Expressões regulares

Introdução

- Expressões aritméticas: (5 + 3) x 4.
  - Resultado: 32.
- Expressões regulares: (0 U 1)0*.
  - Todas as palavras que iniciam com 0 ou 1 seguido de 0's.
- Assim como suprimimos x: (5 + 3) x 4 = (5 + 3)4), também suprimimos o: ((0 U 1)o0* = (0 U 1)0*.
- Onde??? Buscas textuais, comandos AWK e GREP.

Exemplo

- (0 U 1)*: todas as palavras possíveis com 0s e 1s.
- $\Sigma^{*}1$: todas as palavras que terminam com 1.
- $(0\,\Sigma^{*}) \cup (\Sigma^{*}\,1)$: todas as palavras que começam com 0 ou terminam com 1.
- A concatenação sucessiva (*) precede a concatenação (o) que precede a união (U).
- O parêntese modifica precedência.

Definição

- R é uma expressão regular se R é:
  - Qualquer a pertencente a um alfabeto Σ, é uma ER com linguagem possuindo a palavra unitária a, {a}.
  - ε é uma ER com uma palavra vazia,
  - $\emptyset$ é uma ER de linguagem vazia,
  - Se R e S são ER e denotam a linguagem R e S, então:
    - (R+S) é ER e denota a linguagem R U S
    - (RS) é ER e denota a linguagem RS = {uv | u $\in$ R e v $\in$ S} (RoS)
    - (R*) é ER e denota a linguagem R*.

Exemplo com alfabeto {0,1}

- 0*10* = {w | w tem exatamente um 1}.
- $\Sigma^* 1 \Sigma^*$ = {w | w tem ao menos um 1}.
- $\Sigma^* 001 \Sigma^*$ = {w | w contém 001 como subpalavra}.
- $(\Sigma\Sigma)^*$ = {w | w tem tamanho par}.
- $(\Sigma\Sigma\Sigma)^*$ = {w | w tem tamanho múltiplo de 3}.
- 01 U 10 = {01, 10}.
- $0\Sigma^*0 \cup 1\Sigma^*1 \cup 0 \cup 1$ = {w | w começa e termina com o mesmo símbolo}.

Entendendo...

- $R \cup \varepsilon = R$
  - Mas se $R = 0, L(R) = \{0\}$ e $L(R \cup \varepsilon) = \{0, \varepsilon\}$
- $R o \emptyset = R$
  - Mas se $R = 0, L(R) = \{0\}$ e $L(R o \emptyset) = \{\emptyset\}$
- Em compiladores de linguagens de programação:
  - {+,-, ε}(DD* U DD*.D* U D*.DDD*)

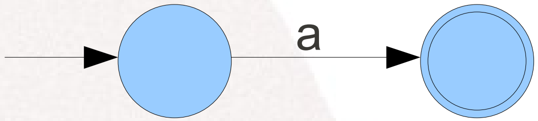
    D = {0,1,2,3,...,9}.
  - Sintax dos tokens feita, pode-se gerar o analisador léxico.

Equivalência com autômatos finitos

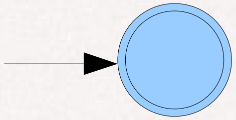
- Poder de descrição.
- Qualquer ER pode-se converter em um AF que reconheça a linguagem descrita e vice-versa.
- Teorema: Uma linguagem é regular sse alguma ER a descreve.
  - Lema 1: Se a linguagem é descrita por uma ER, então ela é regular.
  - Lema 2: Se a linguagem é regular, então há uma ER que a descreva.

Prova: Lema 1

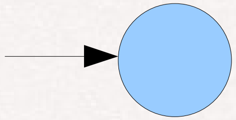
- Tome uma ER 'R' que descreve uma linguagem A; convertendo R em um AFND:
  - 1 - R = a em algum alfabeto.
    - L(R) = {a}.



  - 2 - R = ε
    - L(R) = {ε}



  - 3 - R = $\emptyset$
    - L(R) = {$\emptyset$}



  - 4 – R = R1 U R2
  - 5 – R = R1 o R2
  - 6 – R = R1*

Prova: Lema 1 - Exemplo

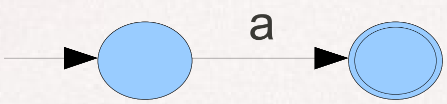
- Converta a ER (ab U a)* para AFND
  - Qual a linguagen aceita?

