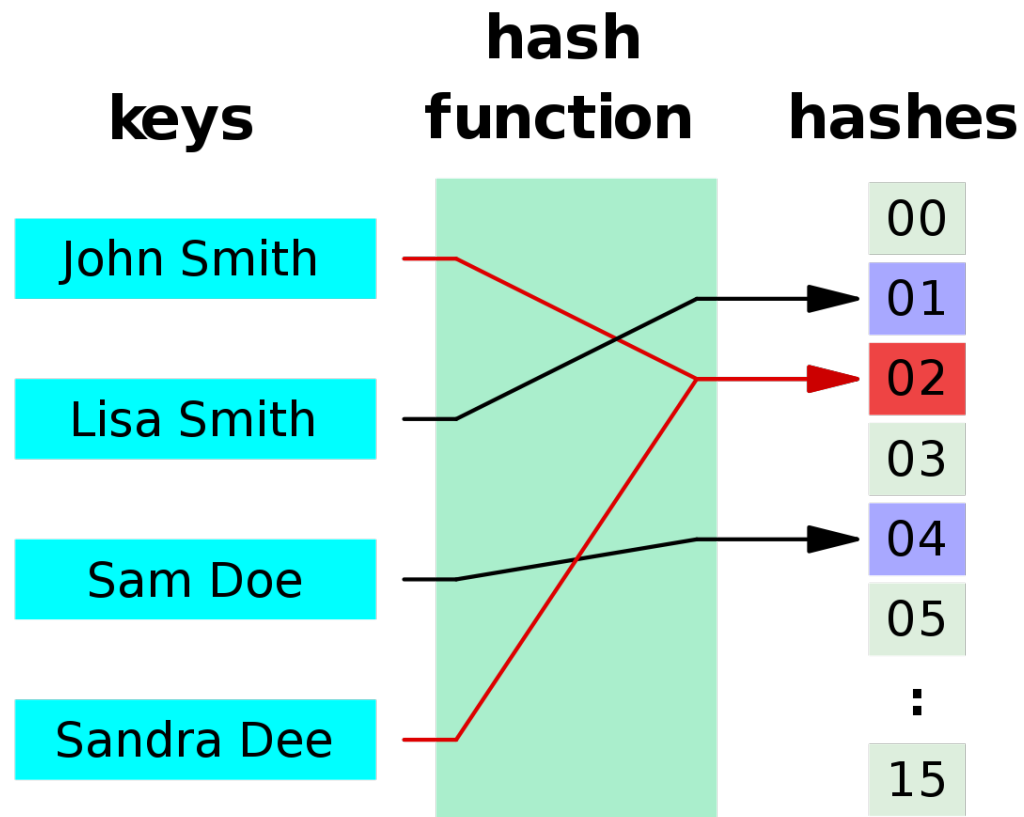


Tabelas de Espalhamento (*Hashing*)

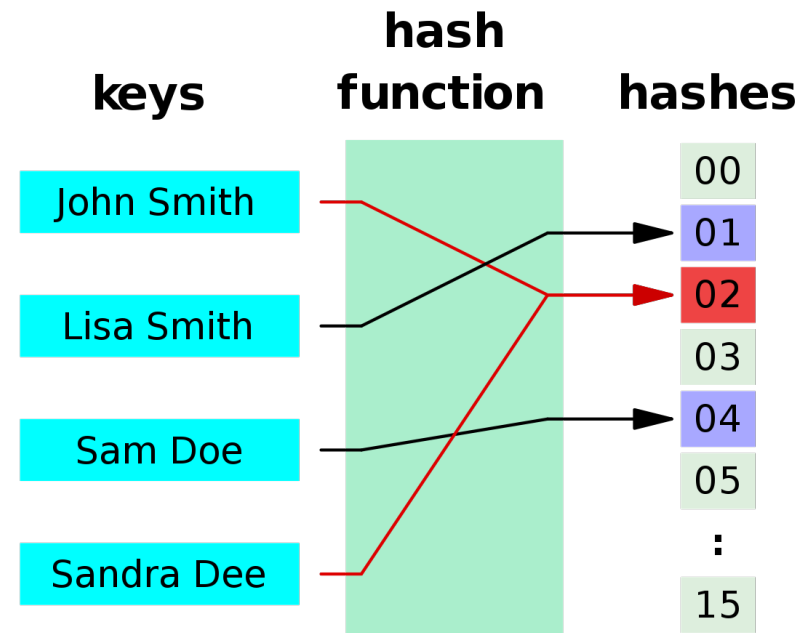
Hélder Seixas Lima
helder.seixas@ifnmg.edu.br

Qual a ideia?



Etapas

- Função de transformação (*hash*)
- Tratamento de colisões



Função *hash* ideal

- Simples de ser computada e barata
- Cada chave de saída deve ter $1/m$ chance de ser produzida, onde m é o tamanho da tabela *hash*

Método da divisão 1/2

- $h(k) = k \bmod m$,
 - Onde h é a função *hash*, k é a chave e m é o tamanho da tabela *hash*
- Evitar m que seja potência de 2
- m deve ser um primo distante de uma potência de 2



Método da divisão 2/2

Potência de 2
2
4
8
16
32
64
128
256
512
1024
...

