Linguagens Formais e Autômatos

Máquinas de Turing

Eduardo Furlan Miranda

Baseado em: GARCIA, A. de V.; HAEUSLER, E. H.
Linguagens Formais e Autômatos. Londrina: EDA, 2017.

Máquina de Turing (MT)

- Modelo teórico de computação proposto por Alan Turing em 1936

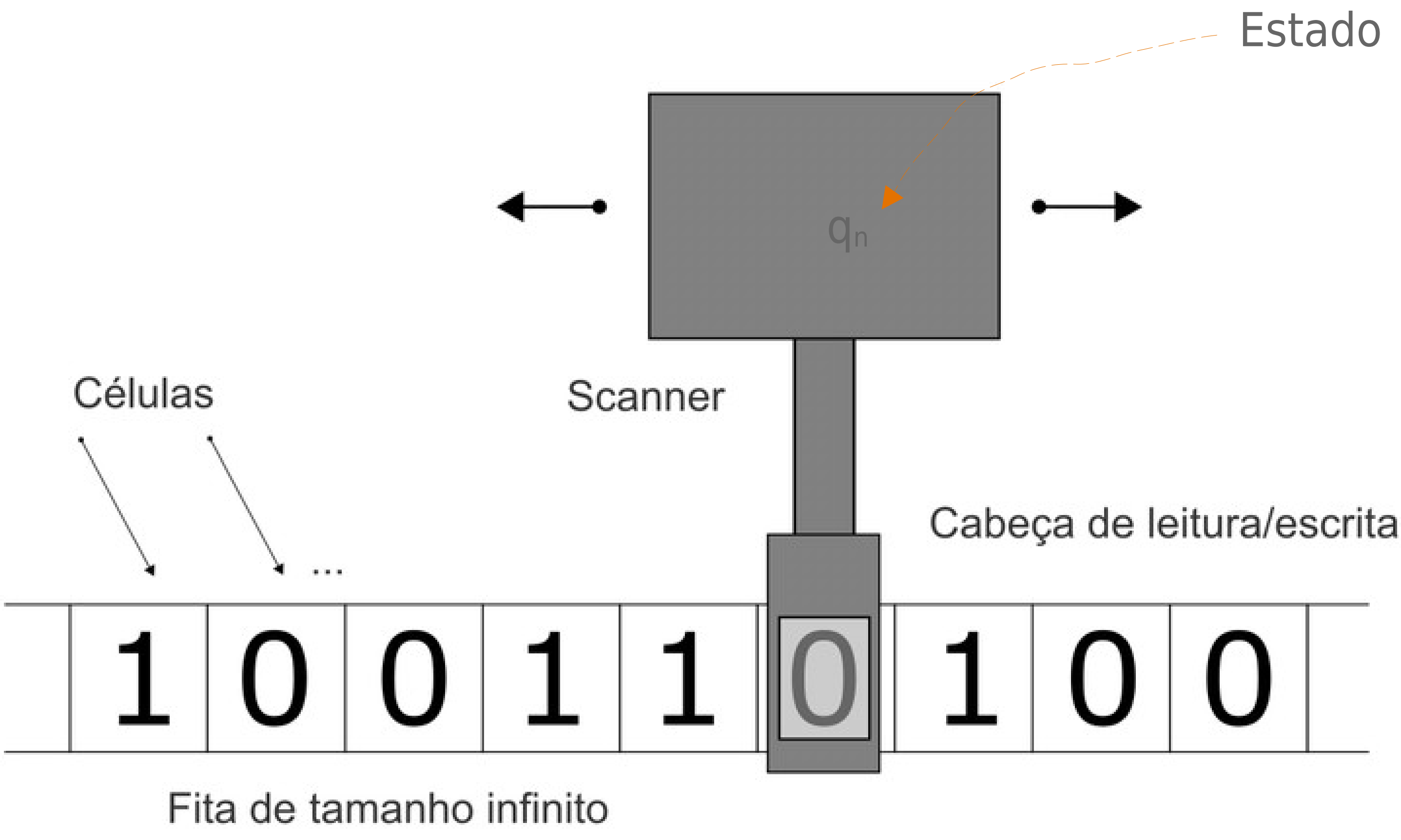
- Foi concebido para formalizar a noção de *procedimento* ou *algoritmo*, estabelecendo assim o conceito de *função computável*

- Uma MT é capaz de resolver qualquer problema que possa ser formalizado matematicamente, incluindo problemas de análise sintática

- Na prática, outras ferramentas e algoritmos são preferidos para a análise sintática, devido à maior eficiência computacional e à complexidade reduzida em comparação com a implementação direta de uma MT

MT como AF

- Podemos entender a MT como um Autômato Finito (AF) com uma estrutura de dados auxiliar (uma fita)