Prova: Lema 1 - Exemplo

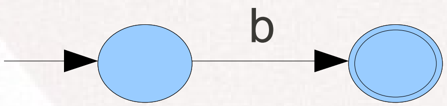
- Converta a ER (ab U a)* para AFND
  - Qual a linguagen aceita? Sucessões de ab ou a.
  - a
  - b
  - ab
  - ab U a
  - (ab U a)*

Prova: Lema 1 - Exemplo

- Converta a ER (ab U a)* para AFND
  - Qual a linguagen aceita? Sucessões de ab ou a.
  - a



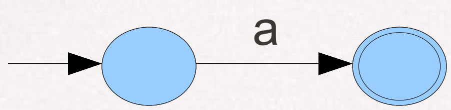
  - b



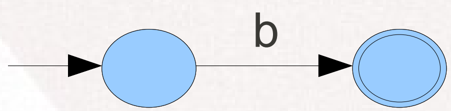
  - ab
  - ab U a
  - (ab U a)*

Prova: Lema 1 - Exemplo

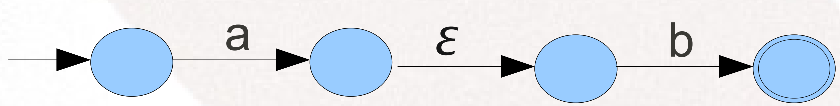
- Converta a ER (ab U a)* para AFND
  - Qual a linguagen aceita? Sucessões de ab ou a.
  - a



  - b



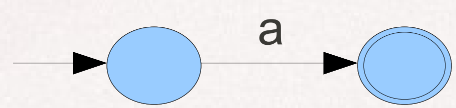
  - ab



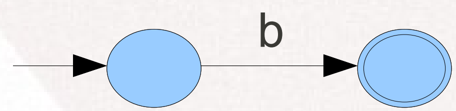
  - ab U a
  - (ab U a)*

Prova: Lema 1 - Exemplo

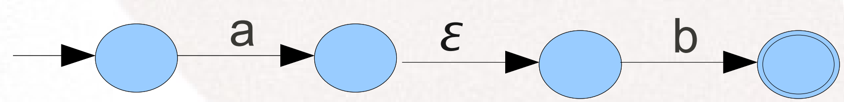
- Converta a ER (ab U a)* para AFND
  - Qual a linguagen aceita? Sucessões de ab ou a.
  - a



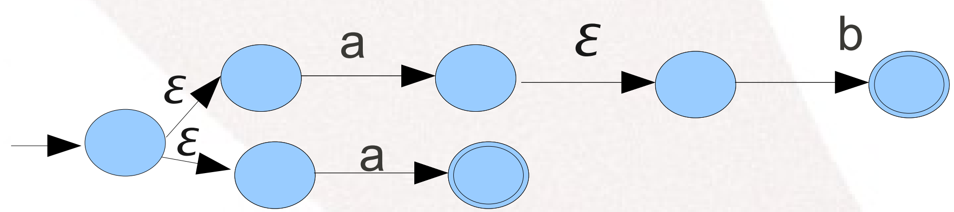
  - b



  - ab



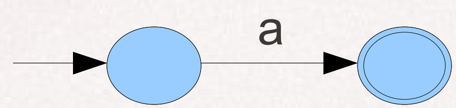
  - ab U a



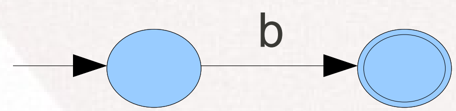
  - (ab U a)*

Prova: Lema 1 - Exemplo

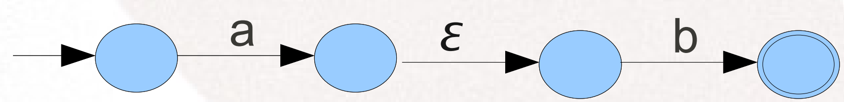
- Converta a ER (ab U a)* para AFND
  - Qual a linguagen aceita? Sucessões de ab ou a.
  - a



