[Aula 06] Linguagem regular: Expressão regular

Prof. João F. Mari
joaof.mari@ufv.br

ROTEIRO

- Expressão regular (ER)
- **[EX]** Expressão regular
- Expressão regular → Linguagem regular
- **[EX]** AFNε a partir de a*(aa + bb)

Expressão regular (ER)

- Toda linguagem regular pode ser descrita por uma:
  - Expressão Regular.
- Formalismo denotacional (ou **gerador**).
- Definida a partir de:
  - Conjuntos básicos (linguagens);
  - Operações de concatenação e união.
- Adequadas para a comunicação:
  - Humano × humano;
  - Humano × máquina.

Expressão regular (ER)

- Base de Indução:
  - ∅ é ER
    - denota a linguagem vazia: ∅
  - ε é ER
    - denota a linguagem { ε }
  - x é ER (para qualquer x ∈ Σ)
    - denota a linguagem { x }
- Passo de Indução: se r e s são ER e denotam as ling. R e S, então:
  - União: (r + s) é ER
    - Denota a linguagem **R U S**
  - Concatenação: (rs) é ER
    - Denota a linguagem **R S = { uv | u ∈ R e v ∈ S }**
  - Concatenação Sucessiva: (r*) é ER
    - Denota a linguagem **R***

Expressão regular (ER)

- Se **r** é ER, a linguagem gerada é dita:
  - Linguagem Gerada por **r**:

L(r) ou **GERA(r)**

- Omissão de parênteses em uma ER é usual:
  - **Concatenação sucessiva**:
    - Precedência sobre **concatenação** e **união.**
  - **Concatenação**:
    - Precedência sobre **união.**

[EX] Expressão regular

ER	Linguagem Gerada
aa	somente a palavra aa
ba*	todas as palavras que iniciam por b, seguido por zero ou mais a
(a + b)*	todas as palavras sobre { a, b }
(a + b)*aa(a + b)*	todas as palavras contendo aa como subpalavra
a*ba*ba*	todas as palavras contendo exatamente dois b
(a + b)*(aa + bb)	todas as palavras que terminam com aa ou bb
(a + ε)(b + ba)*	todas as palavras que não possuem dois a consecutivos

[EX] Expressão regular

- Linguagem gerada pela ER: **(a + b)*(aa + bb)**
  - a e b denotam { a } e { b }, respectivamente
  - a + b denota { a } ∪ { b } = { a, b }
  - (a + b)* denota { a, b }*
  - aa e bb denotam { a } { a } = { aa } e { b } { b } = { bb }

- Respectivamente:
  - (aa + bb) denota { aa } ∪ { bb } = { aa, bb }
  - (a + b)*(aa + bb) denota { a, b }* { aa, bb }

- Portanto, GERA((a + b)*(aa + bb)) é:

{ aa, bb, aaa, abb, baa, bbb, aaaa, aabb, abaa, abbb, baaa, babb, bbaa, bbbb,... }

  - Toda palavra sobre o alfabeto {a, b} cujo sufixo é aa ou bb.

Expressão Regular → Linguagem Regular

- Se **r** é **ER**, então GERA(r) é uma **linguagem regular (LR).**

- Prova (por indução):
- Uma linguagem **é regular** sse é possível construir um:
  - AFD, AFN ou AFNε que reconheça a linguagem.

- Para isso , é necessário mostrar que:
  - Dada uma ER **r** qualquer...
  - é possível construir um autômato finito **M** tal que

ACEITA(M) = GERA(r)

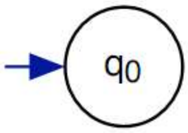
- Demonstração: indução no número de operadores.

Expressão regular → Linguagem regular

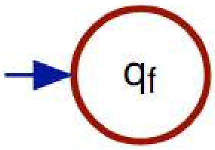
- Base de indução:
  - Seja **r** uma ER com zero operadores.
    - **r** = ∅
      - Autômato???
    - **r** = ε
      - Autômato???
    - **r** = x (x ∈ Σ)
      - Autômato???