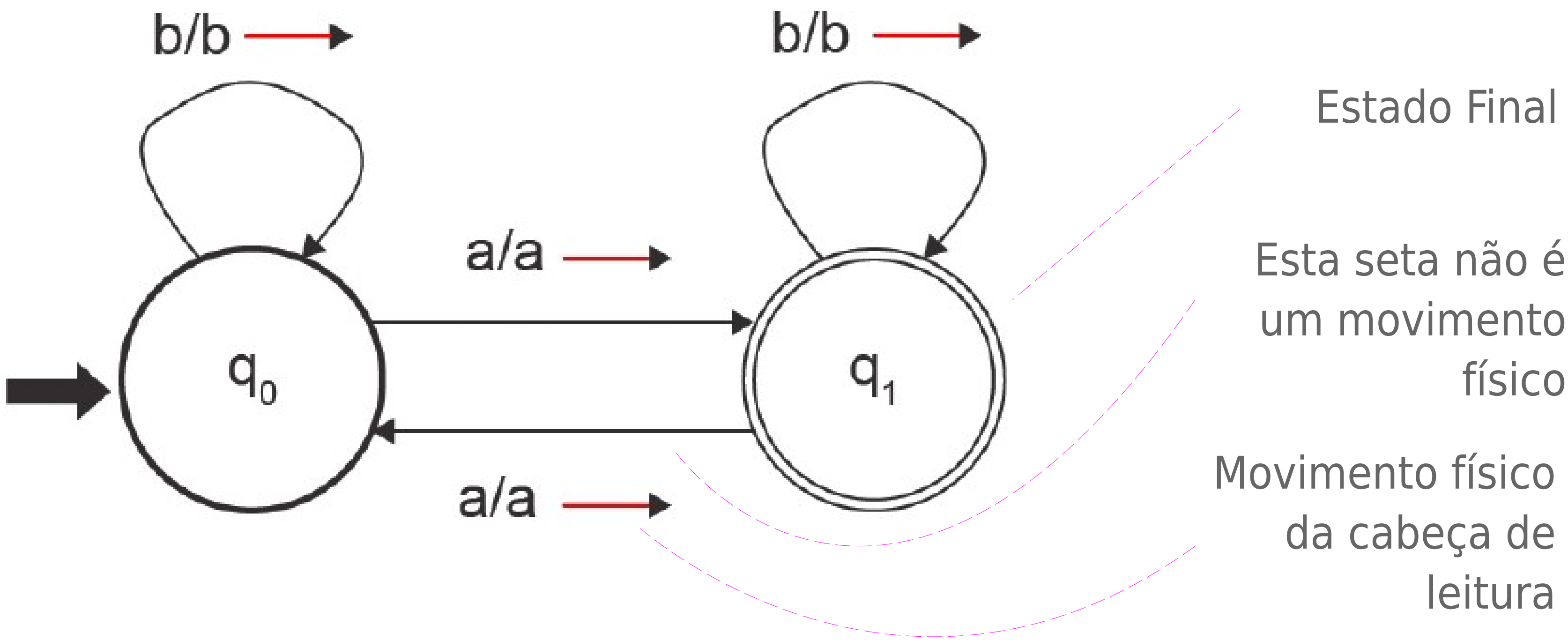


- O dado de entrada está escrito na fita

- A cada transição a máquina pode
  - ler um símbolo do alfabeto
  - escrever sobre o mesmo
  - andar na fita para direita ou para a esquerda
  - mudar de estado

- Vamos supor que a fita é infinita para a direita
  - possui uma primeira posição
  - cada posição da fita sempre possui uma posição à sua direita
- Cada posição da fita armazena um símbolo do alfabeto da fita
- No início da execução, para uma entrada de tamanho **n**,
  - as **n** primeiras posições da fita contém os símbolos da entrada
- As demais posições da fita estão em branco
  - possuem o símbolo branco

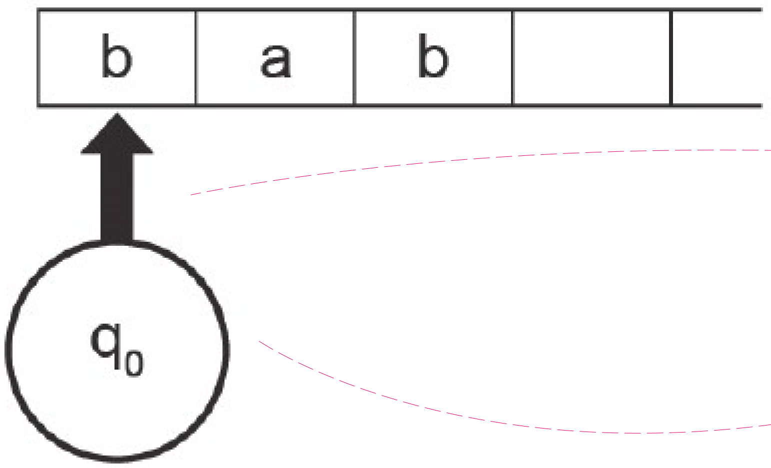
MT se comporta como um AF



- Ilustração de uma MT que reconhece a linguagem
  L = {w | w tem quantidade ímpar de a's }

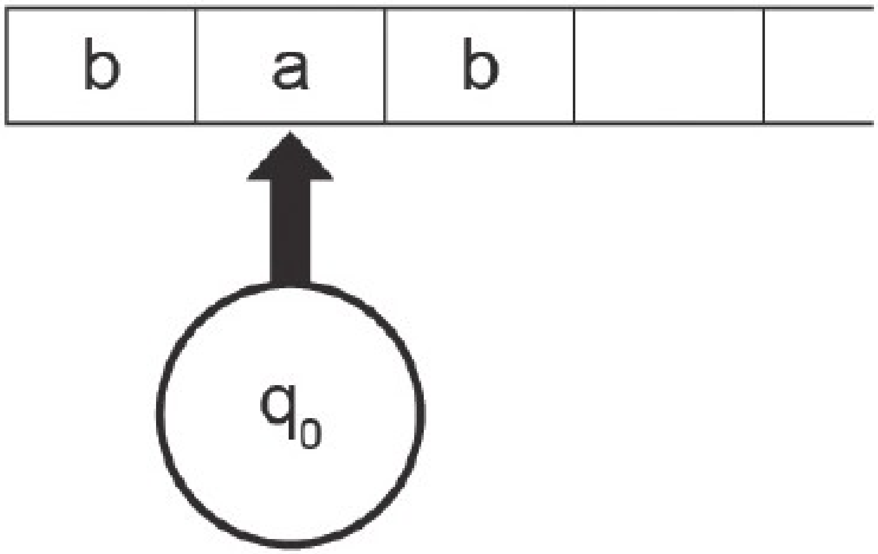
- Neste caso particular a MT anda sempre para a direita (representado pelo símbolo →)