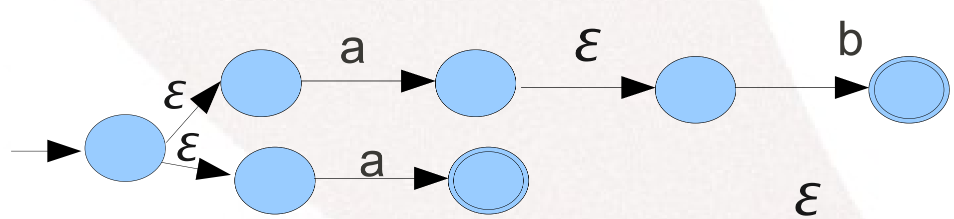
  - b



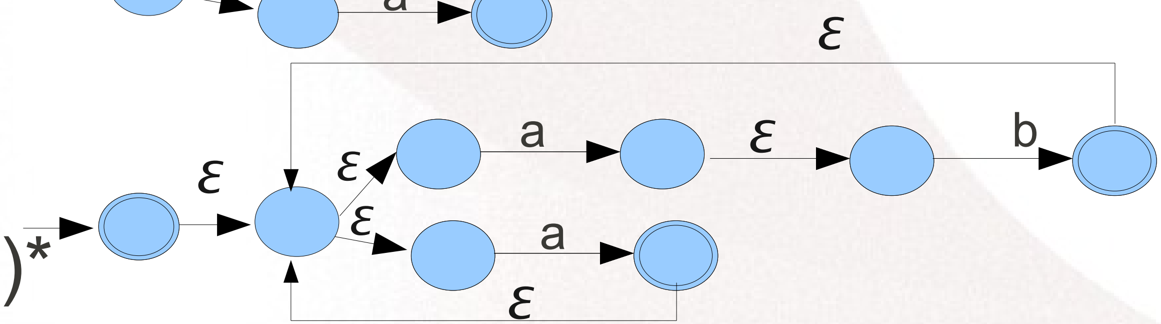
  - ab



  - ab U a



  - (ab U a)*



Prova: Lema 1 - Exemplo

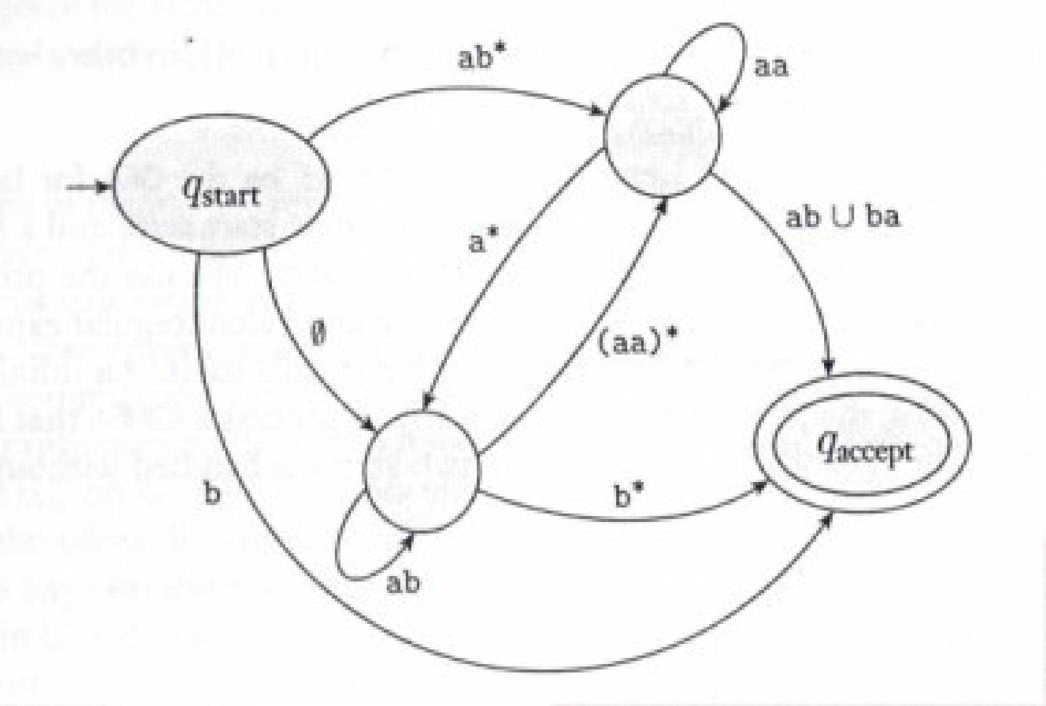
- Converta a ER (a U b)*aba pra AFND.

Prova: Lema 2

- Converta um AFD em ER equivalentes.
  - Autômato finito não-determinístico generalizado (AFNDG).
  - Converter AFD em AFNDG em ER.
  - AFG: são AFND que possuem em suas transições ER ao invés de apenas símbolos do alfabeto ou ε.
  - Como um AFND pode ter várias possibilidades para uma mesma entrada.

Prova: Lema 2

- Converta um AFD em ER equivalentes.



Prova: Lema 2

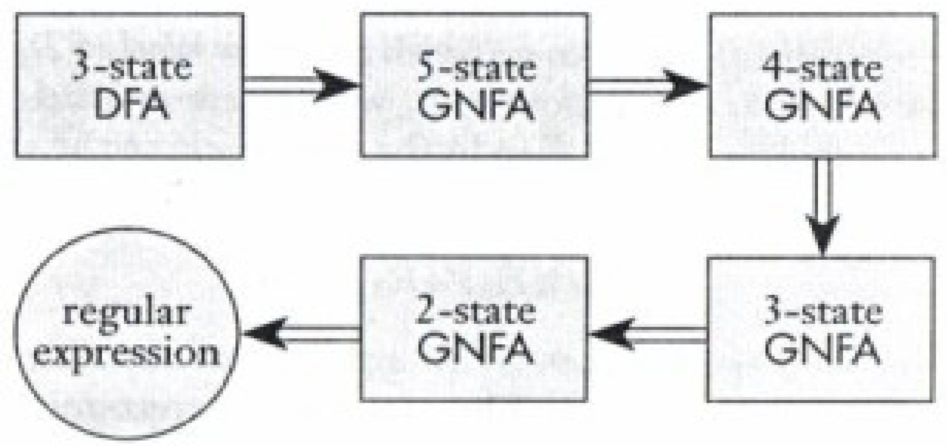
- Por conveniência um AFG deve:
  - O estado inicial tem arcos para todos os outros estados possíveis e não recebe nenhum.
  - Há somente um estado final aceito que recebe arcos de todos os estados e de onde não saem arcos. Estado inicial não pode ser final.
  - Dos estados remanescentes saem arcos em todas as direções, inclusive recursivamente.

