Sifat-sifat Bahasa-bahasa Reguler: Teorema Pumping

Kuliah Teori Bahasa dan Automata
Program Studi Ilmu Komputer
Fasilkom UI

Prepared by:
Suryana Setiawan

Review Bahasa Reguler

- Definisi Bahasa Reguler
- Ekspresi Reguler
- Teorema Kleene:
  - Eskpresi Reguler → FSM
  - FSM → Ekspresi Reguler
- Kelas-kelas ekivalen dari relasi $\approx_L$ berjumlah berhingga.

Bahasa Reguler vs nonreguler

- Sifat-sifat tersebut bersifat positif dalam kelas bahasa regular.
- Untuk membedakan dari bahasa non-regular?
  - TIDAK dapat memenuhi definisi bahasa regular.
  - TIDAK dapat dituliskan ekspresi regulernya.
  - TIDAK dapat dibuatkan FSM untuk mengenalinya.
  - TIDAK dapat ditunjukkan keberhinggan jumlah kelas ekivalensi.
- Kalau kita TIDAK berhasil menemukan suatu FSM atau ekspresi regeuler untuk suatu bahasa, belum tentu bahasa itu non-regular.
- Ketidakberhinggaan pada banyaknya kelas ekivalensi tidak mungkin dibuktikan.

Penggunaan Sifat Untuk Non-reguler

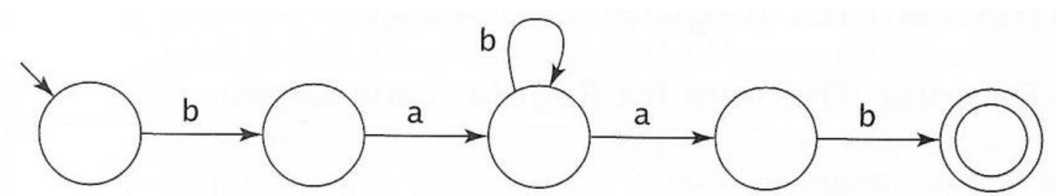
- Bagaimana dengan Bahasa-Bahasa ini?
  - $A^nB^n = \{a^nb^n : n \geq 0\}$
  - Palindrom dengan $\Sigma = \{a, b\}$
  - $L_{AB} = \{w \in \{a, b\}^* : \forall v \in \text{prefix}(w).(\#_a(v) \geq \#_b(v))\}$
- secara intuitif bisa ditunjukkan mesin untuk mengenal A^nB^n memerlukan jumlah status tak berhingga, tapi tidak cukup untuk mengatakan tidak ada FSM.
- Palindrom sudah dikenal sebagai bahasa non-reguler tetapi kita tidak bisa menunjukkan Palindrom tidak memenuhi sifat-sifat tsb.
- Untuk L_{AB} bahkan mungkin kita ragu-ragu, karena intuisi kita tidak jalan.
- Kita masih memerlukan alat pembuktian lain bahwa suatu bahasa non-regular itu benar-benar non-regular!

Rencana Materi

- Sifat atau alat pembuktian itu adalah Teorema Pumping (dikenal juga dengan nama Pumping Lemma).
- Teorema Pumping adalah sifat dari bahasa regular tetapi penggunaannya adalah untuk memeriksa suatu Bahasa non-regular adalah Bahasa non-regular dengan cara proof-by-contradiction,
- Untuk suatu bahasa L:
  - Dengan asumsi L adalah reguler akan ditunjukkan terjadi suatu kontradiksi yang berarti L tidak memenuhi sifat ini, berarti juga bukan bahasa Reguler.
  - Sementara jika L adalah Bahasa regular, maka asumsi itu selalu mengarah pada pemenuhan Teorema Pumping.

Observasi Bahasa Reguler

- Setiap bahasa reguler L dapat diterima oleh suatu DFSM M (Ingat: $|K|$ **berhingga**). Jika L adalah **tak berhingga** maka
  - Pasti sekurangnya terdapat satu loop di dalam M (pigeonhole principle).
- Setiap $w \in L$, $|w| \geq |K|$, memiliki satu atau lebih pola berulang karena dengan adanya substring y ($y \neq \varepsilon$, dan $w = xy^qz$, $q \geq 0$) yang membawa M masuk dalam loop.
  - *Contoh:* $w = babab$ dengan $x = ba$, $y = b$, $z = ab$



Observasi Bahasa Nonreguler

- Observasi dalam slide sebelumnya tidak selalu dapat digunakan untuk memeriksa suatu bahasa L yang sembarang (yang tidak diketahui apakah ada FSM untuk menerimanya).
  - Kalau pasti ada FSM, pasti bahasa reguler!
  - Kalau pasti tidak ada, pasti bukan bahasa reguler!
- Jika tidak pasti? Asumsikan terdapat suatu FSM *M* (tapi entah seperti apa!) dengan jumlah status *k*,
  - Jika observasi itu berlaku, maka itu bahasa reguler.
  - Jika observasi tidak berlaku, bukan bahasa reguler.

Teorema Pumping

- Jika L adalah bahasa reguler, jika dan hanya jika:
  $\exists k \geq 1$ ($\forall w \in L$. ($|w| \geq k$ ($\exists x, y, z$. ($w = xyz \wedge |xy| \leq k \wedge y \neq \varepsilon \wedge \forall q \geq 0$. ($xy^qz \in L$)))))

- Supaya mudah dibaca/dipahami, ditulis ulang sbb:

$\exists k \geq 1$ yang mana (
  $\forall w \in L$, yang mana $|w| \geq k$ (
    $\exists x, y, z$ yang mana (
      $w = xyz$, $|xy| \leq k$, $y \neq \varepsilon$, dan,
      $\forall q \geq 0$ yang mana ($xy^qz \in L$)
)
)
)
)

Batasan-Batasan Teorema Pumping

- Dari observasi sebelumnya dengan mesin dengan K
  - $k = |K|$, string w dengan panjang k atau lebih,
- Terdapat suatu partisi w menjadi xyz
  - y adalah substring yang membawa M melalui siklus,
  - x prefiks dan z sufiks dari w, yang mengapit y.
- Walaupun $|xy| \leq k$
  - karena tepat ada k simbol, jika $|w| = k$ maka siklus dilalui minimal satu kali; apalagi jika jika $|xy| > k$.
- Asalkan $y \neq \varepsilon$
  - karena M deterministik, tidak ada siklus ε
- Maka berlaku $\forall q \geq 0. (xy^q z \in L)$.

Negasi Teorema Pumping

- "*L* bahasa regular *iff* terdapat **suatu** *k* …. **untuk setiap** *w* ….**ada suatu** partisi *x*, *y*, *z* ….. **untuk setiap** *q* ….sehingga sifat pumping **terpenuhi**. "
- **Negasinya**: "*L* bukan reguler *karena* **untuk setiap** *k* …. **ada suatu** *w* …. **untuk semua** partisi *x*, *y*, *z*, ….. **ada suatu** *q*….. dimana sifat pumping **gagal terpenuhi**."

$\forall k \geq 1$ yang mana (
$\quad \exists w \in L$, yang mana $|w| \geq k$ dan
$\quad\quad \forall x, y, z$ yang mana (
$\quad\quad\quad w = xyz$, $|xy| \leq k$, $y \neq \varepsilon$, dan,
$\quad\quad\quad \exists q \geq 0$ yang mana ($xy^q z \notin L$)
$\quad\quad$)
$\quad$)
)
)

Contoh: Bahasa A^nB^n

- Diketahui L adalah $A^nB^n = \{\mathrm{a}^n\mathrm{b}^n : n \geq 0\}$.
  - "Apabila L reguler harusnya terdapat k sehingga setiap string w, dengan $|w| \geq k$, memenuhi kondisi-kondisi dalam teorema pumping."
  - Kontradiksi ditunjukkan dengan memilih suatu w agar gagal dalam pumping!
- Untuk $w = \mathrm{a}^k\mathrm{b}^k$, $|w| = 2k$, yang memenuhi syarat $|w| \geq k$
  - Adakah x, y, dan z dengan $w = xyz$, dimana $|xy| \leq k$, $y \neq \varepsilon$, dan, $\forall q \geq 0$ selalu berlaku $xy^qz \in L$?

Contoh: Bahasa A^nB^n (lanjutan)

- Partisi x, y, dan z dengan $w = xyz$, dimana $|xy| \le k$, $y \ne \varepsilon$
  - Secara umum $xy = \mathrm{a}^j$ dengan $1 \le j \le k$ dan $y = \mathrm{a}^p$, dengan $1 \le p \le j$, sehingga $x = \mathrm{a}^{j-p}$ dan $z = \mathrm{a}^{k-j}\mathrm{b}^k$

k $\qquad\qquad$ k

$w =$ a . . aa. . . . aa. . . aabbb. . . . . bbb

x $\quad$ y $\qquad\qquad$ z

a^{j-p} $\quad$ a^{p} $\qquad\qquad$ $\mathrm{a}^{k-j}\,\mathrm{b}^{k}$

- Selanjutnya, $xy^qz = \mathrm{a}^{j-p}(\mathrm{a}^p)^q\mathrm{a}^{k-j}\mathrm{b}^k = \mathrm{a}^{p.q}\mathrm{a}^{k-p}\mathrm{b}^k$
  - kita pilih $q = 0$, $xy^0z = \mathrm{a}^{p.0}\,\mathrm{a}^{k-p}\mathrm{b}^k = \mathrm{a}^{k-p}\mathrm{b}^k$ yaitu suatu string yang bukan anggota L → **suatu kontradiksi**.

Contoh: Bahasa A^nB^n (coba w lain)

- Mencoba w = $a^{\lceil k/2 \rceil} b^{\lceil k/2 \rceil}$ dengan $|w| = k$ jika k bil genap atau $|w| = k+1$ jika k bil ganjil.
- Jika $|y| = p$, berarti y dapat berupa a^p atau $a^{p-j}b^j$ atau b^p
  - Kasus $y = a^p$, dengan $q = 2$, menghasilkan $a^{\lceil k/2 \rceil - p} (a^p)^2 b^{\lceil k/2 \rceil} = a^{\lceil k/2 \rceil + p} b^{\lceil k/2 \rceil} \notin L$
  - Kasus $y = a^{p-j}b^j$, dengan $q = 2$, menghasilkan $a^{\lceil k/2 \rceil - p + j} (a^{p-j}b^j)^2 b^{\lceil k/2 \rceil} = a^{\lceil k/2 \rceil} b^j a^{p-j} b^{\lceil k/2 \rceil} \notin L$
  - Kasus $y = b^p$, dengan $q = 2$, menghasilkan $a^{\lceil k/2 \rceil} (b^p)^2 b^{\lceil k/2 \rceil - p} = a^{\lceil k/2 \rceil} b^{\lceil k/2 \rceil + p} \notin L$
- Mencoba berbagai w akan tetap menghasilkan kontradiksi!