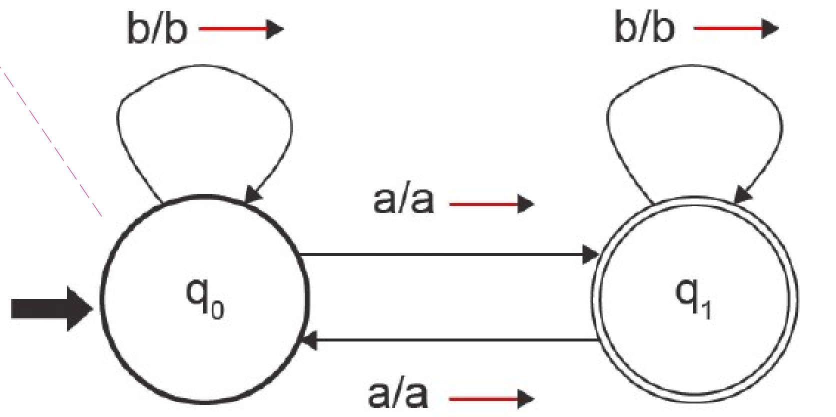
- Neste caso a máquina recebe apenas a cadeia de entrada,
    - e nenhuma informação adicional
- portanto a MT tem o mesmo poder de computação que um AF,
    - a linguagem reconhecida é uma Linguagem Regular (LR)
        - LR é uma linguagem formal que pode ser expressa usando Expressões Regulares (ER), ou seja,
            - uma linguagem produzida utilizando as operações de *concatenação*, *união* e *fecho de Kleene* sobre os elementos de um alfabeto
- Qualquer linguagem que pode ser reconhecida por um AF também é uma LR



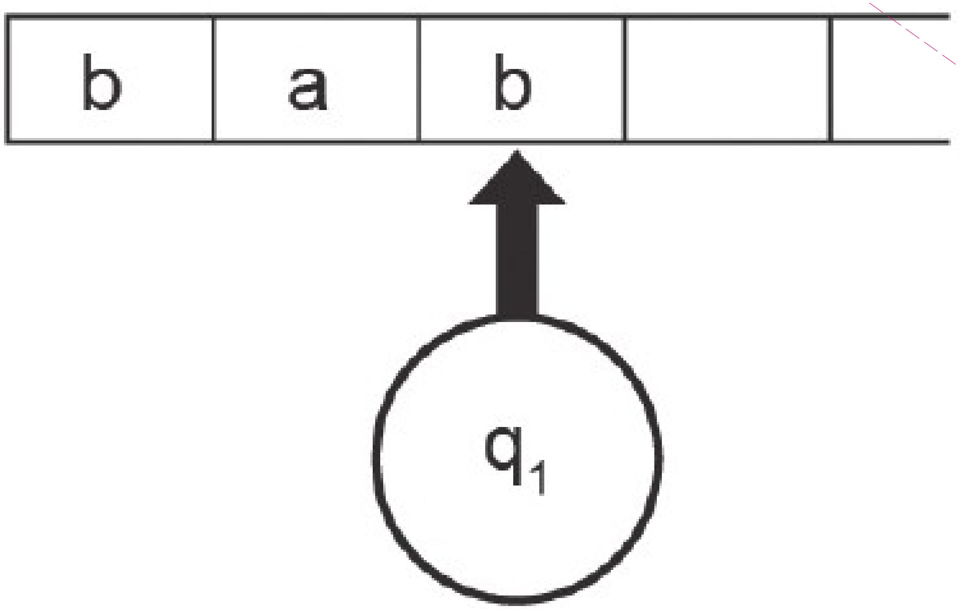
- Suponha que desejamos reconhecer a cadeia bab
- Neste caso, a cadeia estará no início da fita
- A posição da fita sendo lida pela máquina é representada pela seta vertical
  - e o estado no qual a máquina está é representado no círculo

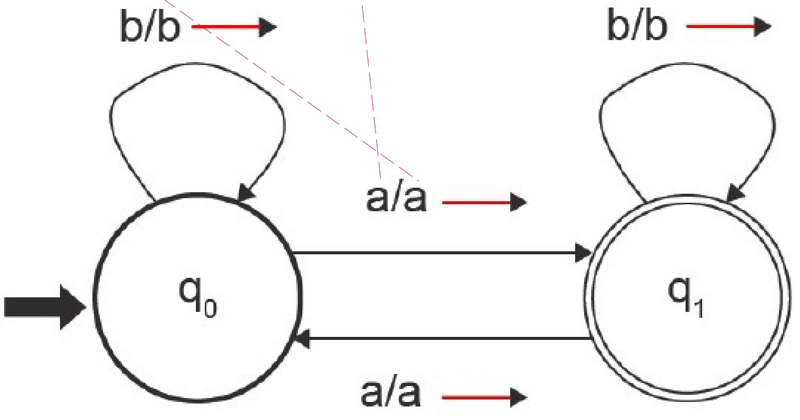
- Estando no estado q_0 e lendo b ,
  - a máquina escreve um b por cima
    - portanto não vai alterar o conteúdo da fita
- Segue para a direita e
  - continua no estado q_0