Expressão regular → Linguagem regular

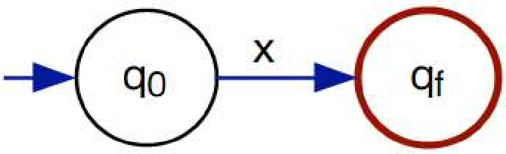
- Base de indução:
  - r é uma ER com zero operadores:
    - r = ∅. Autômato: $M_1 = (\emptyset, \{ q_0 \}, \delta_1, q_0, \emptyset)$



    - r = ε. Autômato: $M_2 = (\emptyset, \{ q_f \}, \delta_2, q_f, \{ q_f \})$



    - r = x (x ∈ Σ). Autômato: $M_3 = (\{ x \}, \{ q_0, q_f \}, \delta_3, q_0, \{ q_f \})$



Expressão regular → Linguagem regular

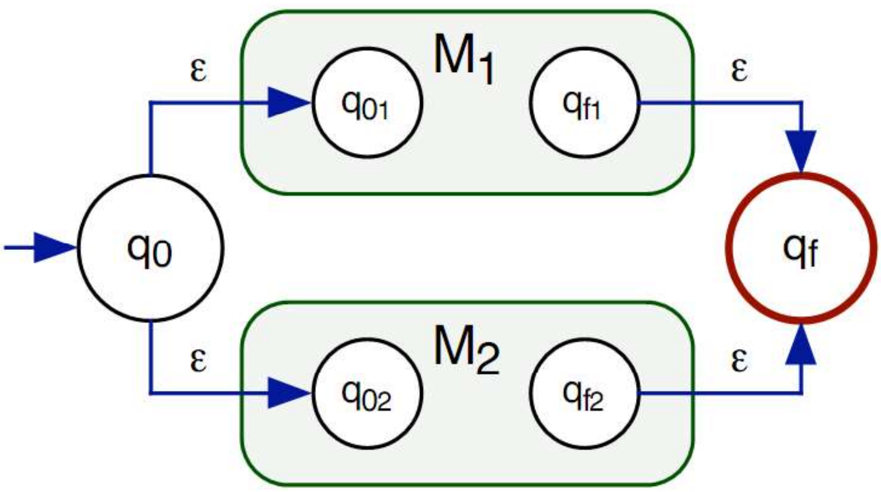
- Hipótese de Indução:
  - **r** é uma **ER** com até n operadores (n > 0);
  - Suponha que é possível definir um **AF** que aceita GERA(r).

- Passo de Indução:
  - **r** é uma ER com n + 1 operadores;
  - **r** pode ser representada por (**r_1** e **r_2** possuem conjuntamente no máximo n operadores)
    - **$r = r_1 + r_2$**
    - **$r = r_1 r_2$**
    - **$r = r_1{}^*$**
  - Por hipótese de indução, existem:

$$M_1 = (\Sigma_1, Q_1, \delta_1, q_{01}, \{ q_{f1} \}) \rightarrow ACEITA(M_1) = GERA(r_1)$$
$$M_2 = (\Sigma_2, Q_2, \delta_2, q_{02}, \{ q_{f2} \}) \rightarrow ACEITA(M_2) = GERA(r_2)$$

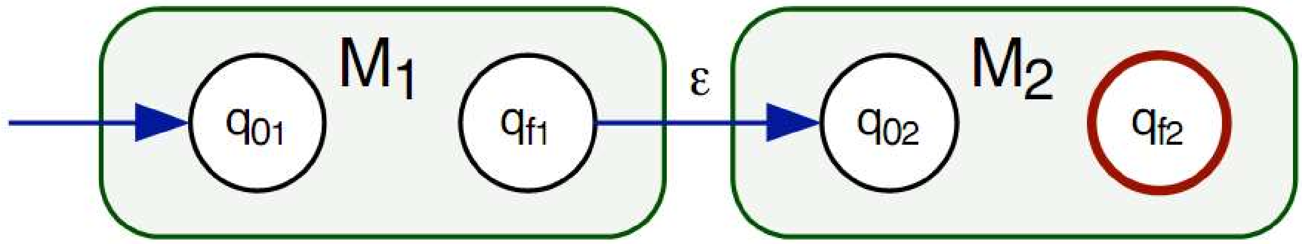
Expressão regular → Linguagem regular

- $r = r_1 + r_2$
  - Autômato $M = (\Sigma_1 \cup \Sigma_2, Q_1 \cup Q_2 \cup \{ q_0, q_f \}, \delta, q_0, \{ q_f \})$



