MODUL 8: PUMPUNG LEMMA UNTUK BAHASA REGULER

Dengan definisi bahasa regular atau teorema Kleene atau keberhinggaan kelas-kelas ekivalensinya sebenarnya sudah cukup untuk memeriksa apakah suatu bahasa merupakan bahasa regular atau bukan. Berikut ini dijelaskan suatu lemma yang juga merupakan sifat khas dari bahasa Regular yang juga dapat digunakan untuk memeriksa apakah suatu bahasa merupakan bahasa regular atau bukan. Lemma ini perlu dijelaskan karena dapat diapdatasi untuk pemeriksaan bahasa-bahasa lainnya yang lebih umum.

Jika $M = (Q, \Sigma, q_0, A, d)$ suatu FA yang mengenali bahasa L. Bila M tidak memiliki siklus maka dengan sendirinya L merupakan himpunan berhingga string dan persoalannya menjadi mudah.

Sebaiknya kasus yang menarik adalah jika memiliki siklus tersebut. Terdapat suatu string $x = a_1a_2\ldots a_ny \in L$, sehingga M bertransisi dari q_0 ke $q_n \in A$, sbb.

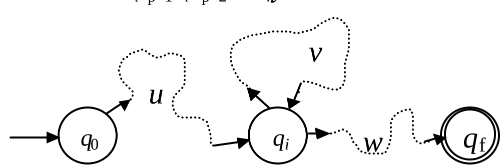
$$q_0 = d^*(q_0, \Lambda),\ \ q_1 = d^*(q_0, a_1),$$
$$q_2 = d^*(q_0, a_1a_2),\ \ \ldots\ \ , q_n = d^*(q_0, a_1a_2\ldots a_n).$$

Dari q_0 sampai dengan q_n terdapat n+1 status sementara bila Q berisikan n status maka menurut prinsip Pigeonhole akan terdapat sekurangnya satu status dalam Q yang muncul dua kali dalam deretan q_i di atas. Misalkan status tersebut adalah q_i yang sama dengan q_{i+p}, di mana $0 \le i < i+p \le n$. maka

$$q_i = d^*(q_0, a_1a_2\ldots a_i) = d^*(q_i, a_{i+1}a_{i+2}\ldots a_{i+p})$$

Untuk menyederhanakan notasi kita tulis $x = uvw$ dan

$$u = a_1a_2\ldots a_i$$
$$v = a_{i+1}a_{i+2}\ldots a_{i+p}$$
$$w = a_{i+p+1}a_{i+p+2}\ldots a_ny$$



Dengan demikian selain $x = uvw$ dikenali oleh M, string-string dalam bentuk uv^mw dengan $m \ge 0$ dapat dikenali oleh M. Ini dinyatakan sebagai teorema sbb.

Teorema Keberulangan: jika L bahasa regular yang dikenali oleh FA dengan n status. Untuk $x \in L$ dengan $|x| \ge n$, dan x dapat dituliskan sebagai $x = uvw$ sedemikian rupa dimana

$$|uv| \le n$$
$$|v| > 0$$

dan untuk sembarang harga $m \ge 0$, $uv^mw \in L$.

Karena m bisa berapa saja berarti juga L merupakan himpunan tak berhingga. Sifat dalam teorema ini belum dapat digunakan untuk memeriksa apakah suatu bahasa sembarang L merupakan regular atau tidak. Dari teorema ini dikembangkan teorema yang pernyataannya lebih "weak" sbb.

Pumping Lemma: Jika L bahasa regular himpunan tak berhingga maka terdapat bilangan bulat n sehingga untuk setiap $x \in L$ dengan $|x| \ge n$, terdapat string u, v, dan w sehingga

$$uvw = x$$
$$|uv| \le n$$
$$|v| > 0$$

dan untuk sembarang harga $m \ge 0$, $uv^mw \in L$.

Apa buktinya? Menurut teorema Kleene jika L suatu bahasa regular maka terdapat FA yang dapat menerimanya. Jika n dalam pumping lemma ini adalah jumlah status FA tersebut maka teorema sebelumnya berlaku.