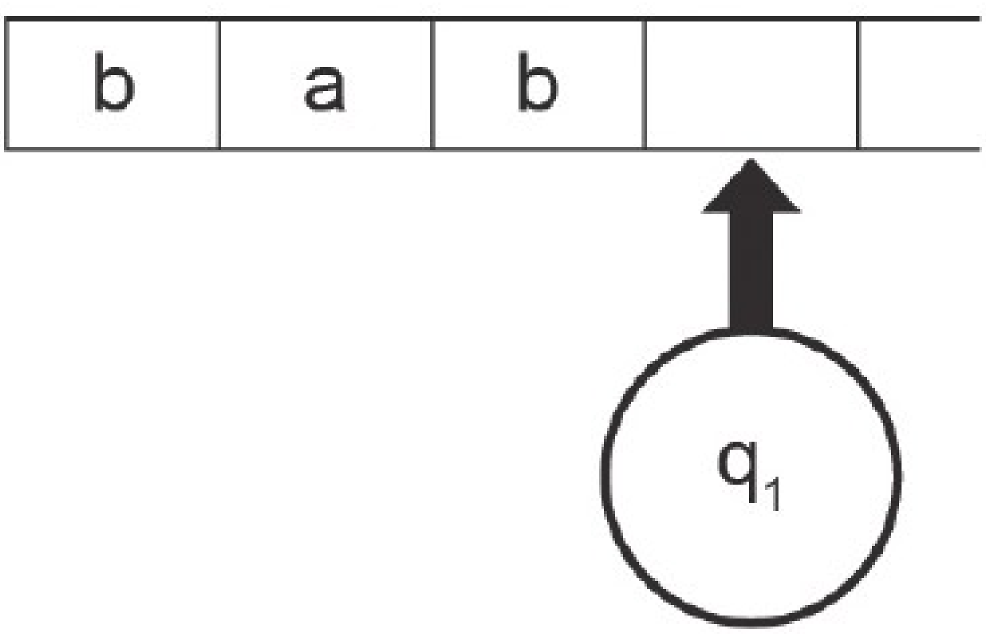


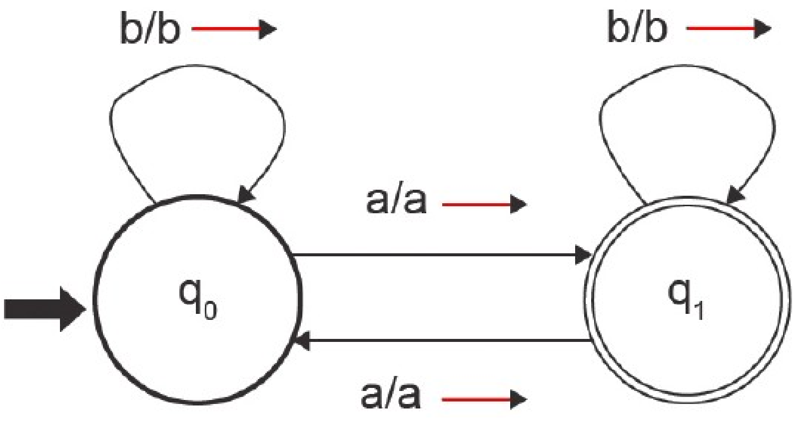
- Neste ponto, estando no estado q_0 e lendo um a, a MT irá escrever um a por cima (portanto, não vai alterar o conteúdo da fita), seguir para a direita (devido ao símbolo →) e mudar para o estado q_1





- **Finalmente**, estando no estado q_1 e lendo um b, a MT irá escrever um b por cima (portanto não vai alterar o conteúdo da fita), seguir para a direita (devido ao símbolo →) e permanecer no estado q_1





Reconheceu

- Neste caso, não se especifica o que fazer quando a MT lê um branco

- Dizemos que a MT parou

- Como ela parou em um estado final (q_1), dizemos que ela reconheceu a sentença bab

MT e LSC

R $a^n b$
LC $a^n b^n$
SC $a^n b^n c^n$
I a^{2^n}

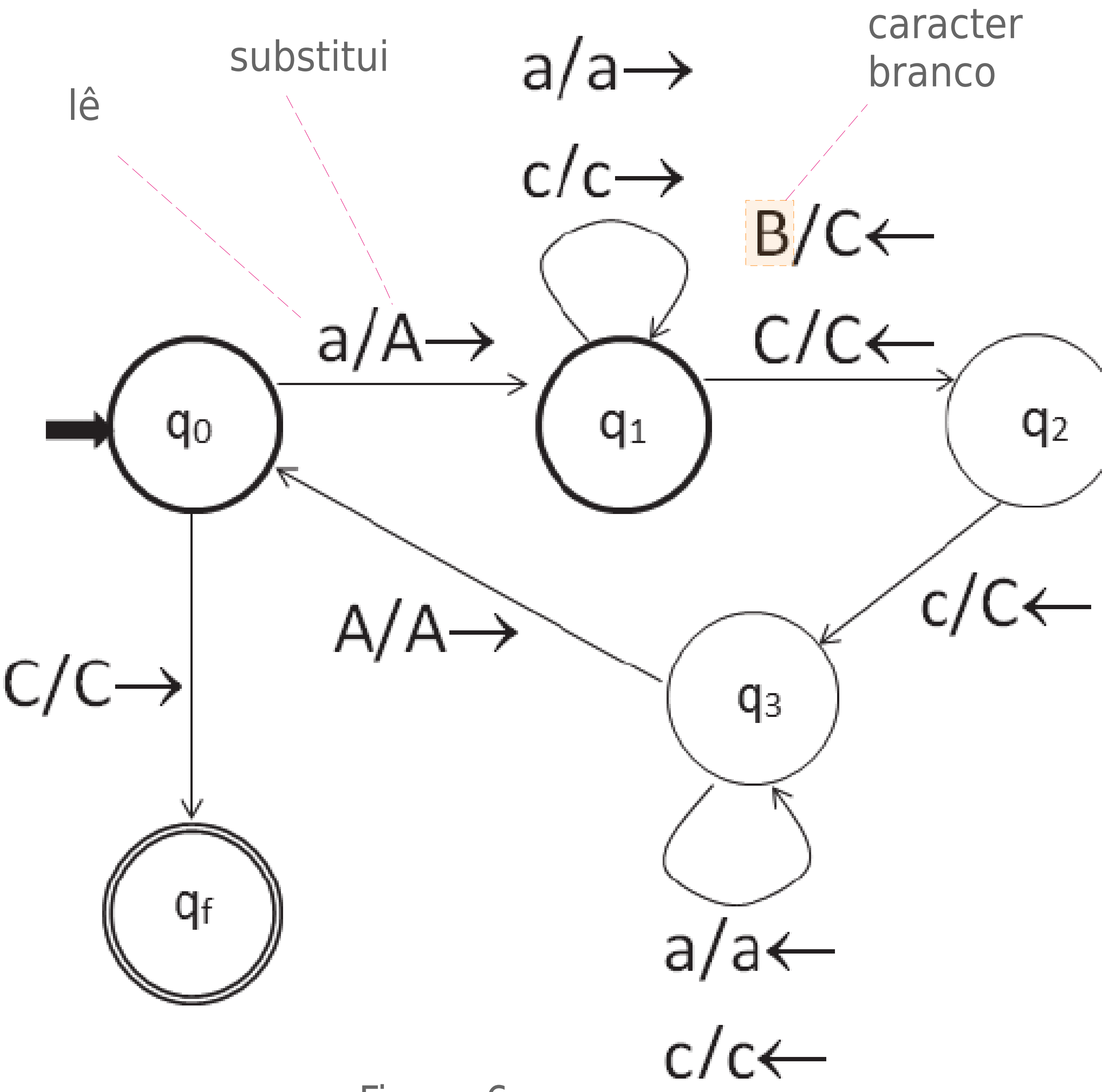


Figura 6

- MT que reconhece a linguagem $L=\{a^n c^n \mid n \geq 1\}$
- Como esta linguagem não é regular, a MT terá obrigatoriamente de se mover para a esquerda

