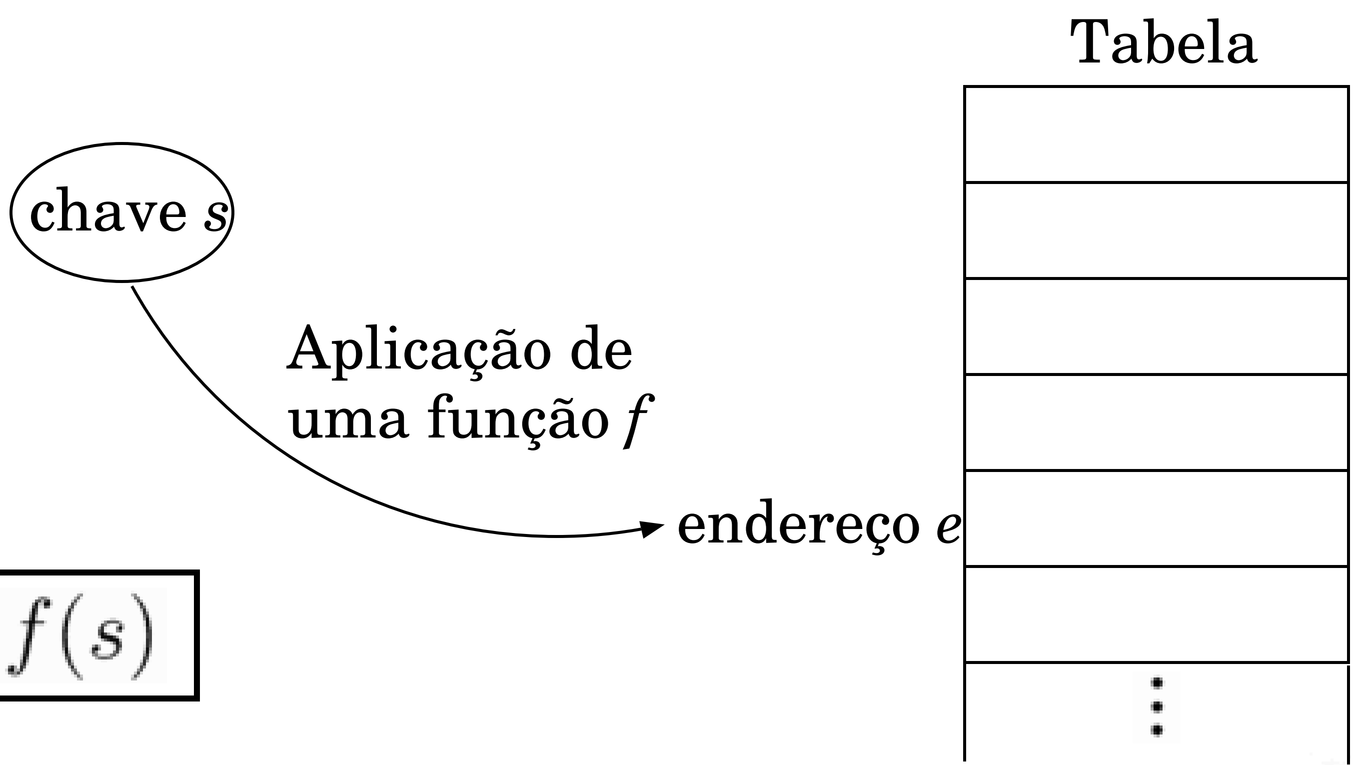
Aula 30: Tabelas de dispersão

- Conceitos e princípios de funcionamento
- Funções de dispersão
- Método da divisão