Pumping lemma ini dapat dipergunakan untuk membuktikan apakah suatu bahasa L merupakan non regular. Pembuktian dengan kontradiksi yaitu pertama dengan mengasumsikan bahwa L memenuhi sifat tersebut lalu menjabarkannya sehingga menghasilkan suatu yang kontradiktif.

Contoh: bahasa $L = \{0^i1^i \mid i \geq 0\}$. Misalkan L adalah regular dan n adalah bilangan bulat yang memenuhi teorema di atas. Kita pilih $x = 0^n1^n$ sehingga $|x| \geq n$. Sekarang kita mencari u, v, dan w yang memenuhi. Menurut teorema haruslah $|uv| \leq n$, maka $uv = 0^k$ dengan $k \leq n$. Menurut teorema haruslah $|v| > 0$, maka $v = 0^j$ dengan $j > 0$. Dengan $uv=0^k$, maka $w = 0^{n-k}1^n$.

Sekarang kita periksa sifat terakhir yang menyatakan bahwa $uv^mw \in L$. Karena $uv^mw = uvv^{m-1}w = 0^k(0^j)^{m-1}0^{n-k}1^n = 0^{k+j(m-1)+n-k}1^n = 0^{n+j(m-1)}1^n$ maka ternyata $uv^mw \notin L$ untuk $m > 1$. Suatu konradiksi.

Bagaimana kalau kita menggunakan x yang lain yang lebih umum? Misalnya dengan $x = 0^t1^t$ di mana $t \geq n/2$ maka ada tiga kemungkinan harga v yaitu berisi hanya berisi simbol 0 ($v = 0^j$), atau berisi simbol 0 dan 1 ($v = 0^i1^j$), atau hanya berisi simbol 1 ($v = 1^j$). Yang pertama telah dijelaskan di atas. Yang ketiga juga dengan cara yang mirip seperti di atas. Yang kedua $uv^mw = 0^{n-i}(0^i1^j)^m1^{n-j} = 0^{n-i}0^i1^j0^i1^j \ldots 0^i1^j1^{n-j} \notin L$ untuk $m \geq 1$.

Kemungkinan apapun yang kita coba akan selalu membawa kepada kontradiksi. Maka ternyata L bukan bahasa regular karena tidak memenuhi teorema Pumping Lema tsb.

Contoh: $L = \{x \in \{0, 1\}^* \mid x$ berisikan jumlah yang sama dari 0 dan 1$\}$.
Kita mencoba dengan memilih $x = 0^n1^n$. dan menentukan $uv = 0^k$, $v = 0^j$ untuk $k \leq n$, dan $j > 0$ sehingga untuk $m \neq 1$ maka terjadi suatu kontradiksi bahwa $uv^mw \notin L$.

Bagaimana jika digunakan $x = (01)^n$? Untuk harga x itu terdapat sejulah kemungkinan harga v sebagai berikut ini.

1. $v = (01)^j$ untuk sejumlah $j > 0$.
2. $v = 1(01)^j$ untuk sejumlah $j \geq 0$.
3. $v = 1(01)^j0$ untuk sejumlah $j \geq 0$.
4. $v = (01)^j0$ untuk sejumlah $j \geq 0$.

Untuk kemungkinan 2 dan 4 kita dapat menunjukkan kontradiksi, namun untuk 1 dan 2 tidak. Mengapa? Karena $(01)^*$ merupakan subset dari L yang dapat kita ketahui merupakan bahasa regular. Jadi disini pemilihan x perlu hati-hati untuk <u>tidak mengggunakan kasus khusus yang termasuk ke dalam kelas bahasa regular</u>.

Contoh: $L = \{0^iy \mid i \geq 0, y \in \{0, 1\}^*$ dan $|y| \leq i\}$. Jika digunakan string x yang hanya string berisi 0 maka kontradiksi tidak dapat ditunjukkan (uv^mw hanya menghasilkan string-string berisi 0 yang lebih panjang yang masih $\in L$).
Mengapa? Karena 0^* adalah bahasa regular.
Sekarang kita coba $x = 0^n1^n$ dan $v = 0^j$. Ternyata $uv^mw \in L$ untuk semua $m > 0$. Apakah ia bahasa regular? Bukan karena untuk $m = 0$ maka $uv^0w \notin L$. Suatu kontradiksi.