Prova: Lema 2

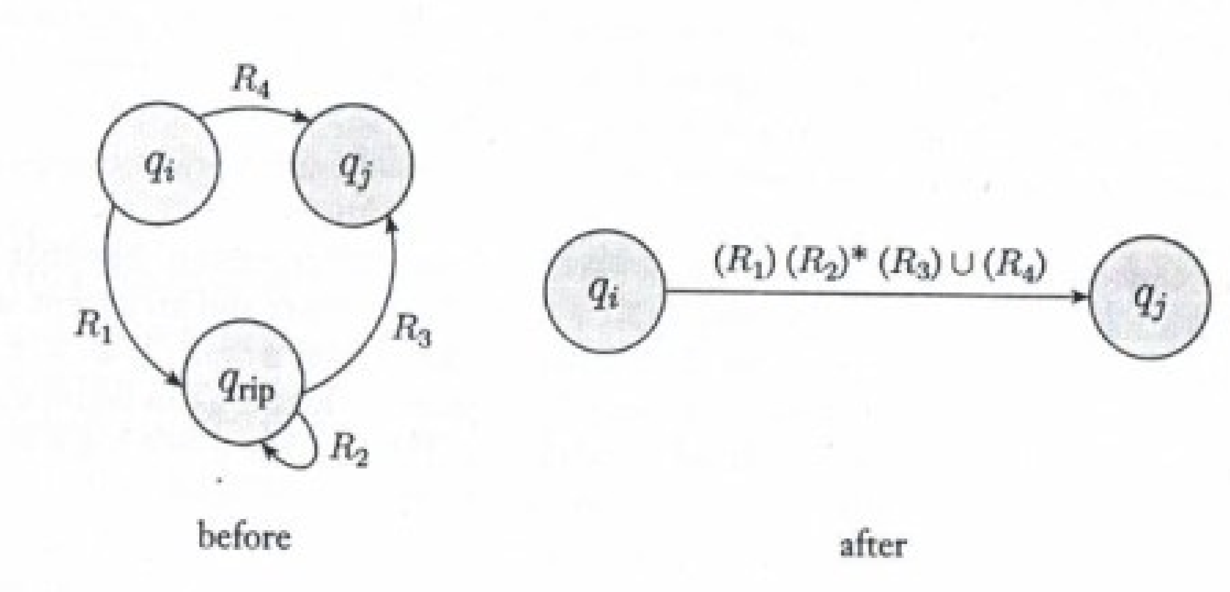
- AFD para AFG:
  - Adicione um estado inicial com uma transição vazia para o antigo q0 e um estado aceito sendo alcançado por transições vazias pelos antigos qf's.
  - Substitua arcos com múltiplas etiquetas pela união delas (vários arcos entre 2 estados, no mesmo sentido).
  - Adicione $\emptyset$ entre os estados que não se conectam. Isso não muda o ACEITA(A) porque essa transição nunca é usada.

Prova: Lema 2

- Converta AFG em ER:
  - O AFG tem pelo menos k=2 estados.
  - Se k>2, reduza o AFG para k-1 estados.
  - Quando k=2, há uma transição entre q0 e qf com a ER desejada.

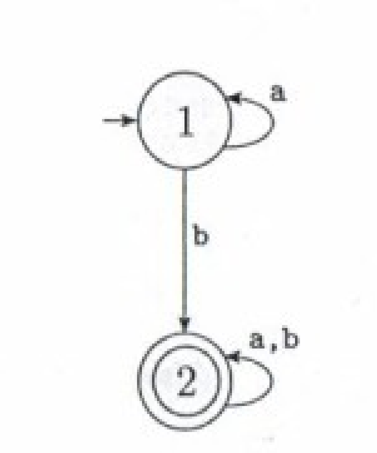


Prova: Lema 2



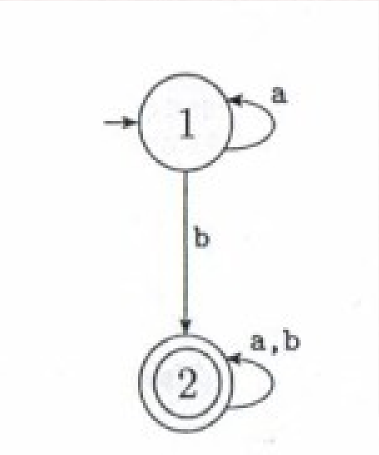
Prova: Lema 2 - Exemplo

- Tome o AFD e converta-o para ER.