- A técnica usada é marcar cada caractere lido:
  - cada a lido é substituído por um A e
  - cada c lido por um C

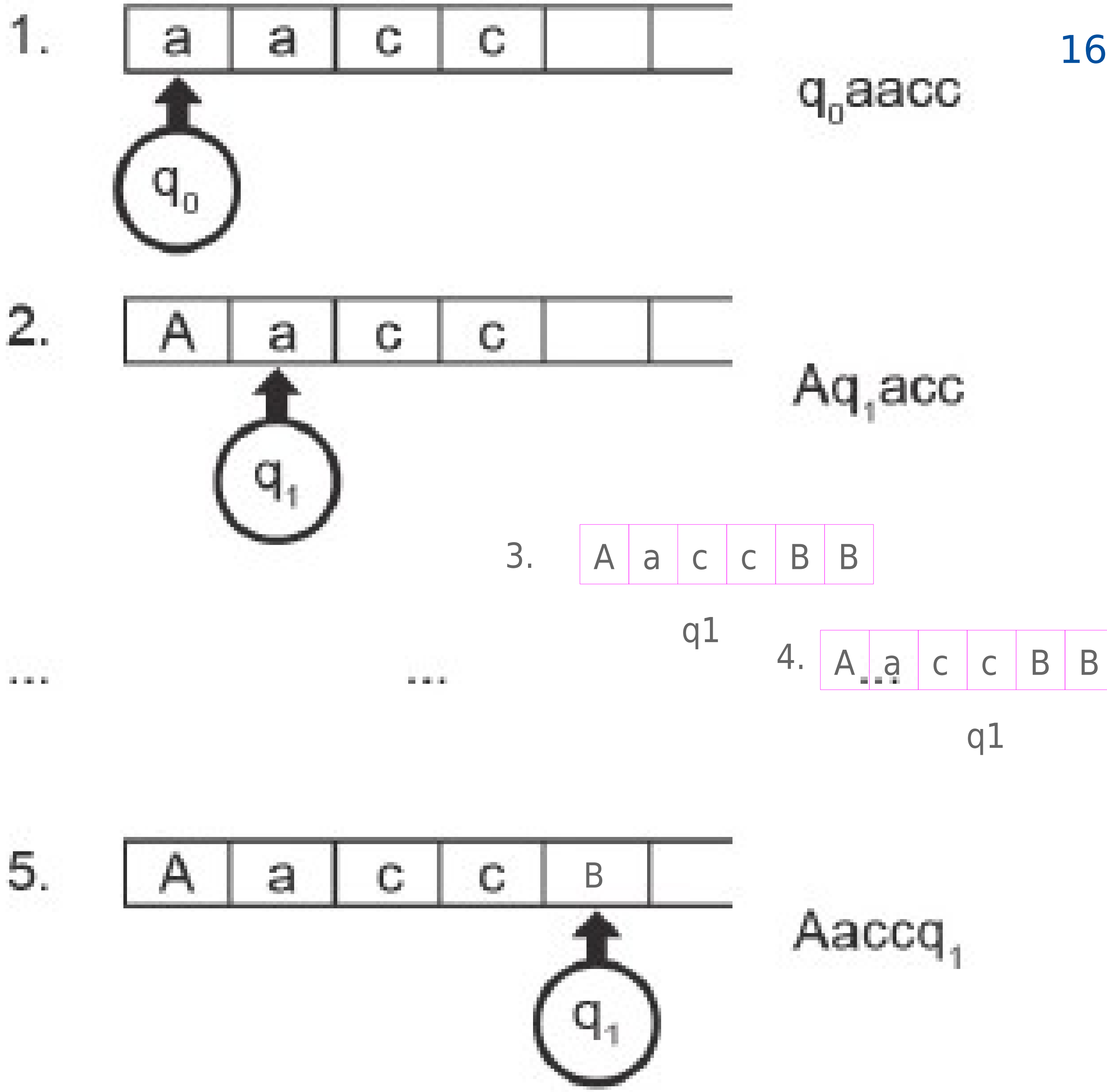
(vide slide anterior)

- O caractere branco é representado por B para facilitar a visualização

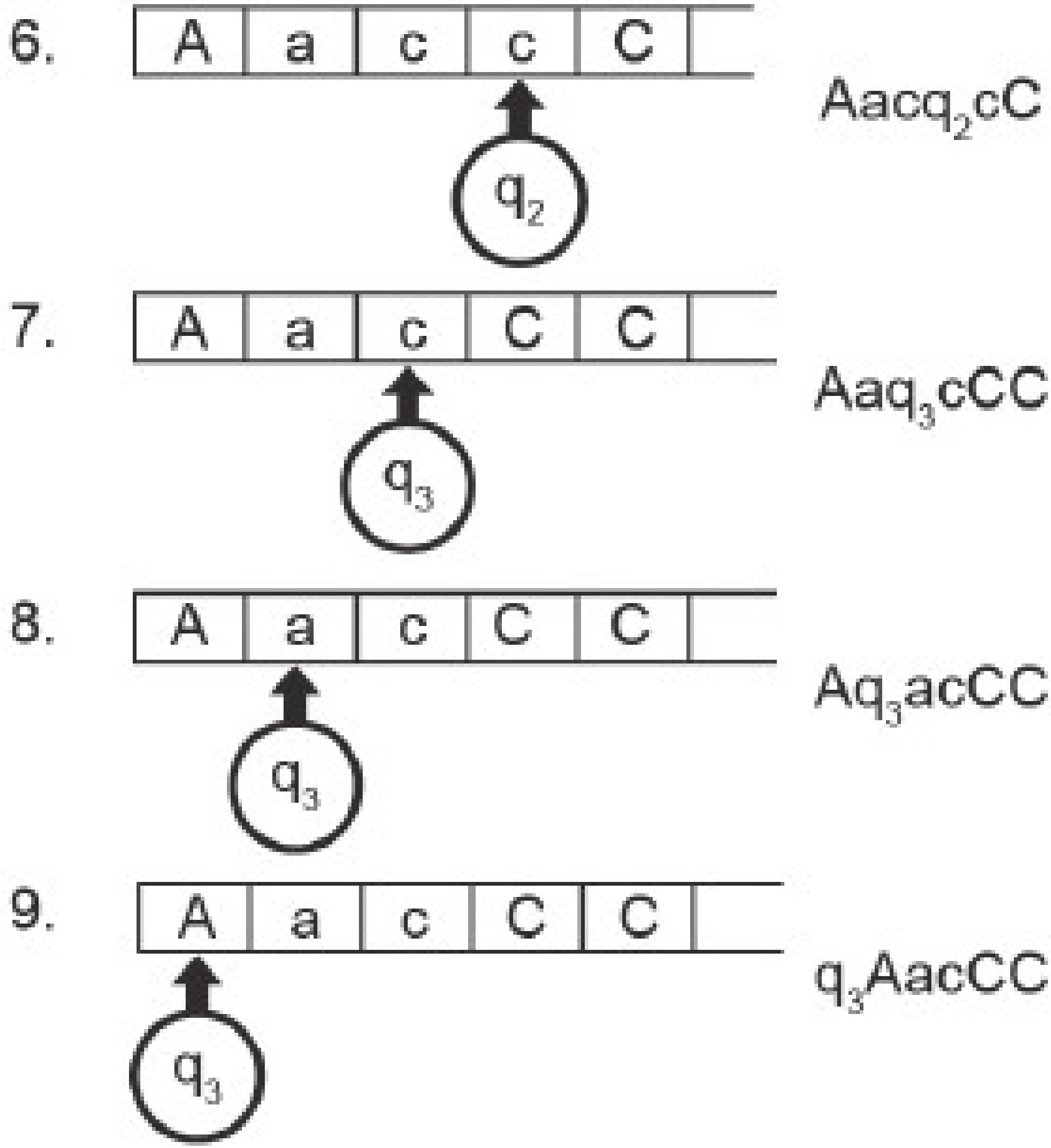
- Por convenção não é permitido escrever o B ,
  - portanto ele também é substituído por C

