil atau sama dengan jumlah simbol pada ruas kanan

Contoh : $A \rightarrow bc$ (reguler)

- $Ab \rightarrow cd$ (context sensitive)
- $AB \rightarrow CD$ (context sensitive)
- $ABC \rightarrow DE$ (bukan context sensitive, karena jumlah simbol pada ruas sebelah kiri lebih banyak dari jumlah simbol pada ruas kanan)
- $Ab \rightarrow cDe$ (context sensitive)
- $bA \rightarrow cd$ (context sensitive)
- $a \rightarrow b$ (Ditolak, karena simbol pada sebelah kiri harus minimal ada sebuah simbol variabel)

TATA BAHASA ALAMI

Simbol pada Sebelah kiri harus minimal ada sebuah simbol variabel

Contoh : $Abcdef \rightarrow g$ (Diterima)

- $aBCdE \rightarrow GHIJKL$ (Diterima)
- $abcdef \rightarrow GHIJKL$ (Ditolak, karena simbol pada sebelah kiri tidak ada sebuah simbol variabel)

TERIMA KASIH