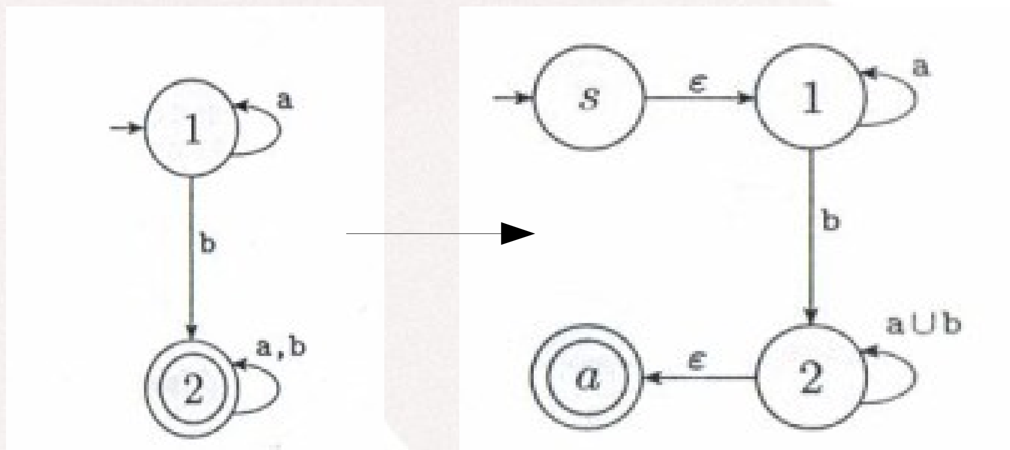
Prova: Lema 2 - Exemplo

- Generalizando o AFD:



Prova: Lema 2 - Exemplo

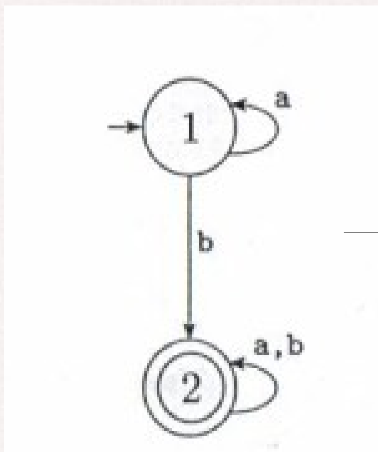
- Generalizando o AFD:

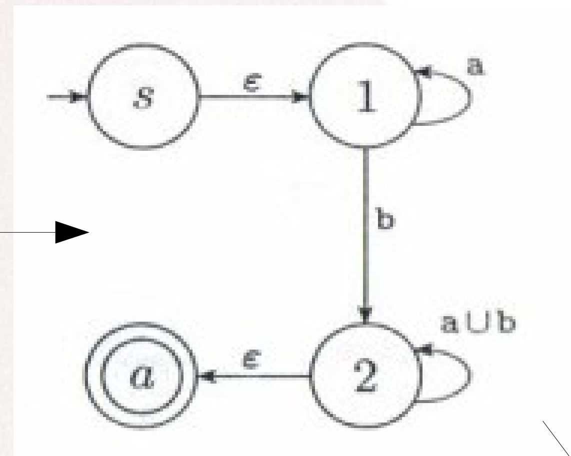


- Etiquetas ∅ foram suprimidas para melhor visualização (ligam estados que não se comunicam).

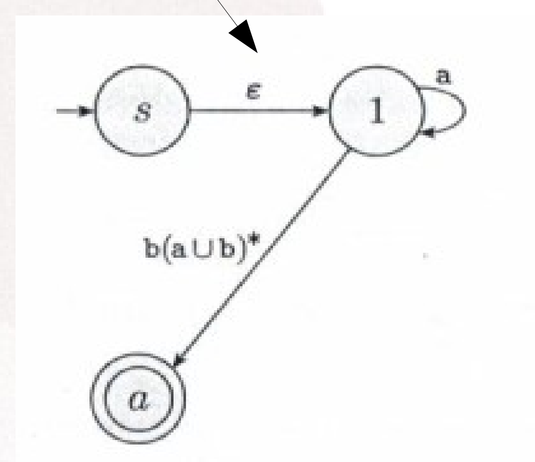
Prova: Lema 2 - Exemplo

- k=4, remova estados até k = 2.