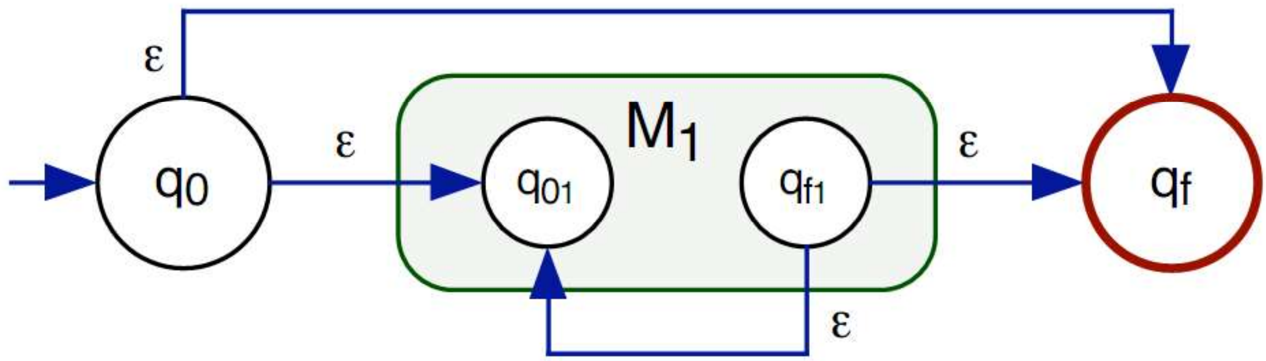
Expressão regular → Linguagem regular

- $r = r_1 r_2$
  - Autômato $M = (\Sigma_1 \cup \Sigma_2, Q_1 \cup Q_2, \delta, q_{01}, \{ q_{f2} \})$