Idéia básica para o uso de tabelas de dispersão

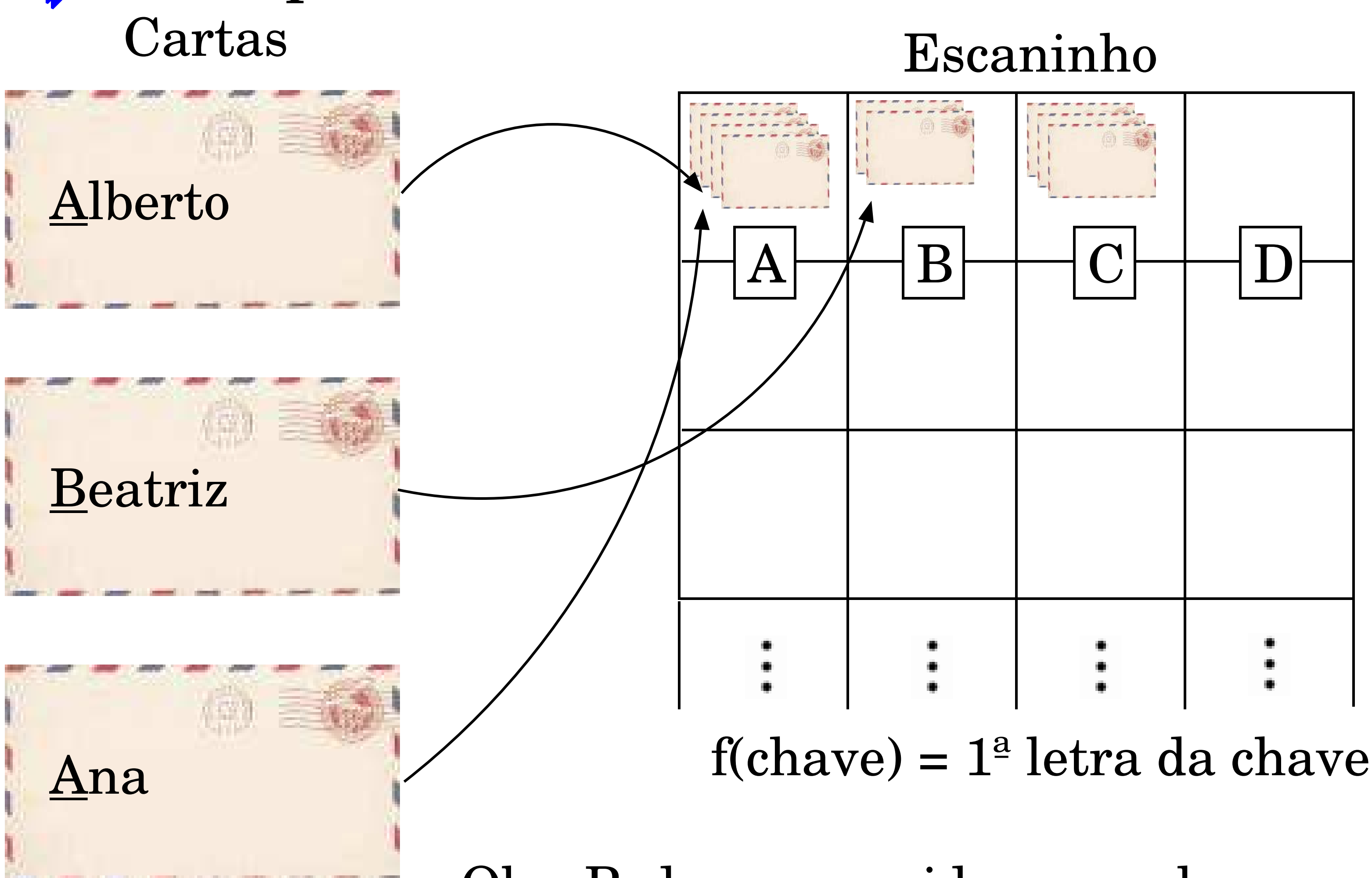
- Transformar as chaves em endereços de uma tabela, como tentativa de fazer a busca de chaves em tempo O(1).



$$\boxed{e = f(s)}$$

Idéia básica para o uso de tabelas de dispersão

➩ Exemplo: "escaninho"