Números primos menores que mil													
2	3	5	7	11	13	17	19	23	29	31	37	41	43
47	53	59	61	67	71	73	79	83	89	97	101	103	107
109	113	127	131	137	139	149	151	157	163	167	173	179	181
191	193	197	199	211	223	227	229	233	239	241	251	257	263
269	271	277	281	283	293	307	311	313	317	331	337	347	349
353	359	367	373	379	383	389	397	401	409	419	421	431	433
439	443	449	457	461	463	467	479	487	491	499	503	509	521
523	541	547	557	563	569	571	577	587	593	599	601	607	613
617	619	631	641	643	647	653	659	661	673	677	683	691	701
709	719	727	733	739	743	751	757	761	769	773	787	797	809
811	821	823	827	829	839	853	857	859	863	877	881	883	887
907	911	919	929	937	941	947	953	967	971	977	983	991	997

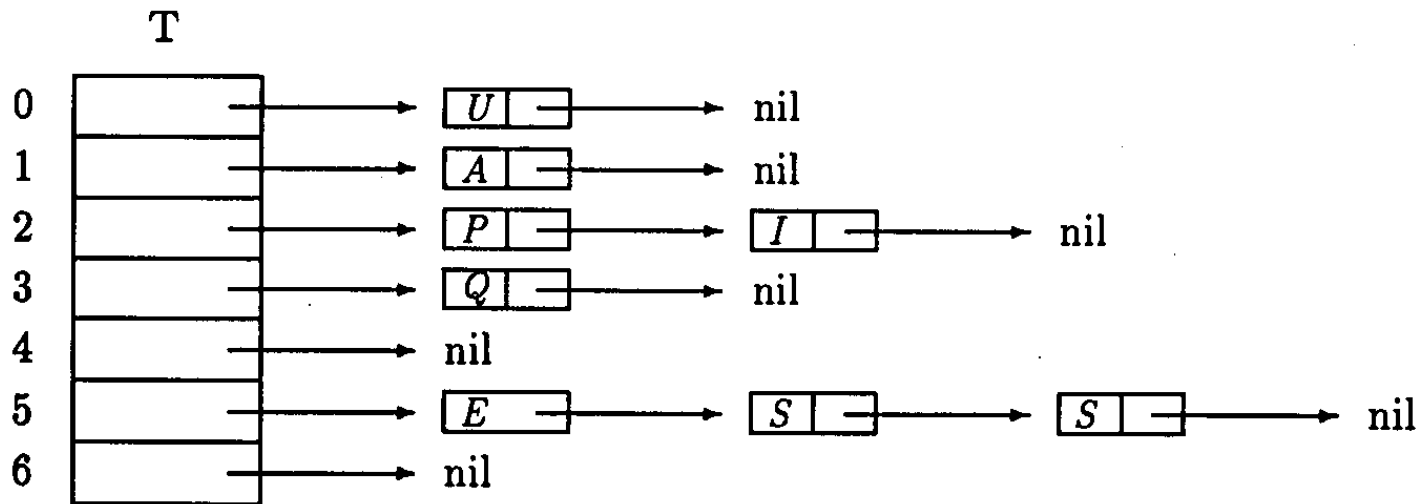
Método da multiplicação

- $h(k) = \lfloor m(kA \bmod 1) \rfloor$,
 - Onde h é a função *hash*, k é a chave, m é o tamanho da tabela *hash* e A é uma constante $0 < A < 1$
- m pode ser qualquer valor
- Sugere-se $A \approx (\sqrt{5} - 1) / 2 = 0,6180339887...$