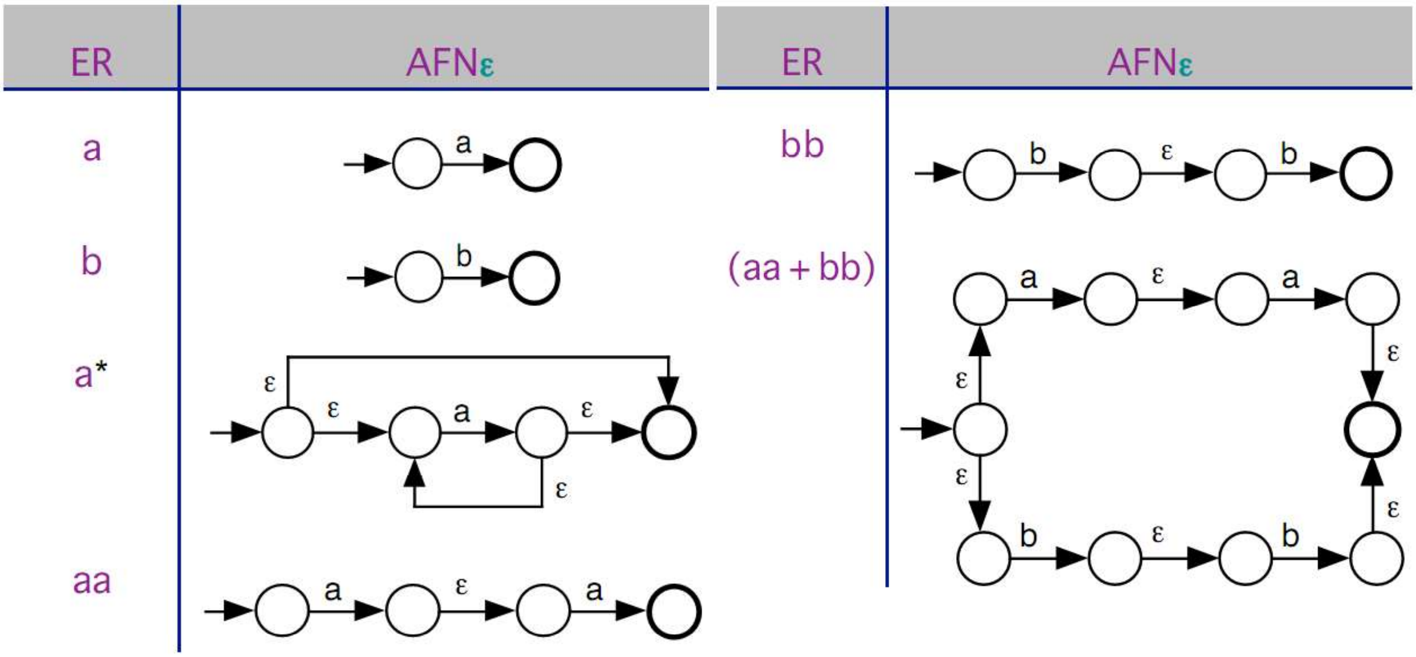


Expressão regular → Linguagem regular

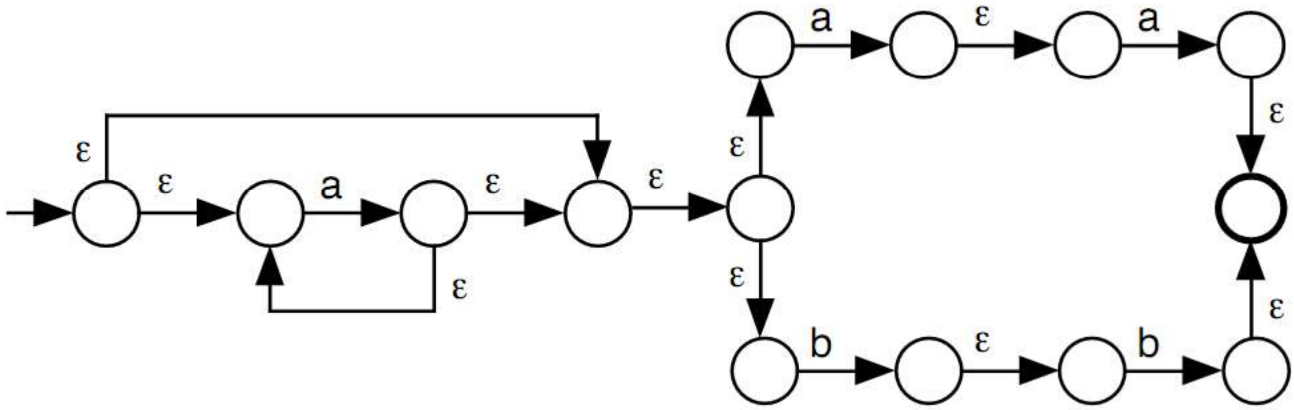
- $r = r_1^*$
  - Autômato (suponha $q_0 \notin Q_1$, $q_f \notin Q_1$)
  - $M = (\Sigma_1, Q_1 \cup \{ q_0, q_f \}, \delta, q_0, \{ q_f \})$



[EX] AFNε a partir de a*(aa + bb)



[EX] AFNε a partir de a*(aa + bb)



Linguagem Regular → Expressão Regular

- Se L é linguagem regular, então existe uma ER r tal que:

GERA(r) = L

BIBLIOGRAFIA

- MENEZES, P. B. **Linguagens formais e autômatos**, 6. ed., Bookman, 2011.
  - Capítulo 3.
  - + Slides disponibilizados pelo autor do livro.

[FIM]

- FIM:
  - **[AULA 06]** LINGUAGENS REGULARES – Expressão regular
- Próxima aula:
  - **[AULA 07]** LINGUAGENS REGULARES – Gramática regular