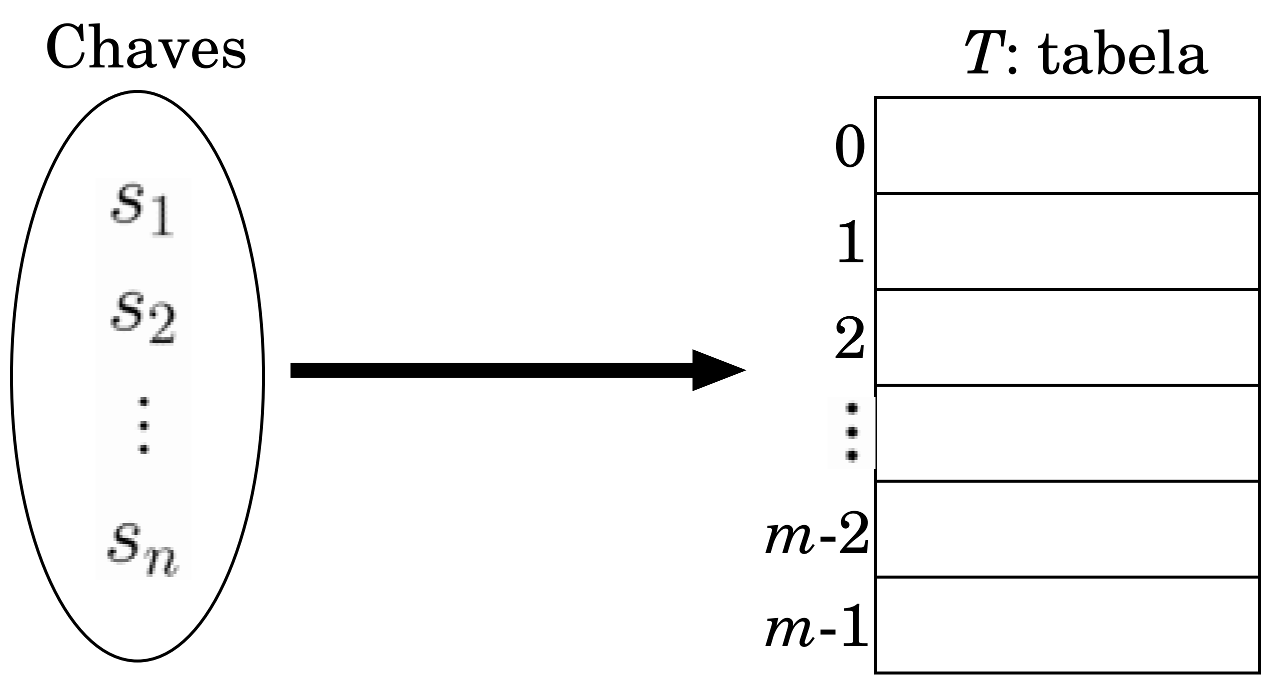


f(chave) = 1ª letra da chave

Obs: Podemos considerar as chaves como valores numéricos.

Princípio de funcionamento

- Suponha que existam $\underline{n}$ chaves a serem armazenadas em uma tabela T com $\underline{m}$ compartimentos numerados 0, 1, 2, ..., m-1

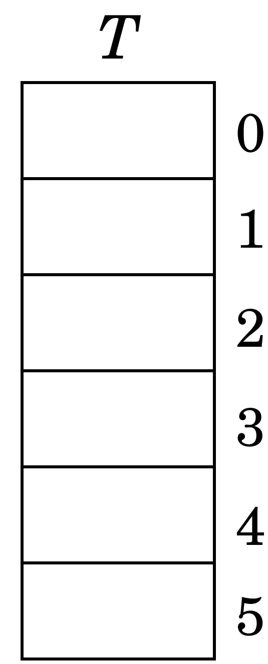


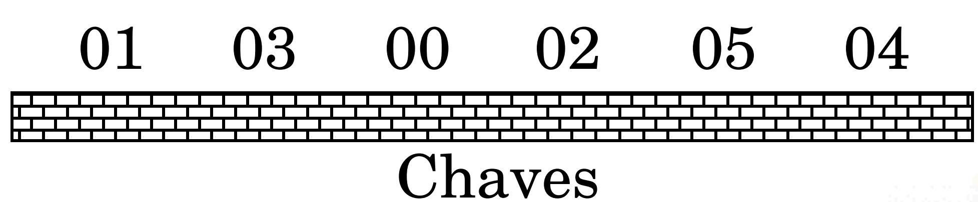
Acesso direto

- O endereço da chave s é o próprio valor de s, isto é

$$\boxed{f(s) = s}$$

- Exemplo 1
  - $n = m$
  - os valores das chaves são 0, 1, ..., m-1
  - cada chave s é armazenada no compartimento s



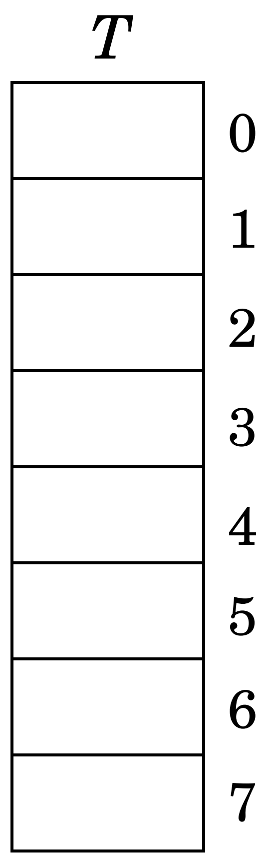


Armazenar Voltar

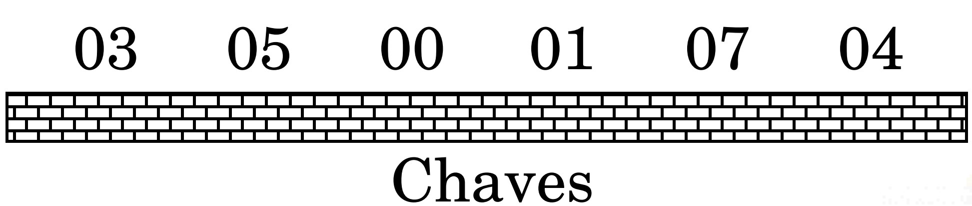
Acesso direto

Exemplo 2

- $n < m$
- m - n é pequeno
- cada chave s é armazenada no compartimento s
- T fica com m - n espaços (compartimentos vazios)



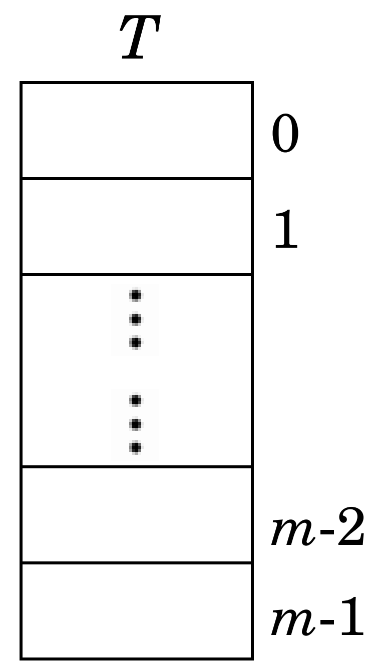
$$\boxed{\begin{array}{l} n = 6 \\ m = 8 \\ m - n = 2 \end{array}}$$

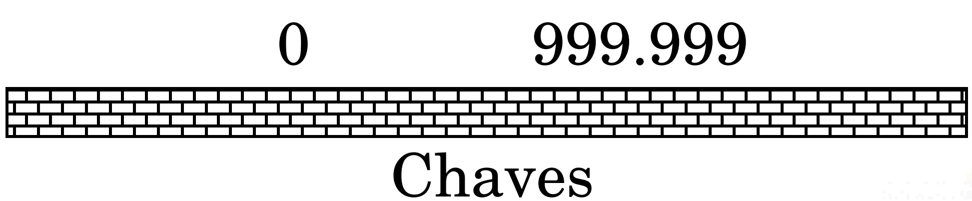


Acesso direto

Exemplo 3

- $n = 2$
- $m = 1.000.000$
- $n << m$
- $m - n$ é grande
- valores das chaves: 0 e 999.999



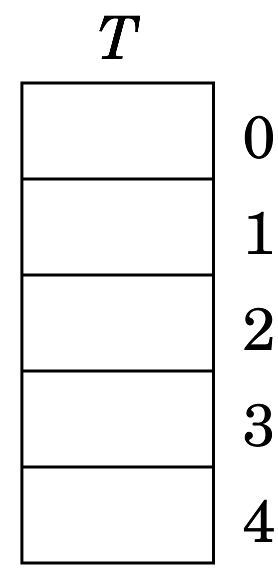


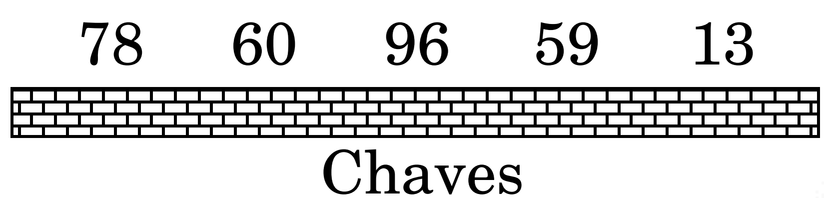
Uso de funções de dispersão geral

- Idéia: transformar a chave x num endereço-base $h(x)$, que é um valor entre 0 e m-1.
  h é chamada função de dispersão.

- Exemplo 4

$$h(x) = x \bmod 5$$





Armazenar Voltar

O fenômeno da colisão

- Ocorre quando a função de dispersão h não é injetora, isto é, existem duas chaves diferentes x e y com o mesmo valor de endereço-base. Temos portanto:

$$\boxed{h(x) = h(y) \text{ e } x \neq y}$$

- Neste caso, x e y são chamadas chaves sinônimas em relação a h.

- As técnicas que lidam com chaves sinônimas englobam-se sob o método geral denominado tratamento de colisões.

Condições que deve satisfazer uma boa função de dispersão

- Produzir um número baixo de colisões.
- Ser facilmente computável.
- Ser uniforme, isto é: a função h deve dar a todos os compartimentos de T a mesma probabilidade de serem escolhidos.

  Se T tem m compartimentos, então cada um deles tem probabilidade $1/m$ de ser endereço-base de alguma chave.

Método da divisão

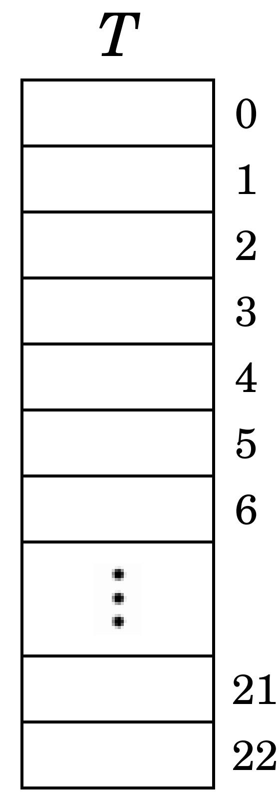
- Consiste em calcular o endereço-base da chave x através da função de dispersão

$$\boxed{h(x) = x \bmod m}$$

onde m é o tamanho da tabela T.

- <u>Exemplo 5</u>

$$h(x) = x \bmod 23$$



44 46 49 68 71 97

Chaves

Armazenar Voltar

Exercício final

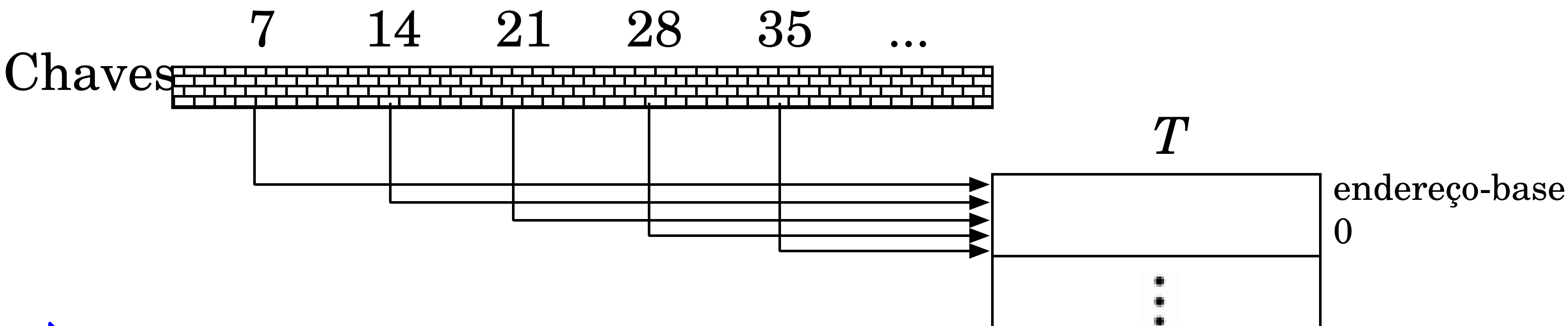
- Suponha um conjunto de n chaves formado pelos n primeiros múltiplos do número 7.

  Quantas colisões seriam obtidas mediante a aplicação das funções de dispersão seguintes:

  (a) $x \bmod 7$
  (b) $x \bmod 14$
  (c) $x \bmod 5$

Resolução do exercício

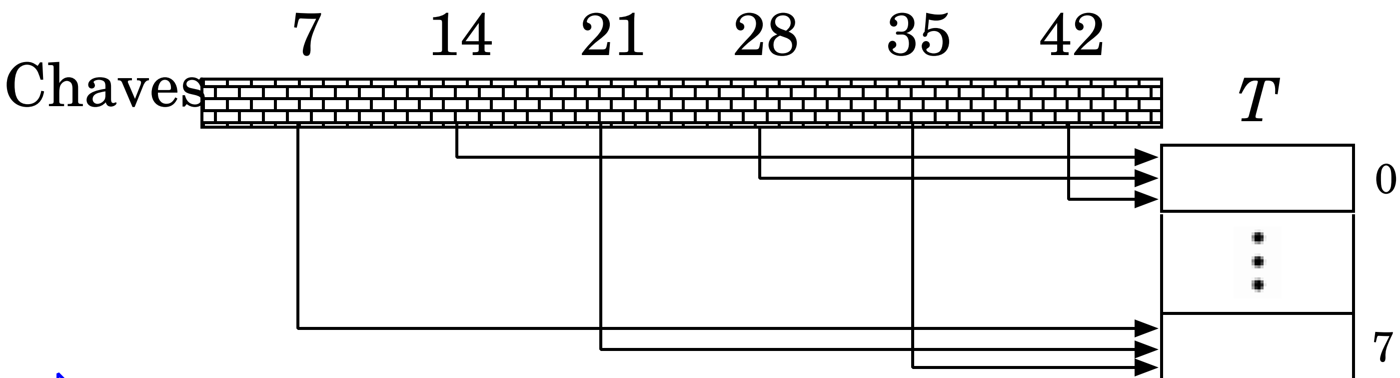
- Item (a)
  Para a função $h(x) = x \bmod 7$, todas as chaves são sinônimas!
  Observe:



- Se temos n chaves, seriam então produzidas $\underline{n\text{-}1}$ colisões, pois:
  - a primeira chave, 7, é colocada no endereço 0, que está inicialmente vazio
  - a partir daí, as n-1 chaves restantes geram n-1 colisões

Resolução do exercício

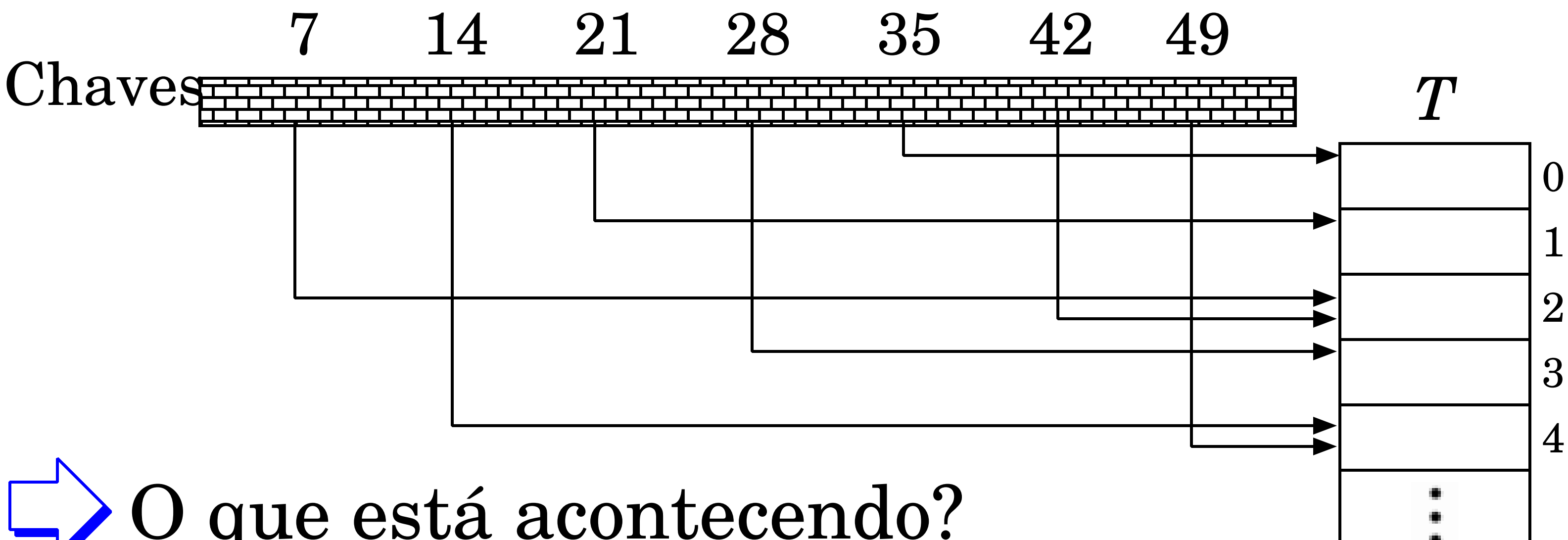
- Item (b)
  Para a função $h(x) = x \bmod 14$, observe o que acontece:



- As chaves múltiplas de 7 têm endereço-base 7 e as chaves múltiplas de 14 têm endereço-base 0
- Logo, para $n \leq 2$, temos 0 colisões (as chaves 7 e 14 não geram colisões)
- Para $n > 2$, temos n-2 colisões (as n-2 maiores chaves ou colidem com a chave 7 ou com a chave 14)

Resolução do exercício

- **Item (c)**
  Para a função $h(x) = x \bmod 5$, observe o que acontece:



- O que está acontecendo?
  Você consegue deduzir?
- Mostre que o número de colisões é

$$\begin{cases} 0, & \text{se } n \leq 5 \\ n-5, & \text{se } n > 5 \end{cases}$$