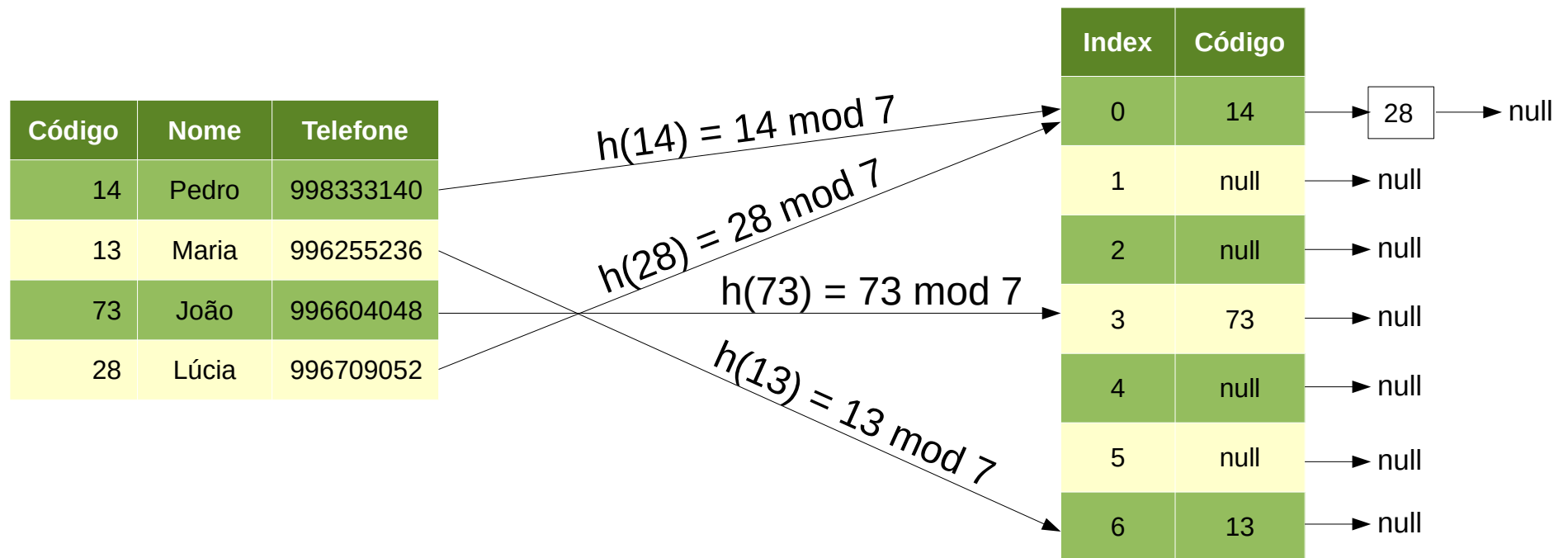
Mas minha chave é uma string?

- Codificação da tabela ASCII
 - **JOSÉ** $\rightarrow 74 + 79 + 83 + 45 = \mathbf{281}$
 - **MARIA** $\rightarrow 77 + 65 + 82 + 73 + 65 = \mathbf{362}$
- Multiplique a posição para evitar colisões por anagramas
 - **JOSÉ MARIA** $\rightarrow 1*74 + 2*79 + 3*83 + 4*45 + 5*77 + 6*65 + 7*82 + 8*73 + 9*65 = \mathbf{3179}$
 - **MARIA JOSÉ** $\rightarrow 1*77 + 2*65 + 3*82 + 4*73 + 5*65 + 6*74 + 7*79 + 8*83 + 9*45 = \mathbf{3136}$

Tratamento de colisões: listas encadeadas



Exemplo: clientes de uma loja



Análise: pesquisa, inserção e remoção

- Melhor caso: $O(1)$
- Pior caso: $O(n)$
- Caso médio: $O(1 + n/m) = O(1)$

$\underbrace{\hspace{10em}}$
Ziviani (2017)

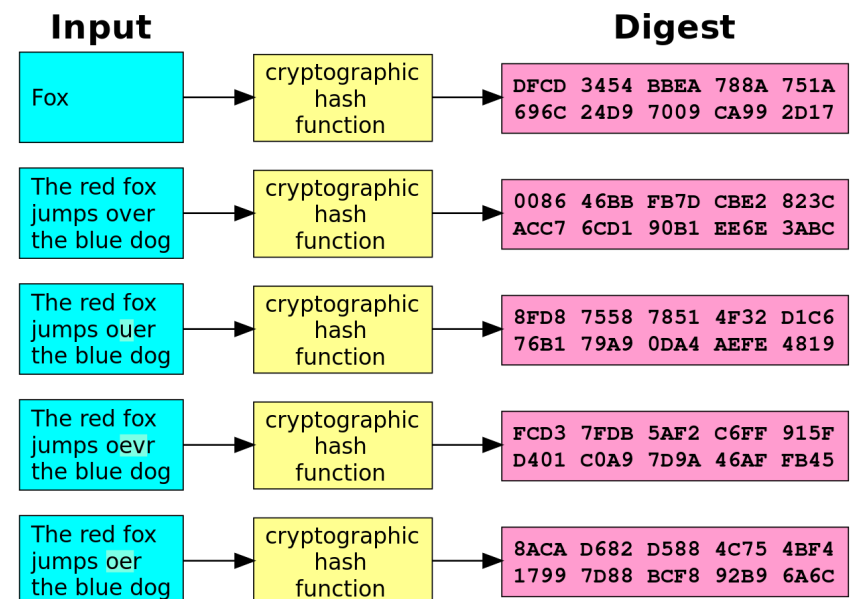
$\underbrace{\hspace{10em}}$
Cormem et al (2012)

Outros tópicos sobre *hash*

- *Hashing* universal
- Endereçamento aberto
- *Hashing* perfeito

Outras aplicações de *hashing*

- Dígito verificador
- Verificar integridade de dados
- Segurança de senhas
- Algoritmos MD5 e SHA256



Referências

- CORMEN, T.; LEISERSON, C.; RIVEST, R.; STEIN, C.. **Algoritmos**: teoria e prática. e. ed. Rio de Janeiro: Elsevier, 2012.
- ZIVIANI, N. **Projeto de Algoritmos**: com Implementação em Pascal e C. 3. ed. São Paulo: Cengage Learning, 2017.

Dúvidas

- Email
 - helder.seixas@ifnmg.edu.br
- Fóruns
- Monitoria

