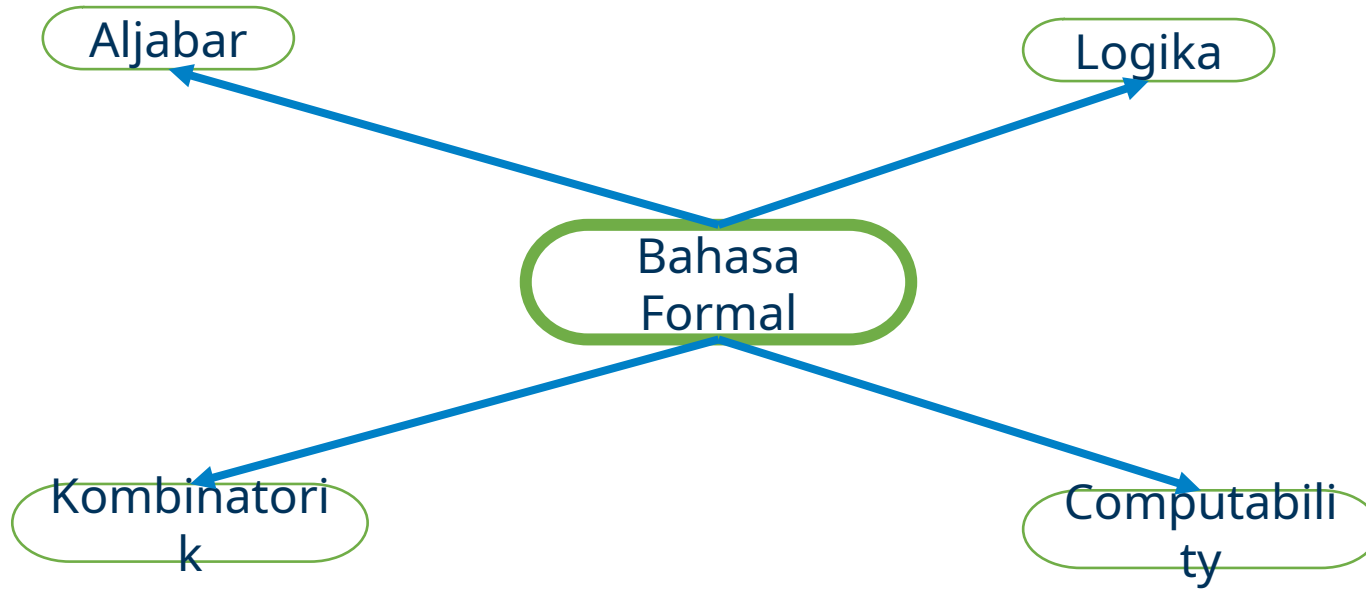




Kuliah 1

Pendahuluan

Pengertian Bahasa Formal, alfabet, dan Bahasa



Komponen Bahasa Formal

Pendahuluan



UKRIDA

- Bahasa adalah suatu sistem yang meliputi pengekpresian gagasan, fakta, konsep, termasuk sekumpulan simbol simbol dan aturan untuk melakukan manipulasinya.
- Bahasa bisa juga disebut sebagai rangkaian simbol simbol yang mempunyai makna.
- Bahasa formal adalah kumpulan kalimat dimana kalimat dalam sebuah bahasa dibangkitkan oleh sebuah tata bahasa (grammar) yang sama.

- Teori bahasa formal adalah bagian dari matematika diskrit yang memiliki koneksi dengan ilmu yang lain seperti aljabar, logika, kombinatorik, dan computability.
- Teori bahasa formal (*formal language*) merupakan landasan ilmu komputer teoritis. Namun, asal dan motivasinya berasal dari sumber yang berbeda:
 - a. Switching sirkuit sebagai model untuk teknik elektro
 - b. Grammar sebagai model untuk struktur bahasa alami
 - c. Model untuk fenomena biologis:
 - Jaringan syaraf yang mengarah pada otomata.
 - Lindenmayer sistem sebagai model untuk pertumbuhan organisme
 - d. Model dibagian yang berbeda dari teori bahasa pemrograman: parsing, kompilasi, pengeditan teks.
 - e. Model untuk pertanyaan matematis (dan filosofi) dari komputasi.
- Area aplikasi baru adalah kriptografi dan grafik komputer.

- Teori bahasa membicarakan bahasa formal (*formal language*) untuk kepentingan perancangan kompilator (*compiler*) dan pemroses naskah (*text processor*).
- Bahasa formal dibangkitkan oleh dua atau lebih tata bahasa yang berbeda.
- Dikatakan bahasa formal karena grammar diciptakan mendahului pembangkitan setiap kalimatnya.

- Bahasa manusia bersifat sebaliknya; grammar diciptakan untuk meresmikan kata kata yang hidup di masyarakat.
- Bidang teori bahasa formal mempelajari aspek aspek sintaksis murni bahasa bahasa yaitu pola struktur internal.
- Dalam bidang ilmu komputer, bahasa formal sering digunakan sebagai dasar untuk mendefenisikan bahasa pemrograman dan sistem lain dimana kata kata bahasa dikaitkan dengan makna atau semantik tertentu.

- Teori bahasa formal biasanya memperhatikan dirinya sendiri dengan bahasa formal yang dijelaskan oleh beberapa aturan sintaksis, seperti rangkaian string panjang hingga (mungkin tak terbatas).
- Dalam prakteknya ada banyak bahasa yang dapat dijelaskan oleh aturan seperti bahasa reguler atau bahasa bebas konteks.
- Gagasan tata bahasa formal mungkin lebih dekat dengan konsep intuitif “bahasa” yang dijelaskan oleh aturan aturan sintaksis.

- Sebuah alphabet dalam konteks bahasa formal dapat berupa apa saja, meskipun sering menggunakan alphabet sebagai set karakter seperti ASCII.
- Abjad juga bisa tidak terbatas; misalnya logika orde pertama sering diekspresikan menggunakan alphabet yang selain simbol seperti \wedge , \neg , tanda kurung, mengandung banyak elemen x_0, x_1, x_2, \dots yang memainkan peran variabel.
- Unsur unsur alphabet disebut huruf huruf.
- Abjad yang dilambangkan dengan simbol adalah himpunan tak kosong dari simbol simbol.

Contoh

- Alphabet biner adalah = $\{0,1\}$
- Alphabet huruf kecil adalah = $\{a, b, c, d, \dots, z\}$
- Alphabet bilangan asli < 9 adalah = $\{1, 2, 3, 4, 5, 6, 7, 8\}$

Untai atau kata (String)



UKRIDA

- Sebuah kata yang terbentuk dari kumpulan alphabet dapat berupa urutan, atau string, huruf apapun yang terbatas.
- Himpunan semua kata diatas alphabet Σ biasanya dilambangkan dengan Σ^* (menggunakan bintang Kleene)
- Untuk alphabet apapun khanya ada satu kata dengan panjang 0, kata kosong, yang sering dilambangkan dengan $\lambda, \epsilon, \Lambda, \Delta$.
- Dengan penggabungan satu atau dua kata untuk membentuk kata baru yang panjangnya adalah jumlah dari dari kata kata asli.
- Hasil penggabungan kata dengan kata kosong adalah kata aslinya.
- Untai kadang disebut kata atau word, adalah barisan berhingga simbol simbol yang berasal dari suatu alphabet.

Contoh

- 1011 adalah untaian yang berasal dari alfabet $\Sigma = \{0,1\}$
- kelas, jakarta, mobil adalah untaian yang berasal dari alfabet $\Sigma = \{a, b, c, \dots, z\}$
- 22244775 adalah untaian yang berasal dari alfabet $\Sigma = \{1, 2, 3, \dots, 9\}$

- Sebuah kata (string) adalah urutan item yang terbatas , yang disebut simbol atau huruf yang dipilih dari himpunan yang terbatas tertentu yang disebut alphabet.
- Contoh alphabet umum adalah huruf dalam alphabet (+ ruang interword, tanda baca, dll), bit 0 dan 1.
- Panjang kata w adalah jumlah simbol didalamnya, dilambangkan dengan $|w|$.
- Panjang kata kosong adalah 0

- Jika ada simbol k dalam alphabet, maka ada kata kata kn panjang n .
- Jadi ada:

$$\sum_{i=0}^n k^i = \frac{k^{n+1} - 1}{k - 1}$$

- Kata kata panjang paling banyak n , jika $k > 1$, dan $n+1$ kata jika $k=1$.
- Himpunan semua kata adalah tak terbatas, yaitu dapat diberikan sebagai daftar tak terbatas.

- Operasi dasar kata kata adalah gabungan, yaitu menulis kata sebagai suatu gabungan.
- Penggabungan kata $w1$ dan $w2$ dilambangkan hanya dengan $w1w2$.
- Contoh penggabungan dalam alphabet $\{a,b,c\}$:
 - ❖ $w1 = aacbba, w2 = caac, w1w2 = aacbbaacaac$
 - ❖ $w1 = aacbba, w2 = \Lambda, w1w2 = aacbba$
 - ❖ $w1 = \lambda, w2 = caac, w1w2 = caac$

- Penggabungan adalah asosiatif, yaitu $w_1(w_2w_3) = (w_1w_2)w_3$.
- Sebagai konsekuensinya penggabungan berulang dapat ditulis tanpa menggunakan tanda kurung.
- Disisi lain **penggabungan** biasanya **tidak komutatif**, $w_1w_2 \neq w_2w_1$, dalam kasus penggabungan aljabar uniter jelas komutatif.
- The n^{th} (*concatenation*) dari kata w adalah $w^n = ww \dots w \rightarrow n \text{ copy}$
- Terutama $w^1 = w$ dan $w^0 = \Lambda$, dan selalu $\Lambda^n = \Lambda$.
- The mirror image atau pembalikan dari kata $w = a_1, a_2, \dots, a_n$ adalah kata $\hat{w} = a_n, \dots, a_2, a_1$.

- **Untai kosong**

- Untai kosong adalah untaian yang tidak mempunyai simbol yang berasal dari alfabet.
- Untai kosong (null string) dilambangkan dengan λ , ϵ , Λ , Δ .

- **Panjang Untai**

- Panjang untaian adalah jumlah simbol yang membentuknya.

- **Contoh:**

- Panjang untaian 1011, ditulis $|1011| = 4$
- Panjang untaian sttrqw, ditulis $|sttrqw| = 6$
- Panjang untaian λ , ditulis $|\lambda| = 0$

- **Pangkat alphabet**

- Himpunan seluruh untaian yang berasal dari alphabet tertentu dapat dinyatakan dalam notasi eksponensial.
- Σ^k didefinisikan sebagai himpunan untaian dengan panjang k , yang masing masing simbolnya berasal dari Σ .

1. Σ^* (Kleen Star) didefinisikan sebagai himpunan seluruh untaian mulai dari untaian kosong sampai untaian panjang tertentu.

$$\Sigma^* = \Sigma^0 + \Sigma^1 + \Sigma^2 + \Sigma^3 + \dots = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \Sigma^3 \cup \dots$$

2. Σ^+ didefinisikan sebagai himpunan seluruh untaian tanpa untaian kosong (null string)

$$\Sigma^+ = \Sigma^1 + \Sigma^2 + \Sigma^3 + \dots = \Sigma^1 \cup \Sigma^2 \cup \Sigma^3 \cup \dots$$

• Contoh

Jika $\Sigma = \{0, 1\}$, maka :

$$\Sigma^0 = \{\lambda\}$$

$$\Sigma^1 = \{0, 1\}$$

$$\Sigma^2 = \{00, 01, 10, 11\}$$

$$\Sigma^3 = \{000, 001, 010, 011, 100, 101, 110, 111\}$$

$$\Sigma^* = \{\lambda, 0, 1, 00, 01, 10, 11, 000, \dots\}$$

$$\Sigma^+ = \{0, 1, 00, 01, 10, 11, 000, \dots\}$$

- Rentengan Untai (*string Concatenation*)

Misal w_1 dan w_2 adalah untaian. Rentengan untaian w_1 dan w_2 menghasilkan untaian $w_1 w_2$

Contoh.

➤ $w_1 = xx, w_2 = xyx$, maka $w_1 w_2 = xxxyx$

➤ $w_1 = aab, w_2 = abb$, maka $w_1 w_2 = aababb$

➤ $w_1 = \lambda, w_2 = xy$, maka $w_1 w_2 = xy$

➤ $w_1 = abba, w_2 = \lambda$, maka $w_1 w_2 = abba$

➤ $w_1 = \lambda, w_2 = \lambda$, maka $w_1 w_2 = \lambda$

Diberikan dua string : $x = abc$, dan $y = 123$

- Prefix string w adalah string yang dihasilkan dari string w dengan menghilangkan nol atau lebih simbol-simbol **paling belakang** dari string w tersebut.

Contoh: abc , ab , a , dan ϵ adalah semua $\text{prefix}(x)$

- Properprefix string w adalah string yang dihasilkan dari string w dengan menghilangkan satu atau lebih simbol-simbol **paling belakang** dari string w tersebut.

Contoh : ab , a , dan ϵ adalah semua $\text{Properprefix}(x)$

- Postfix atau sufix string w adalah string yang dihasilkan dari string w dengan menghilangkan nol atau lebih simbol-simbol **paling depan** dari string w tersebut.

Contoh : abc , bc , c , dan ϵ adalah semua $\text{prefix}(x)$

- **Properpostfix** atau Propersuffix string w adalah string yang dihasilkan dari string w dengan menghilangkan **satu** atau lebih simbol-simbol **paling depan** dari string w tersebut.

Contoh : bc , c , dan ϵ adalah semua $\text{Properpostfix}(x)$

Operasi dasar Untai (String)



UKRIDA

- Head string w adalah simbol paling depan dari string w

Contoh : a adalah $\text{Head}(x)$

- Tail string adalah string yang dihasilkan dari string w dengan menghilangkan simbol paling depan dari string w tersebut.

Contoh : bc adalah $\text{Tail}(x)$

- Substring string w adalah string yang dihasilkan dari string w dengan menghilangkan nol atau lebih simbol-simbol paling depan dan/atau simbol-simbol paling belakang dari string w tersebut.

Contoh : $abc, ab, bc, a, b, c,$ dan ϵ adalah semua $\text{Substring}(x)$

- Propersubstring string w adalah string yang dihasilkan dari string w dengan menghilangkan satu atau lebih simbol-simbol paling depan dan/atau simbol-simbol paling belakang dari string w tersebut.

Contoh : ab, bc, a, b, c , dan ϵ adalah semua Propersubstring(x)

- Subsequence string w adalah string yang dihasilkan dari string w dengan menghilangkan nol atau lebih simbol-simbol dari string w tersebut.

Contoh : abc, ab, bc, ac, a, b, c , dan ϵ adalah semua subsequence(x)

Operasi dasar Untai (String)



UKRIDA

- ProperSubsequence string w adalah string yang dihasilkan dari string w dengan menghilangkan satu atau lebih simbol-simbol dari string w tersebut.

Contoh : ab, bc, ac, a, b, c , dan ϵ adalah semua ProperSubsequence(x)

- **Concatenation** adalah penyambungan dua buah string. Operator concatenation adalah `concat` atau tanpa lambang apapun.

Contoh : `concat(xy)=xy=abc123`

- **Alternation** adalah pilihan satu diantara dua buah string. Operator alternation adalah `alternate` atau `|`.

Contoh : `alternate(xy)=x|y=abc atau 123`

- Bahasa adalah kumpulan kata kata dari beberapa alphabet.
- Seringkali bahasa tunggal $\{w\}$ diidentifikasi dengan satu satunya kata w , dan bahasa dilambangkan hanya oleh w .
- Seperangkat notasi digunakan untuk bahasa : \subseteq (inklusi), \subset (penyertaan yang benar), \cup (union), \cap (intersection), $-$ (difference) dan $-$ (pelengkap terhadap himpunan semua kata diatas alphabet).
- Milik kata w dalam bahasa L dilambangkan dengan $w \in L$.

- Bahasa semua kata diatas alphabet Σ , khususnya Λ , dilambangkan dengan Σ^* .
- Bahasa dari semua kata tidak kosong diatas alphabet Σ yang dilambangkan dengan Σ^+ .
- Jadi $L = \Sigma^* - L$ dan $\Sigma^+ - \{\Lambda\}$.
- Ada jumlah bahasa yang tidak terbatas pada alphabet manapun sehingga bahasa tidak dapat diberikan dalam daftar tak terbatas.

- Bahasa menurut kamus Websters adalah “ the body of word and methods of combining words used and understood by a considerable community”, sedangkan menurut Kamus Besar Bahasa Indonesia (KBBI), bahasa adalah:
 1. Sistem lambang bunyi yang arbitrer, yang digunakan oleh anggota suatu masyarakat untuk bekerjasama, berinteraksi, dan mengidentifikasikan diri.
 2. Percakapan (perkataan) yang baik, tingkah laku yang baik, sopan santun, perkataan perkataan yang dipakai oleh suatu bangsa.
- Definisi diatas mengungkapkan bahwa suatu bahasa adalah kalimat atau perkataan.

- Kalimat dalam sebuah bahasa dibentuk dengan menggabungkan satu atau lebih kata kata.
- Perhatian dalam pembentukan kalimat hanya tertuju pada sintaks atau bentuk kalimat, bukan pada semantik atau makna kalimatnya.
- Bahasa komunikasi yang digunakan oleh manusia, yaitu bahasa ucap seperti bahasa indonesia, Inggris, Jerman, dan sebagainya disebut sebagai bahasa alami atau bahasa natural (**Natural Language**).

- Sintaks bahasa alami sangat rumit dan tidak mungkin untuk menspesifikasikan semua aturan sintaksnya.
- Bahasa yang kaidah sintaksnya dapat dispesifikasikan secara matematis dengan baik disebut bahasa formal.
- Bahasa formal dapat didefinisikan secara abstrak sebagai “ sistem matematik”.
- Kaidah sintaks dalam teori bahasa formal tidak hanya bermanfaat bagi studi linguistik bahasa alami seperti penterjemahan secara otomatis dari suatu bahasa ke bahasa lain, tetapi juga berguna untuk studi bahasa pemrograman.

1. Tentukan nilai dari operasi berikut : $a + cb^*$

Jawab:

$$b^* = e, b, bb, bbb, bbbb, bbbbbb, \dots$$

$$cb^* = c, cb, cbb, cbbb, cbbbbb, cbbbbb, \dots$$

$$\text{Maka } a + cb^* = a, c, cb, cbb, cbbb, cbbbbb, cbbbbb, \dots$$

2. Tentukan nilai dari operasi berikut : $a(b + c)$

Jawab:

$$= a.b + a.c$$

$$= ab, ac$$

3. Tentukan nilai dari operasi berikut : $(ab)^*a$

Jawab:

$$(ab)^* = \varepsilon, ab, abab, ababab, \dots$$

$$(ab)^*a = a, aba, ababa, abababa, \dots$$

$$\text{Maka } (ab)^*a = a, aba, ababa, abababa, \dots$$

4. Tentukan nilai dari operasi berikut : $0.1^*(0 + 1)$

Jawab:

$$1^* = \varepsilon, 1, 11, 111, \dots$$

$$0.1^* = 01^* = 0, 01, 011, 0111, \dots$$

$$\text{Maka, } 0.1^*(0 + 1) = 00, 01, 010, 011, 0110, 0111, 01110, \dots$$

1. Tentukan nilai dari operasi berikut : $(a + b)^* a$

Jawab:

2. Tentukan nilai dari operasi berikut : $(a + c)b^*$

Jawab:

3. Tentukan nilai dari operasi berikut: $1(0 + 1)^* + 10^*$

Jawab:

jflap

Thank You

ukrida.ac.id



UKRIDA
Universitas Kristen Krida Wacana