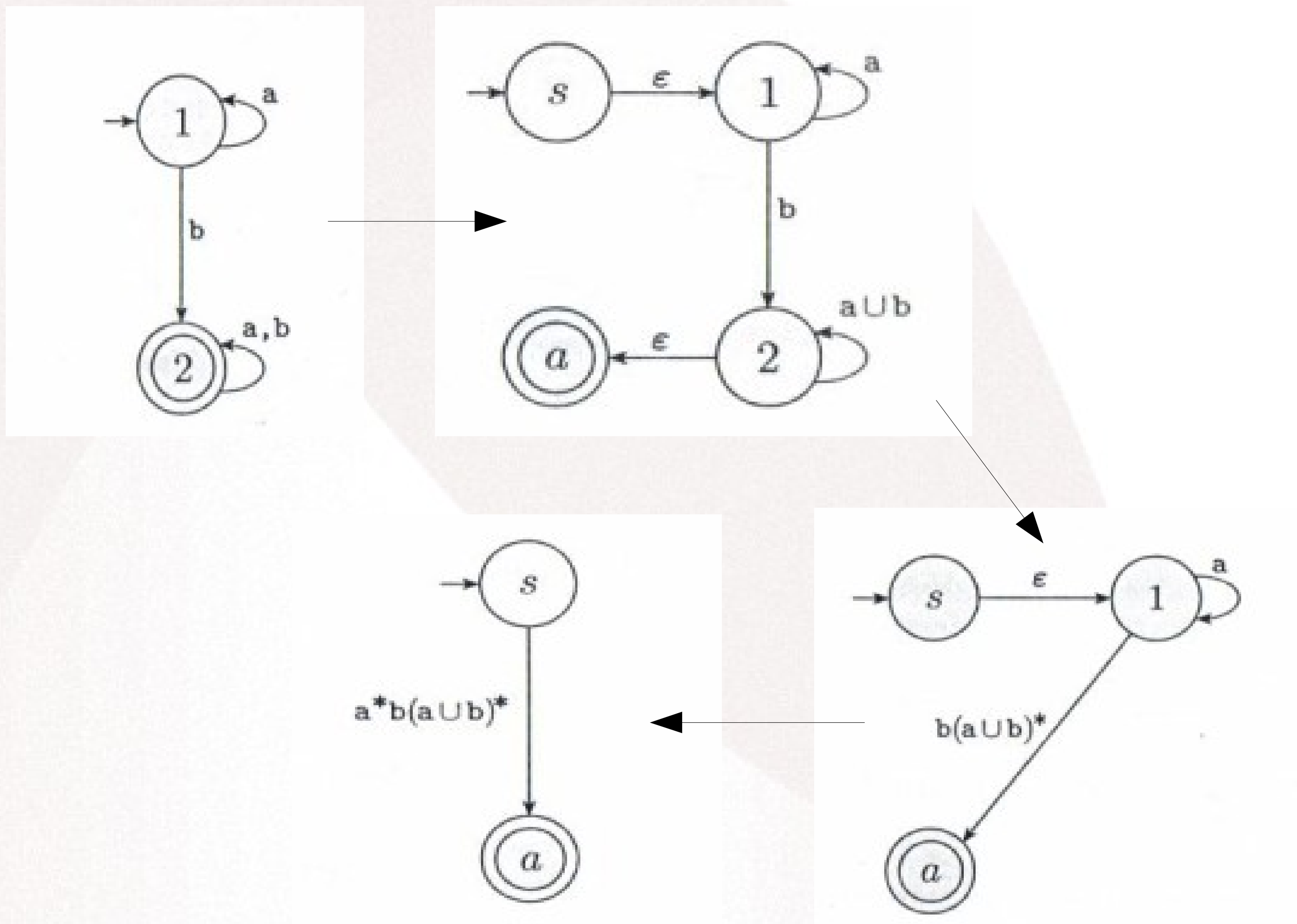


- R1 = b, R2 = aUb, R3 = ε e R4 = $\emptyset$.
- (b)(aUb)*(ε)U $\emptyset$ = b(aUb)*.



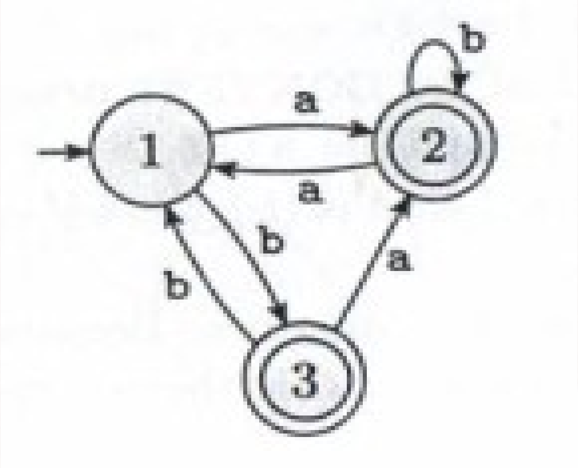
Prova: Lema 2 - Exemplo

- k=3, remova estados até k = 2.