Kadang-kadang kita mungkin menggunakan bentuk Pumping Lemma yang lebih lemah. Dengan pumping lemma diperlemah pembuktian menjadi lebih sederhana karena sifat-sifat yang diperiksa lebih sedikit, namun belum tentu mampu mmeriksa setiap nonregular. Berikut ini dinyatakan kembali Pumping Lemma yang diperlemah dalam dua versi.

Pumping Lemma Diperlemah: Seandainya L adalah bahasa regular impunan tak berhingga, maka akan terdapat string u, v, dan w dimana $|v| > 0$ sehingga $uv^mw \in L$ untuk setiap $m \geq 0$.

Perbedaan dengan Pumping lemma sebelumnya adalah disini berapa besarnya batas bawah panjang x (harga n) yang digunakan tidak penting. Pumping lemma ini dapat digunakan untuk memeriksa $\{0^i1^i \mid i \geq 0\}$ tetapi tidak dapat digunakan untuk $\{x \in \{0, 1\}^* \mid x$ berisikan 0 dan 1 dalam jumlah yang sama$\}$ mau pun L = $\{0^iy \mid i \geq 0, y \in \{0, 1\}^*$ dan $|y| \leq i\}$.

Pumping Lemma Lebih Diperlemah: Seandainya L adalah bahasa regular himpunan tak berhingga, maka akan terdapat bilangan bulat p, dan q dengan $q > 0$ sehingga untuk setiap $m \geq 0$, L berisi suatu string dengan panjang $p + mq$.

Dengan kata lain, himpunan bilangan bulat

$$lengths(L) = \{|x| \mid x \in L\}$$

adalah deret hitung $p + mq$ dimana $m \geq 0$.

Contoh $L = \{0^n \mid n$ bilangan prima$\} = 0^2, 0^3, 0^5, 0^7, 0^{11}, \ldots\}$
Untuk L ini *lengths*(L) = himpunan semua bilangan prima. Sementara deret hitung berbentuk $p+mq$ bisa juga memiliki harga nonprima maka L nonregular. Misalnya jika $m = p+2q+2$ maka $p+mq = p+ (p+2q+2)q = (p+2q) + (p+2q)q = (p+2q)(1+q)$ memiliki faktor-faktor $(p+2q)$ dan $(1+q)$. Sehingga $p+mq$ bukan bilangan prima.

Catatan yang penting dipahami disini adalah bahwa keberhinggaan kelas-kelas ekivalensi perlu dan cukup untuk pemeriksaan apakah suatu bahasa adalah regular. Teorema keberulangan menunjukkan kondisi-kondisi yang perlu dan cukup untuk suatu bahasa dapat dikatakan regular. Sementara Pumping Lemma hanya menyebutkan kondisi-kondisi yang perlu saja (belum tentu cukup) untuk memperlihatkan bahwa suatu bahasa adalah regular. Jadi bisa terjadi walaupun kondisi-kondisi dalam pumping lemma terpenuhi dengan pemilihan x dan v yang benar ternyata bahasa ybs adalah non regular.

Contoh: $L = \{a^i b^j c^j \mid i \geq 1$ dan $j \geq 0\} \cup \{b^i c^k \mid j, k \geq 0\}$. Untuk $n = 1$ akan dilihat $x \in L$ dan $|x| \geq n$. Terdapat dua kasus. Pertama jika $x = a^i b^j c^j$ dengan $i > 0$, maka dengan $u = \Lambda$, $v = a$, $w = a^{i-1} b^j c^j$, setiap string $uv^m w$ masih $\in \{a^i b^j c^j \mid i \geq 1$ dan $j \geq 0\}$. Kedua, jika $x = b^i c^j$ maka dengan $u = \Lambda$, v berisi simbol pertama dalam x dan menghasilkan setiap string $uv^m w$ masih $\in \{b^i c^k \mid j, k \geq 0\}$. Akan tetapi L tidak regular seperti yang dapat diketahui melalui keberhinggan kelas-kelas ekivalensi.