(vide slide anterior)

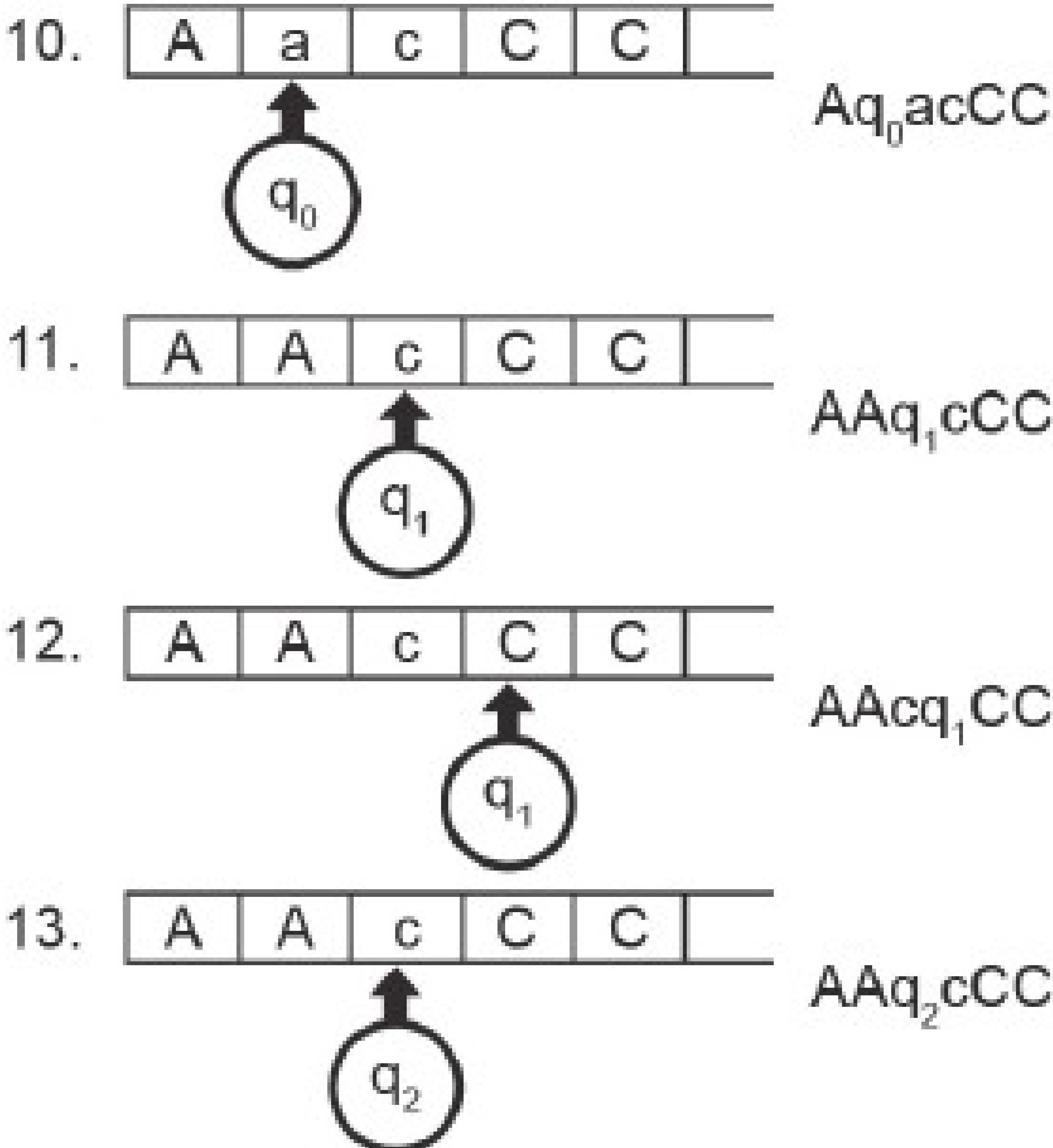
- Podemos representar a configuração do autômato em um dado ponto por $\alpha q \beta$, onde
  - α é o conteúdo da fita antes da cabeça de leitura
  - q é o estado no qual a máquina se encontra
  - β é o conteúdo da fita a partir da cabeça de leitura até o primeiro caractere branco (B)
- A configuração inicial da MT do slide a seguir, ao reconhecer a cadeia aacc, é $q_0$aacc
- A figura do próximo slide ilustra o reconhecimento passo a passo



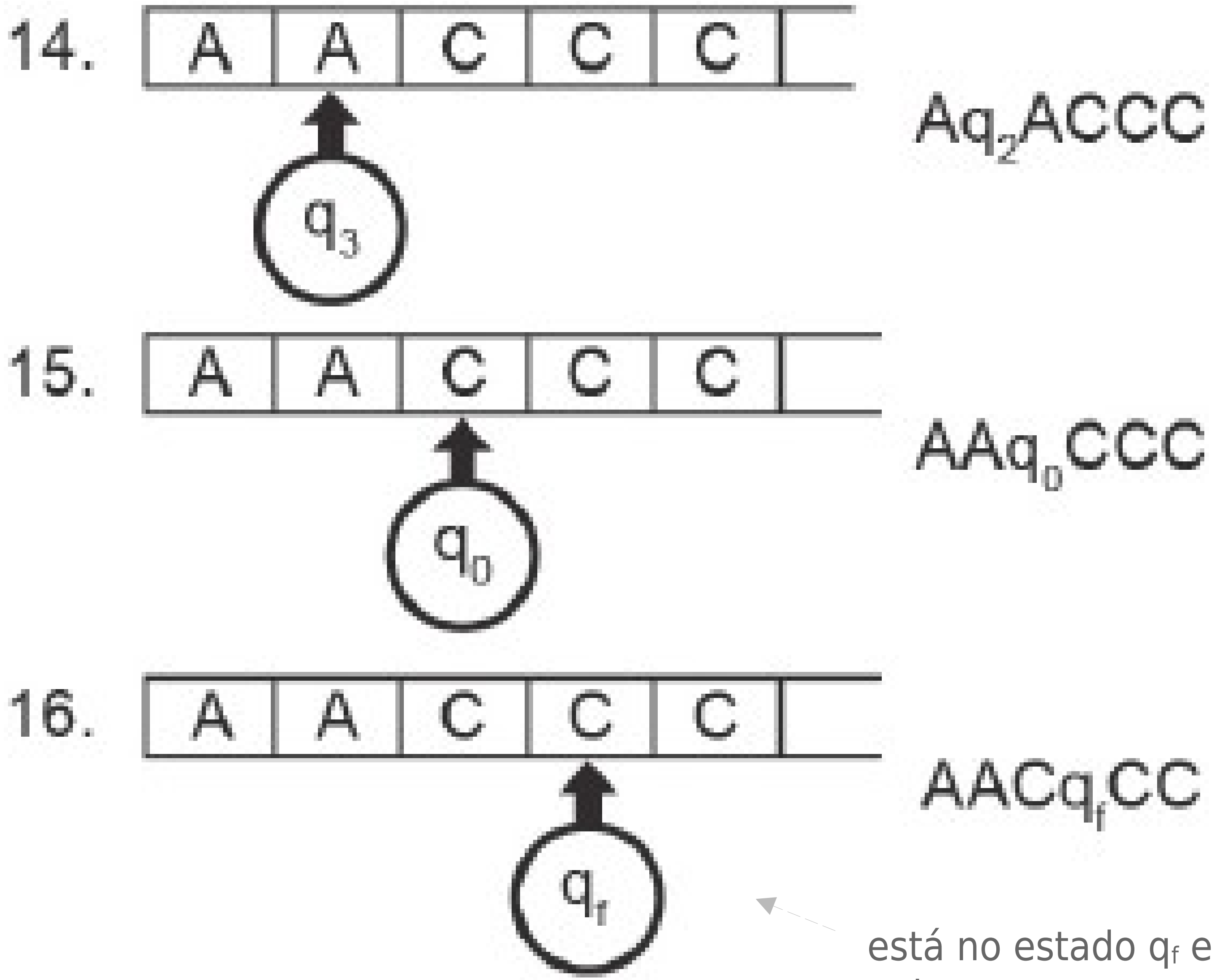
1.
a a c c
q_0
q_0aacc
2.
A a c c
q_1
Aq_1acc
3.
A a c c B B
q1
4.
A a c c B B
q1
...
...
...
5.
A a c c B
q_1
$Aaccq_1$



6.
A a c c C
$Aacq_2cC$
q_2
7.
A a c C C
Aaq_3cCC
q_3
8.
A a c C C
Aq_3acCC
q_3
9.
A a c C C
q_3AacCC
q_3


10.
A a c C C
q_0
Aq_0acCC
11.
A A c C C
q_1
AAq_1cCC
12.
A A c C C
q_1
$AAcq_1CC$
13.
A A c C C
q_2
AAq_2cCC

Reconhece



está no estado q_f e não tem outra regra

MT e AF

- Formalmente, uma MT possui alguns elementos em comum com um AF
  - Q : um conjunto finito de estados
  - Σ : um alfabeto de entrada sobre o qual as cadeias a serem reconhecidas estão definidas
  - $F \subseteq Q$: um conjunto de estados finais
  - $q_0 \in Q$: o estado inicial

- Como a MT possui uma fita que pode ser lida e escrita, ela possui os seguintes elementos diferentes de um AF
  - Γ : um conjunto finito de símbolos que podem ser escritos na fita
  - $B \in \Gamma$: o símbolo 'branco' tal que $B \notin \Sigma$, e B não pode ser escrito na fita
  - A fita é infinita para a direita e é inicializada com a palavra a ser reconhecida seguida de infinitos brancos
  - δ : função de transição
    - $Q \times \Gamma$ em $\wp(Q \times (\Gamma - \{B\}) \times \{\leftarrow, \rightarrow\})$

(continua no próximo slide)

alfabeto de entrada

Função de transição (δ) de uma MT

- $\delta : Q \times \Gamma$ em $\wp(Q \times (\Gamma - \{B\}) \times \{\leftarrow, \rightarrow\})$
  - “:” : significa "é uma função de" ou "mapeia"
  - Q : conjunto finito de estados
  - Γ : alfabeto da fita

    $\times$ = produto cartesiano
  - $Q \times \Gamma$: conjunto de todos os pares possíveis formados por um estado de Q e um símbolo de Γ
  - $\wp$: conjunto potência (conjunto de todos os seus subconjuntos)
    - caracteriza uma Máquina de Turing Não Determinística
  - Q : o novo estado (estado destino)
  - $\Gamma - \{B\}$: alfabeto da fita Γ excluindo o símbolo de branco B
  - $\{\leftarrow, \rightarrow\}$: conjunto das direções possíveis para o movimento da cabeça de leitura/escrita: esquerda ($\leftarrow$) ou direita ($\rightarrow$)
  - $(Q \times (\Gamma - \{B\}) \times \{\leftarrow, \rightarrow\})$: conjunto de onde os elementos dos subconjuntos (retornados pela função δ) são retirados

Função de transição

(repetição do slide anterior)

- δ : uma função de $Q \times \Gamma$ em $\wp(Q \times (\Gamma - \{B\}) \times \{\leftarrow, \rightarrow\})$

- Cada transição pertence a $Q \times (\Gamma - \{B\}) \times \{\leftarrow, \rightarrow\}$
  - Possui um estado destino, um símbolo não branco da fita (símbolo que irá sobrescrever o símbolo lido), e uma direção (esquerda ou direita) para a qual a cabeça de leitura irá se mover

- Caso a cabeça se mova à esquerda do limite inicial da fita, nada acontece, a máquina não vai para uma nova configuração

- Um estado e um símbolo sendo lido da fita podem estar associados a um conjunto de transições
  - Se este conjunto possuir sempre no máximo uma transição, dizemos que a MT é determinística

Definição

- Formalmente, definimos uma MT M como uma tupla
  $M = (Q, \Sigma, \Gamma, \delta, q_0, B, F)$
  - onde os componentes foram descritos anteriormente

- Seja uma MT que está em uma configuração $\alpha\mathbf{q}a\beta$, e possui a transição
  $(\mathbf{r}, x, \rightarrow) \in \delta(\mathbf{q}, a)$ onde $q, r \in Q$; $a, x \in \Gamma$; $\alpha, \beta \in \Gamma^*$
  e se move para a configuração $\alpha x\mathbf{r}\beta$
  - A MT muda de uma configuração para outra seguindo uma regra de transição que inclui mudar de estado, escrever um símbolo e mover a cabeça de leitura/escrita

- Neste caso descrevemos como $\alpha\mathbf{q}a\beta \vdash \alpha x\mathbf{r}\beta$

estado

símbolo

relação de transição

a cabeça de leitura agora está posicionada após o x devido ao “→”

$\alpha \mathbf{q} a \beta \vdash \alpha x \mathbf{r} \beta$

(repetição do slide anterior)

- $(\mathbf{r}, x, \rightarrow) \in \delta(\mathbf{q}, a)$ onde $q, r \in Q$; $a, x \in \Gamma$; $\alpha, \beta \in \Gamma^*$
  - Configuração atual
    - α e β são sequências de símbolos na fita
    - **q** é o estado atual da máquina
    - a é o símbolo na célula atual
  - Transição
    - Existe a transição $(\mathbf{r}, x, \rightarrow)$ na função δ, que diz que
      - Se a máquina está no estado **q** e lê o símbolo a
      - ela escreve o símbolo x no lugar de a
      - move a cabeça de leitura/escrita para a direita (→)
      - e muda para o estado **r**
  - Nova configuração ($\alpha x \mathbf{r} \beta$)
    - α permanece igual
    - a é substituído por x
    - **r** é o novo estado
    - e a cabeça de leitura/escrita move-se para a célula à direita de x

- De forma análoga, se a máquina está em uma configuração $\alpha b\mathbf{q}a\beta$, possui a transição
  $(\ \mathbf{r}, x, \leftarrow\ ) \in \delta(\ \mathbf{q}, a)$ onde $q, r \in Q$; $a, x \in \Gamma$; $\alpha, \beta \in \Gamma^*$
  e se move para a configuração $\alpha\mathbf{r}bx\beta$

- Neste caso, escrevemos $\alpha b\mathbf{q}a\beta \vdash \alpha\mathbf{r}bx\beta$

- Se a MT não pode se mover em uma dada configuração, dizemos que a máquina parou nesta configuração deriva em zero ou mais passos

- Chamamos $\vdash^*$ à aplicação de 0 ou mais vezes da relação de transição $\vdash$

- A MT $M = (Q, \Sigma, \Gamma, \delta, q_0, B, F)$ reconhece uma sentença w se, começando do estado inicial q_0, e lendo a sentença w a partir da fita, a máquina pode fazer uma série de transições, denotadas por $\vdash^*$, até alcançar uma configuração $\alpha\, q_f\, \beta$, onde q_f é um estado de aceitação (ou estado final) pertencente ao conjunto F
  - ou, $q_0\, w \vdash^* \alpha\, q_f\, \beta$ onde $q_f \in F$, e a MT para na configuração $\alpha\, q_f\, \beta$

- A linguagem reconhecida pela MT é o conjunto de todas as sentenças reconhecidas pela MT

Linguarem reconhecida pela MT

- Dada uma MT $M = (Q, \Sigma, \Gamma, \delta, q_0, B, F)$, definimos a linguagem reconhecida pela MT como

  $T(M) = \{ w \in \Sigma^* \mid \exists\, q_f \in F ,\ q_0\, w \vdash^* \alpha\, q_f\, \beta$ e M para em $\alpha\, q_f\, \beta \}$

  - ou seja, a linguagem reconhecida pela Máquina de Turing M, denotada por $T(M)$, é o conjunto de todas as strings w pertencentes a Σ^* (ou seja, todas as strings de entrada possíveis), tal que existe ($\exists$) um estado final q_f pertencente ($\in$) a F, e a partir da configuração inicial $q_0\, w$, a MT transiciona em zero ou mais passos ($\vdash^*$) para uma configuração $\alpha\, q_f\, \beta$, onde α e β são strings de símbolos da fita, e a MT para nessa configuração $\alpha\, q_f\, \beta$

MT não determinística

- Segundo a definição, para uma MT não determinística (MTND), basta um dos caminhos possíveis fazer a MT parar em um estado final para que a MT reconheça a cadeia

- Na MTND, em vez de especificar uma única ação para cada combinação de estado e símbolo, a função de transição pode especificar um conjunto de ações possíveis