Prova: Lema 2 - Exemplo

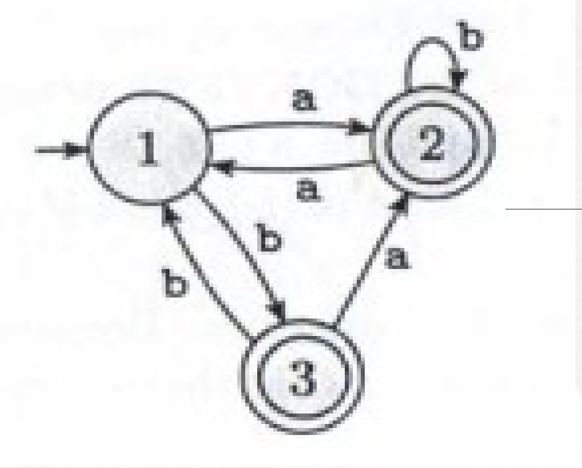
- Encontre a ER a partir do AFD abaixo:

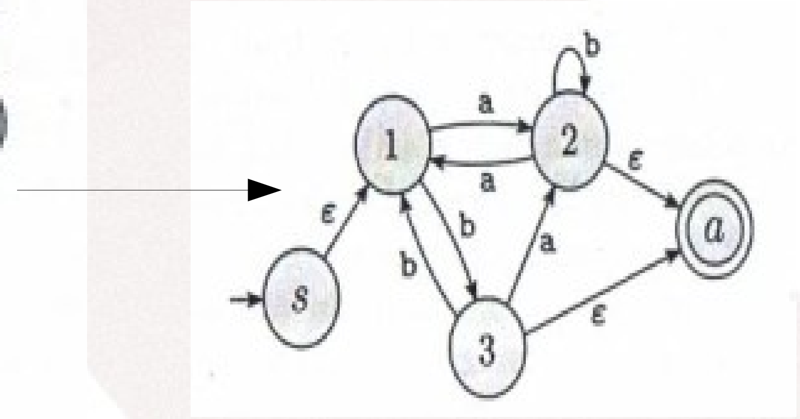


Já tentou fazer???
Se sim, pode continuar e olhar a resposta ;)

Prova: Lema 2 - Exemplo

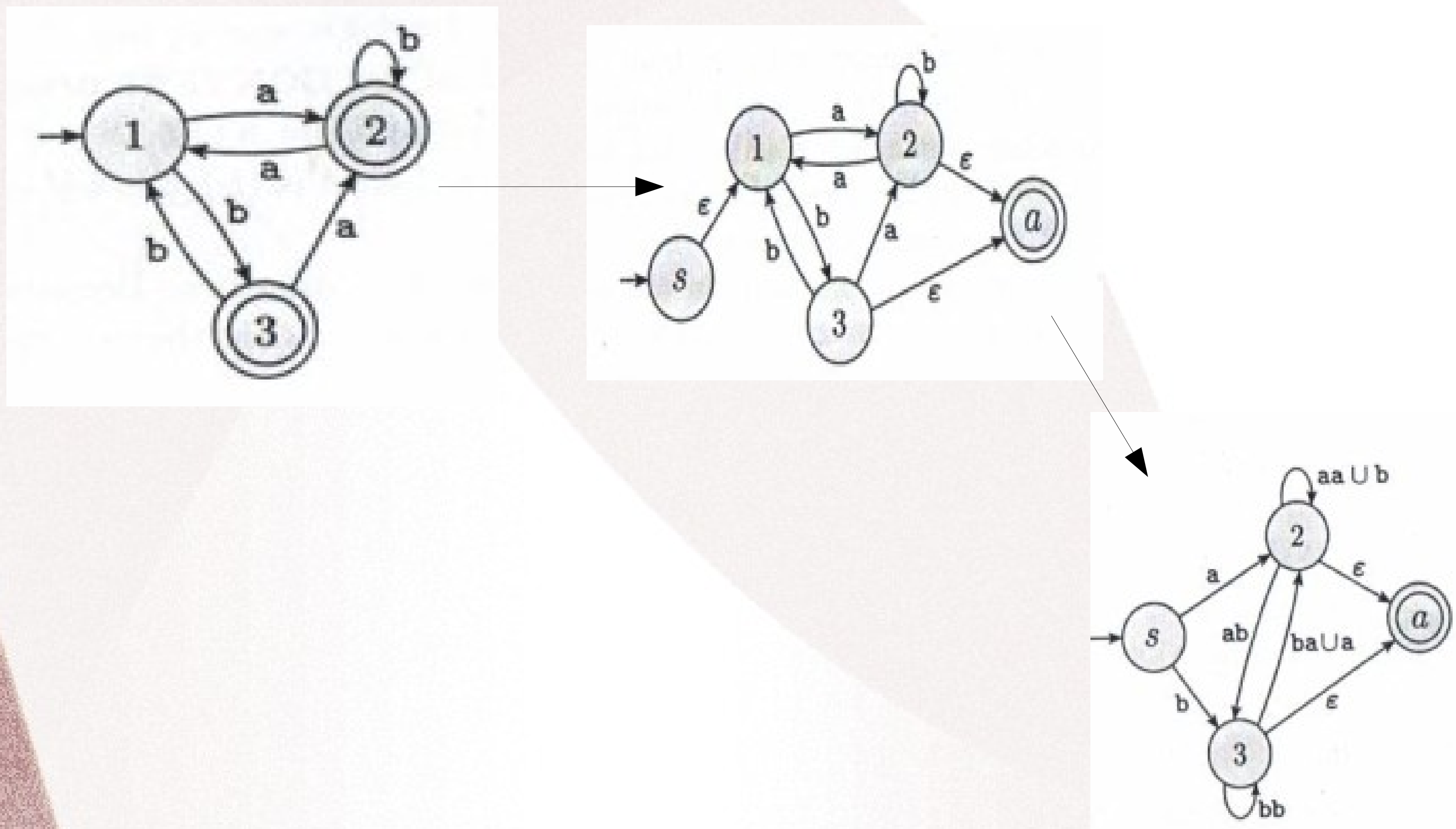
- Encontre a ER a partir do AFD abaixo:





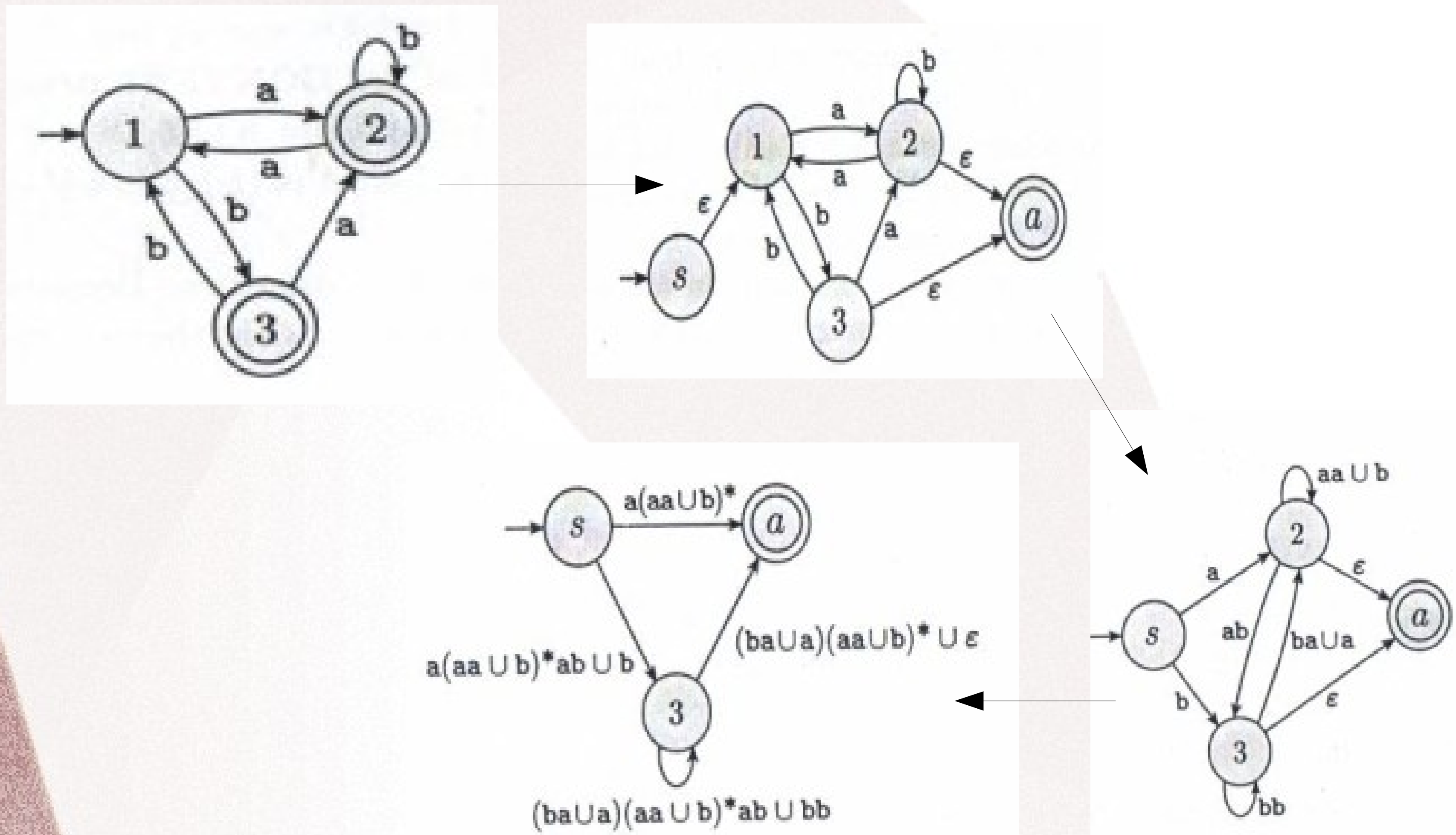
Prova: Lema 2 - Exemplo

- Encontre a ER a partir do AFD abaixo:



Prova: Lema 2 - Exemplo

- Encontre a ER a partir do AFD abaixo:



Prova: Lema 2 - Exemplo

- Encontre a ER a partir do AFD abaixo: