Algorithms and Data Structures

Andrzej Pisarski

Plan of the lecture

1. Hash Tables
2. Hashing with Linear Probing
3. Hashing with Quadratinc Probing
4. Hashing with Double Hashing
5. Hashing with Separate Chaining
6. Efficiency

Hash Tables

The basic idea:

Take a record (key) and convert it using some function to numeric value (hash key)
some function = hash fuction

Example of hash fuction (division method):

hash key (index) = key % number of slots

Hash Tables

Example of hash fuction (division method):

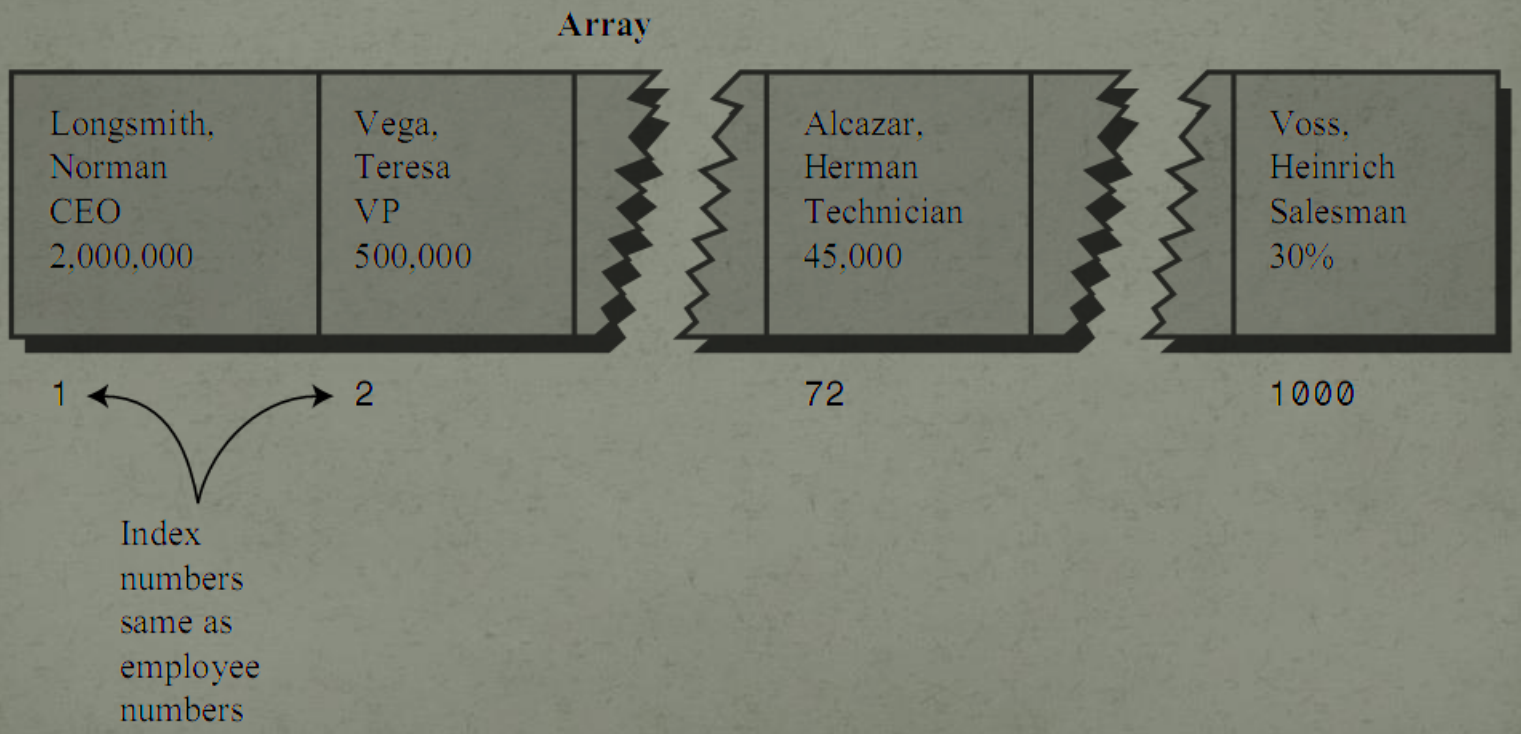
hash key (index) = key % number of slots

0	1	2	3	4	5	6	7	8
81	-1	20	39	40	-1	-1	7	-1

0 = 81 % 9
2 = 20 % 9
3 = 39 % 9
4 = 40 % 9
7 = 7 % 9

Hash Tables

Some other examples of keys:
- employee numbers



(*example from R. Lafore book*)

Hash Tables

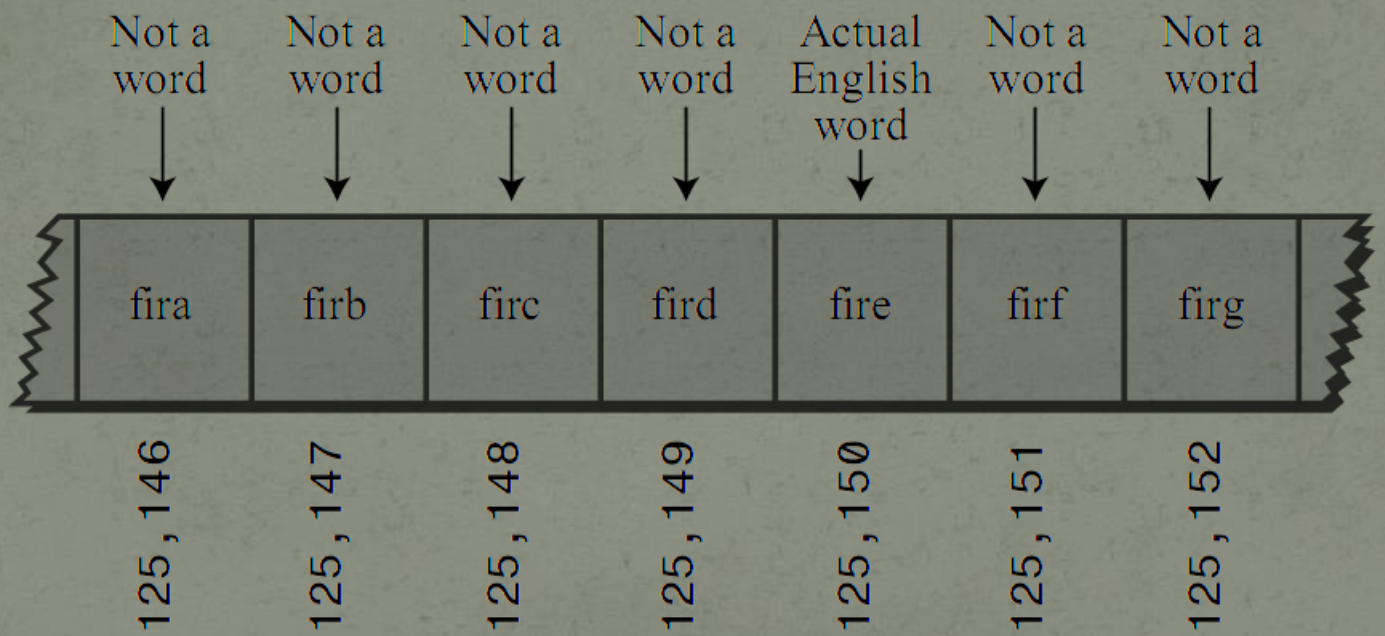
Some other examples of keys:
- employee numbers,
- words (how to convert words to hash key?)

Sign	Code page ANSI ASCII		
	binary	decimal	aplhabet index
c	01100011	99	3
a	01100001	97	1
t	01110100	116	20
s	01110011	115	19

5467 = 5*10^3 + 4*10^2 + 6*10^1 + 7*10^0
cats = 3*27^3 + 1*27^2 + 20*27^1 + 19*27^0 = 60337

Hash Tables

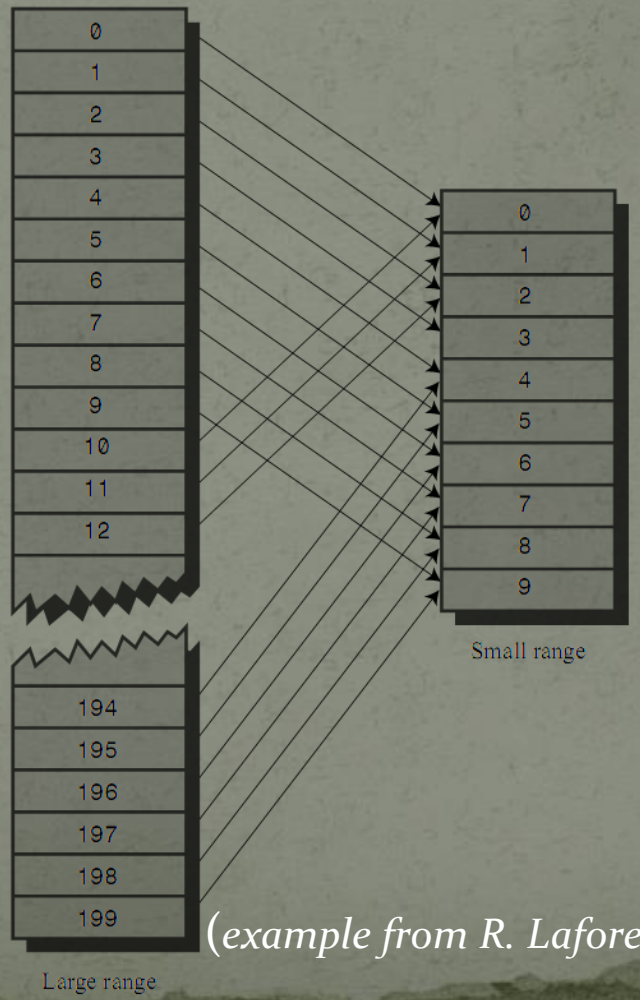
Every potential word from „a” to „zzzzzzzzzz” (10 „z”)
requaire array size:
26*27^9 + 26*27^8 + … + 26*27^1 + 26*27^0 > 26*27^9 = 198 * 10^12



(*example from R. Lafore book*)

Hash Tables

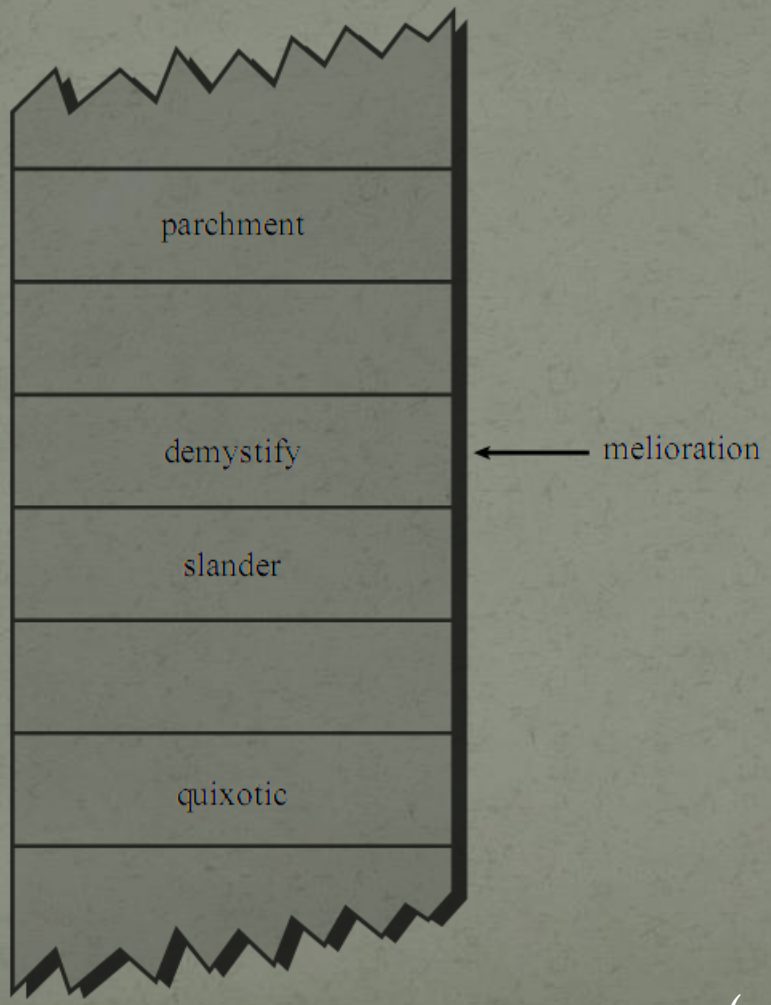
hugeNumber = 26*27^9 + 26*27^8 + … + 26*27^1 + 26*27^0
arraySize = numberWords * 2;
arrayIndex = hugeNumber % arraySize;
(hash function)



(*example from R. Lafore book*)

Hash Tables

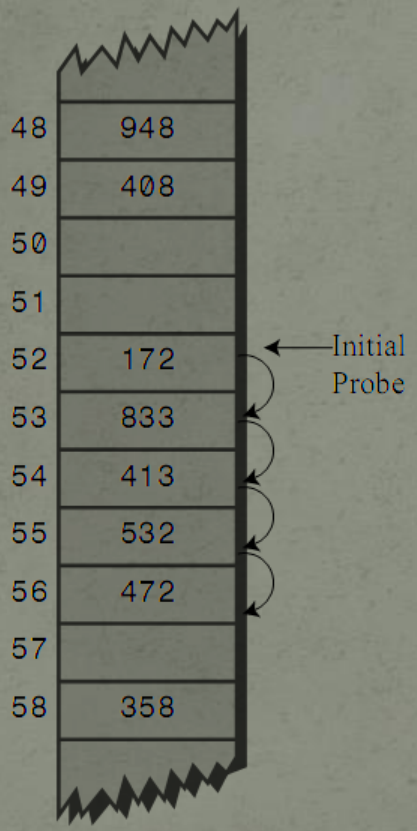
Collisions



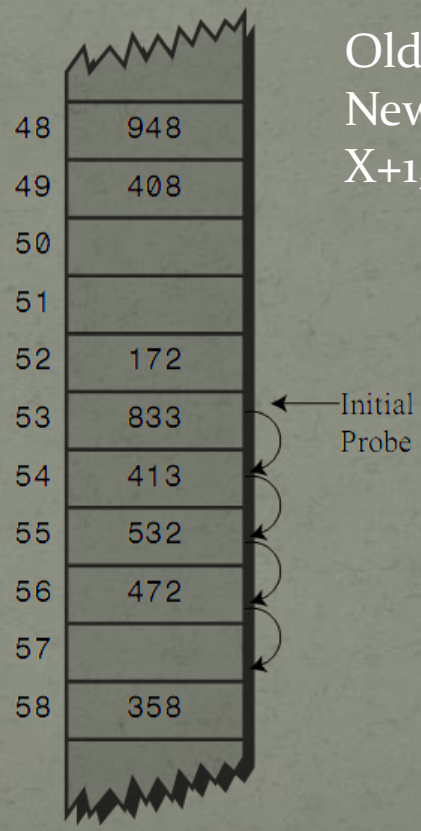
(*example from R. Lafore book*)

Hashing with Linear Probing

Collisions : inserting and finding



a) Successful search for 472



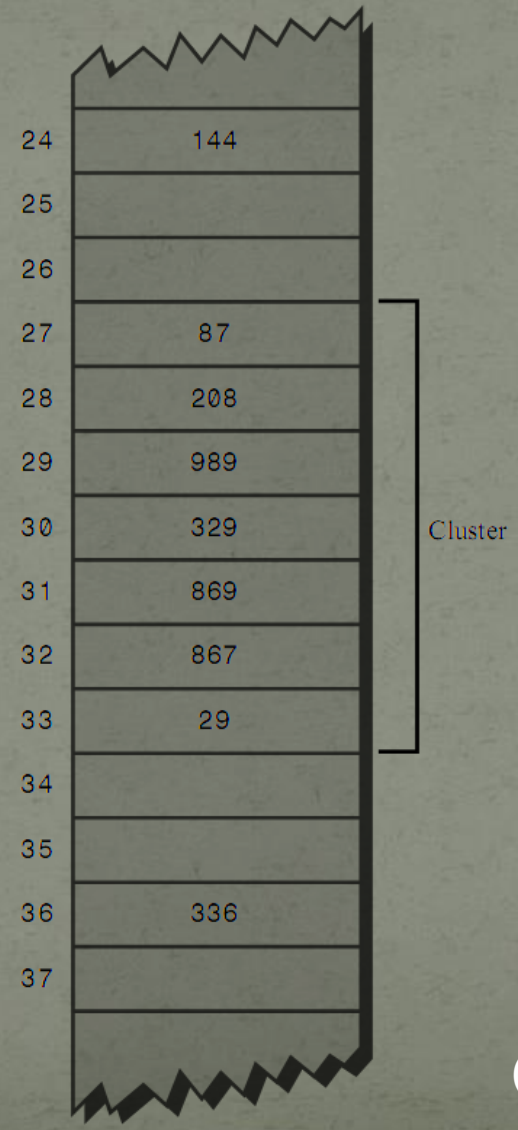
b) Unsuccessful search for 893

Old hash key = X
New hash key:
X+1, X+2, X+3, …, X+5

(*example from R. Lafore book*)

Hashing with Linear Probing

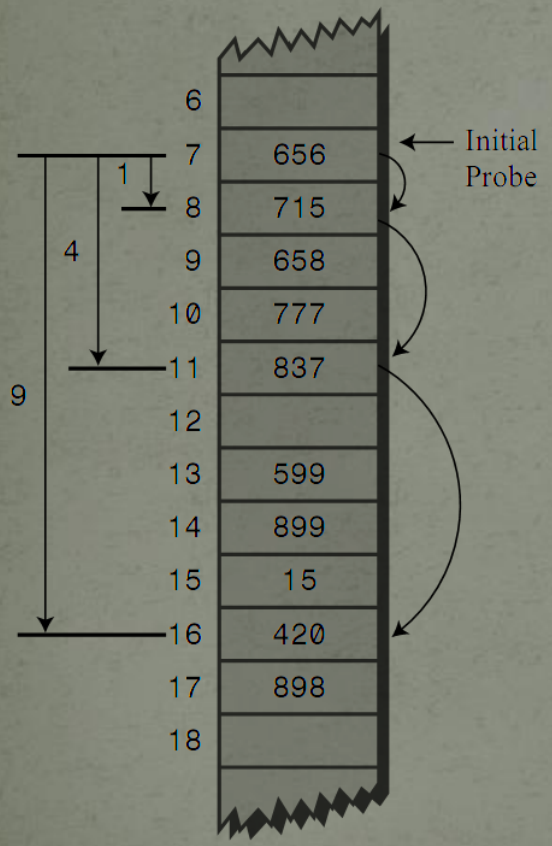
Clustering



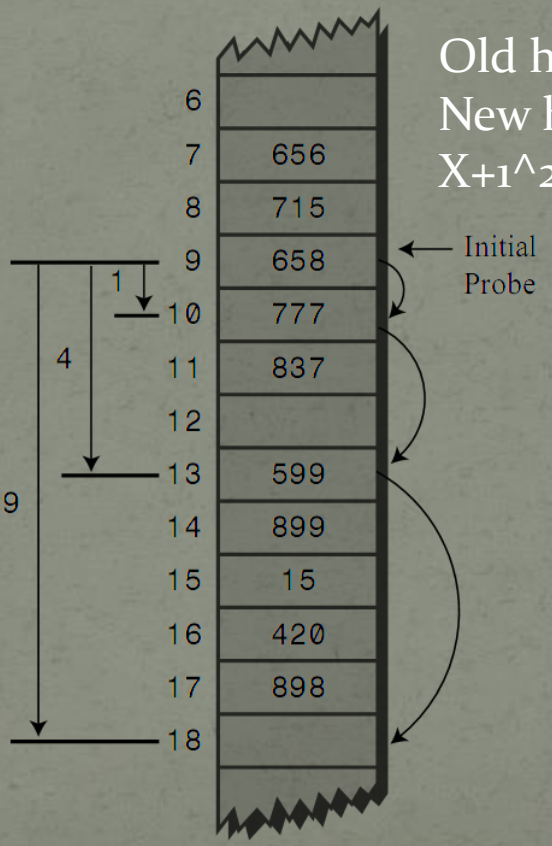
(*example from R. Lafore book*)

Hashing with Quadratic Probing

Quadrating Probing



a) Successful search for 420



b) Unsuccessful search for 481

Old hash key = X
New hash key:
X+1^2, X+2^2, X+3^2, …, X+5^2

(*example from R. Lafore book*)

Hashing with Quadratic Probing

Quadrating Probing why to use?
• solve problem with primary clustering in linear probe

Can cause problem for keys which are hashing function
return this same hash key

Example: Let's say for 184, 302, 420 and 544 hash function return 7.
Then 302 will require a 1-step probe, 420 a 2-step-probe ,
544 will require 3-step probe. This is a secondary clustering.

To sole problem with secondary clustering as well primary can be used:
double hashing.

(*example from R. Lafore book*)

Hashing with Double Hashing

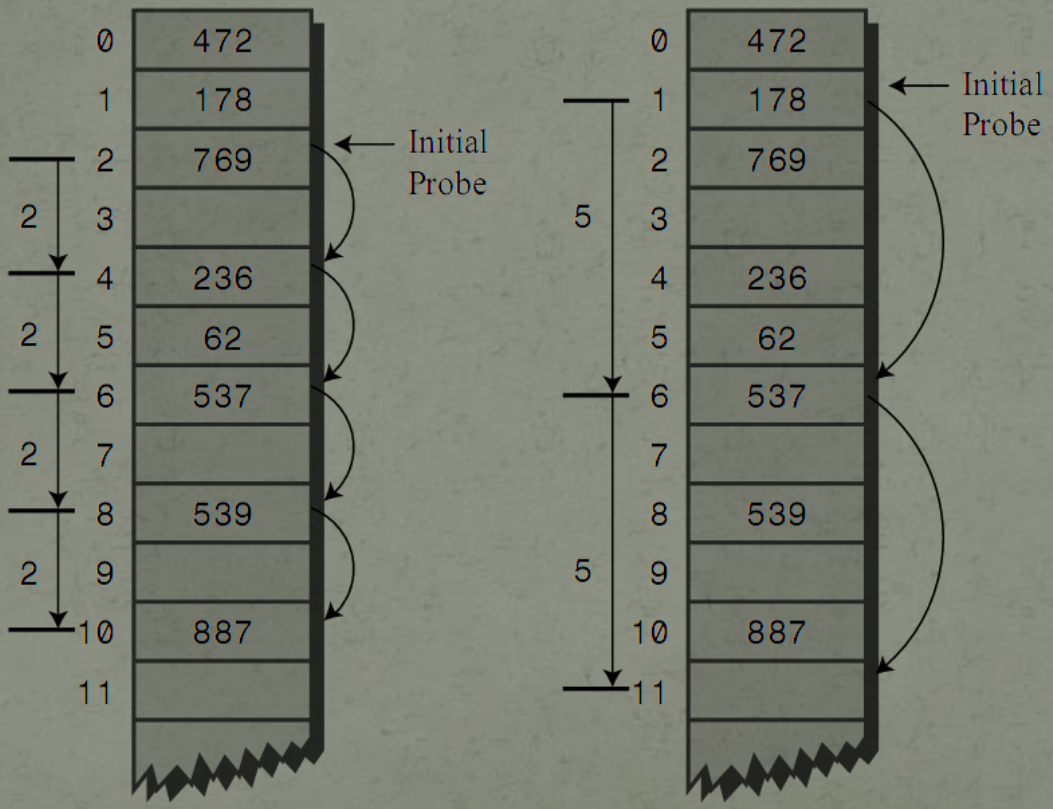
The basic idea is to use a **different** hash function a second time :
stepSize = constant – (key % constant);

constant = prime number less than an array size; example:
stepSize=5 – (key % 5);

(*example from R. Lafore book*)

Hashing with Double Hashing

The basic idea is to use a **different** hash function a second time :



(*example from R. Lafore book*)

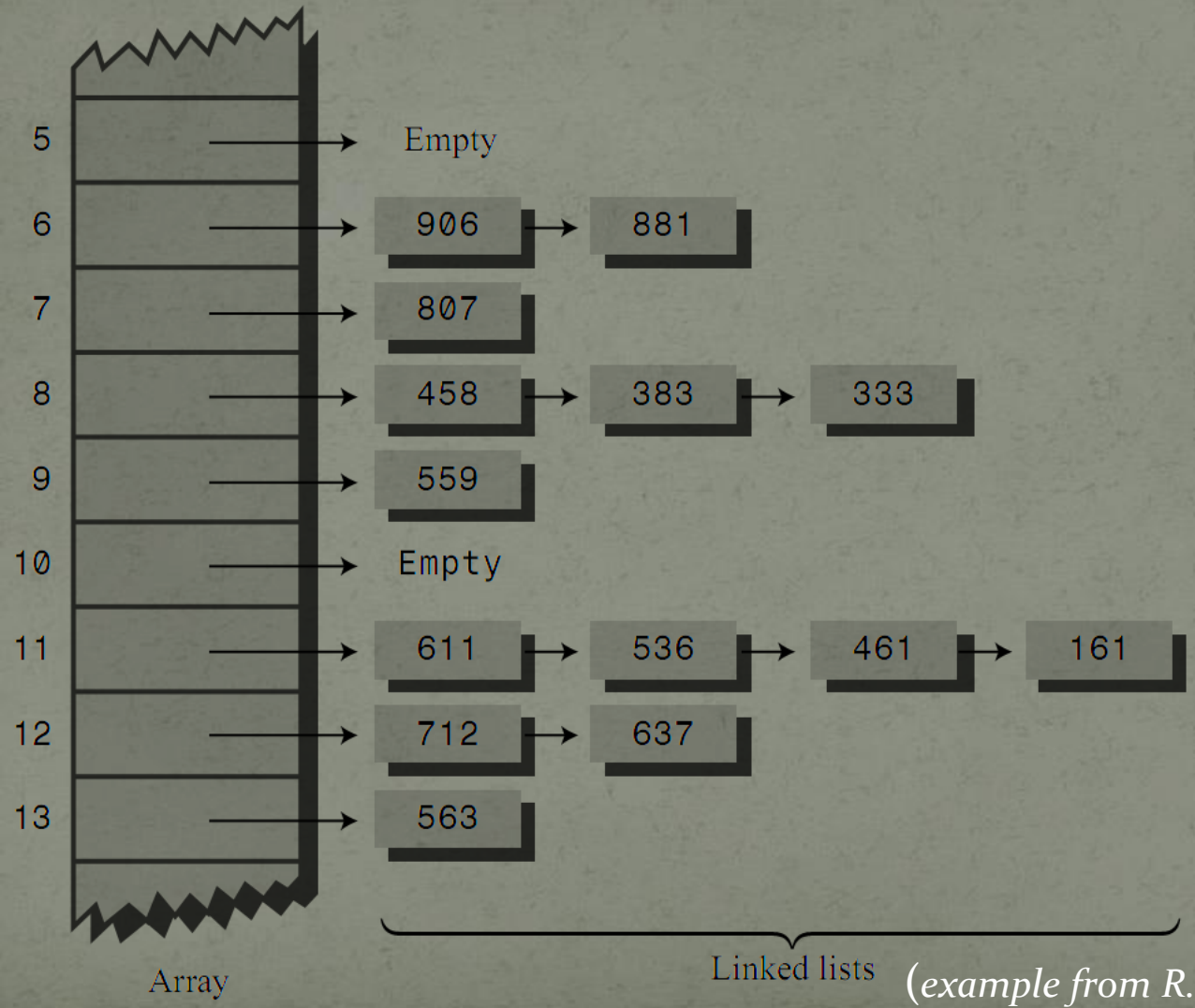
Hashing with Double Hashing

Double hashing requires an array size to be a prime number.
Example1 : arraySize = 15 (0, 1, 2,…,12, 13, 14), stepSize = 5
Intial index (hash key of particular key) = 0;
Probe sequence: 0, 5, 10, 0, 5, 10, …

Example2 : arraySize = 13 (0, 1, 2,…,10, 11, 12), stepSize = 5
Intial index (hash key of particular key) = 0;
Probe sequence: 0, 5, 10, 2, 7, 12, 4, 9, 1, 6, 11, 3, … (*all indexes* of the array).

(*example from R. Lafore book*)

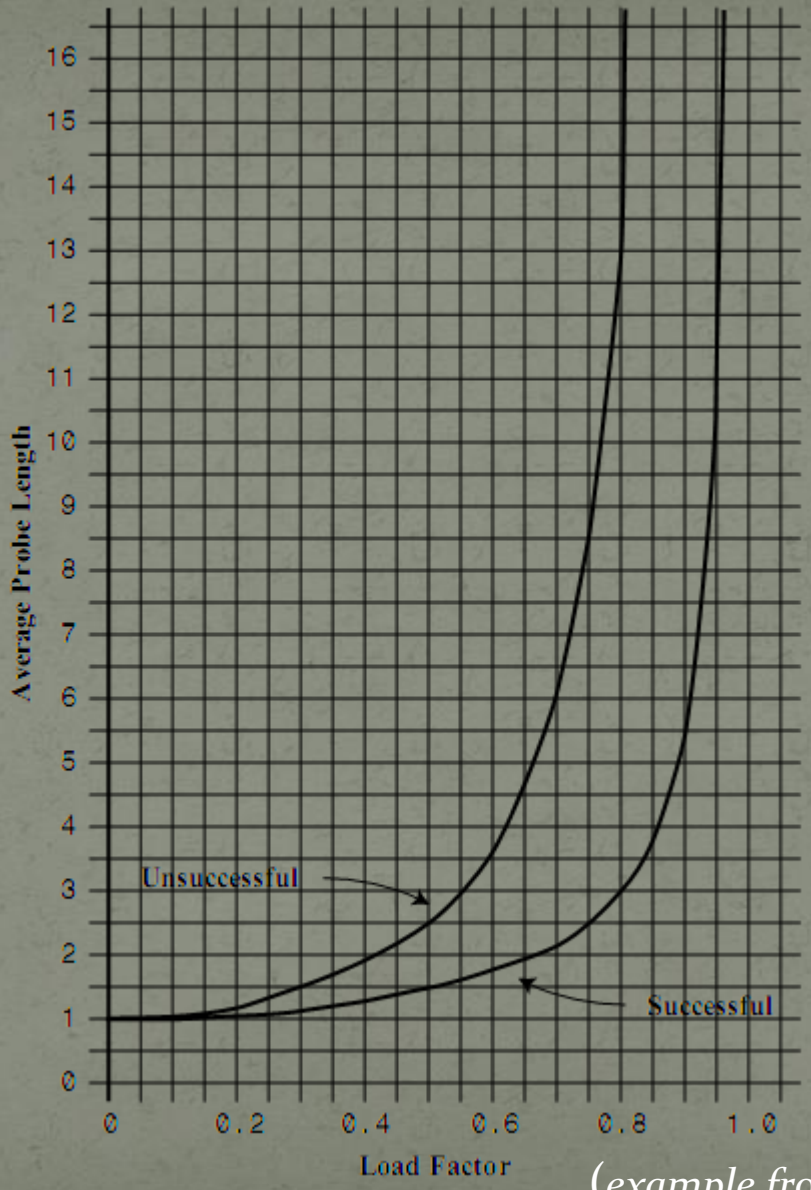
Hashing with Separate Chaining



(*example from R. Lafore book*)

Efficiency

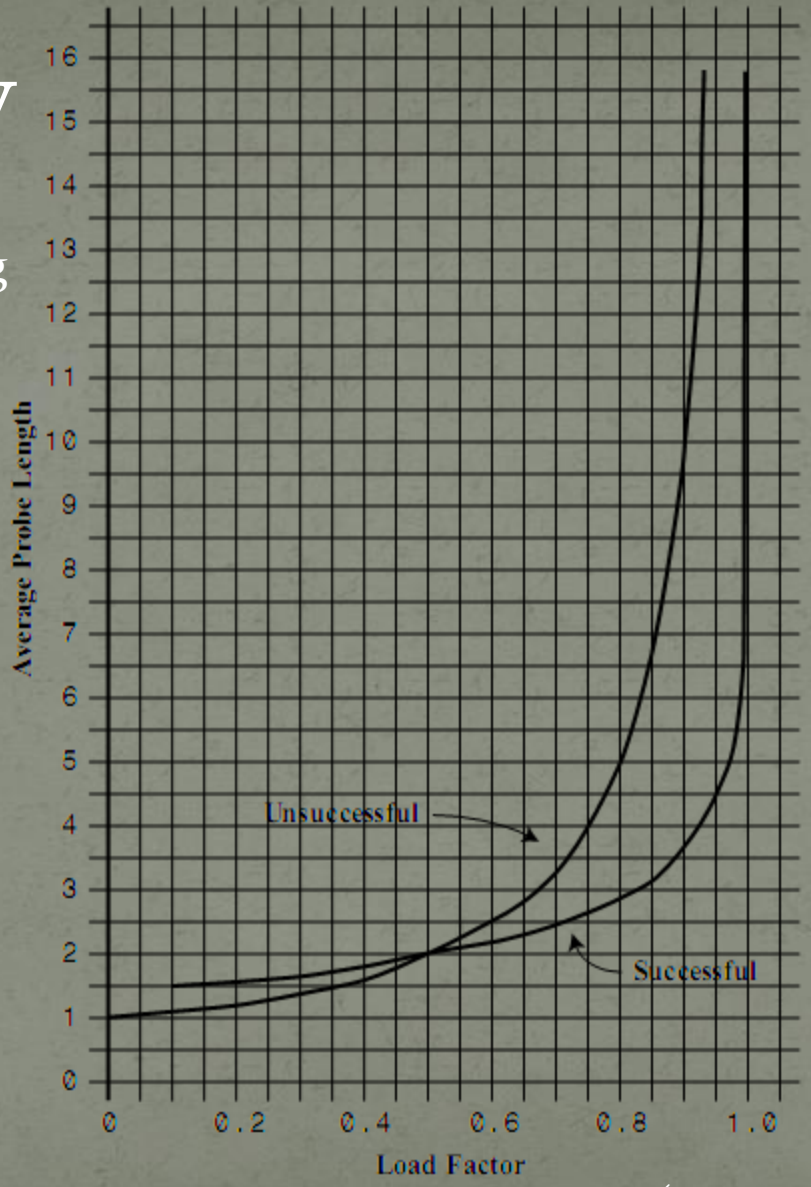
Linear Probing



(*example from R. Lafore book*)

Efficiency

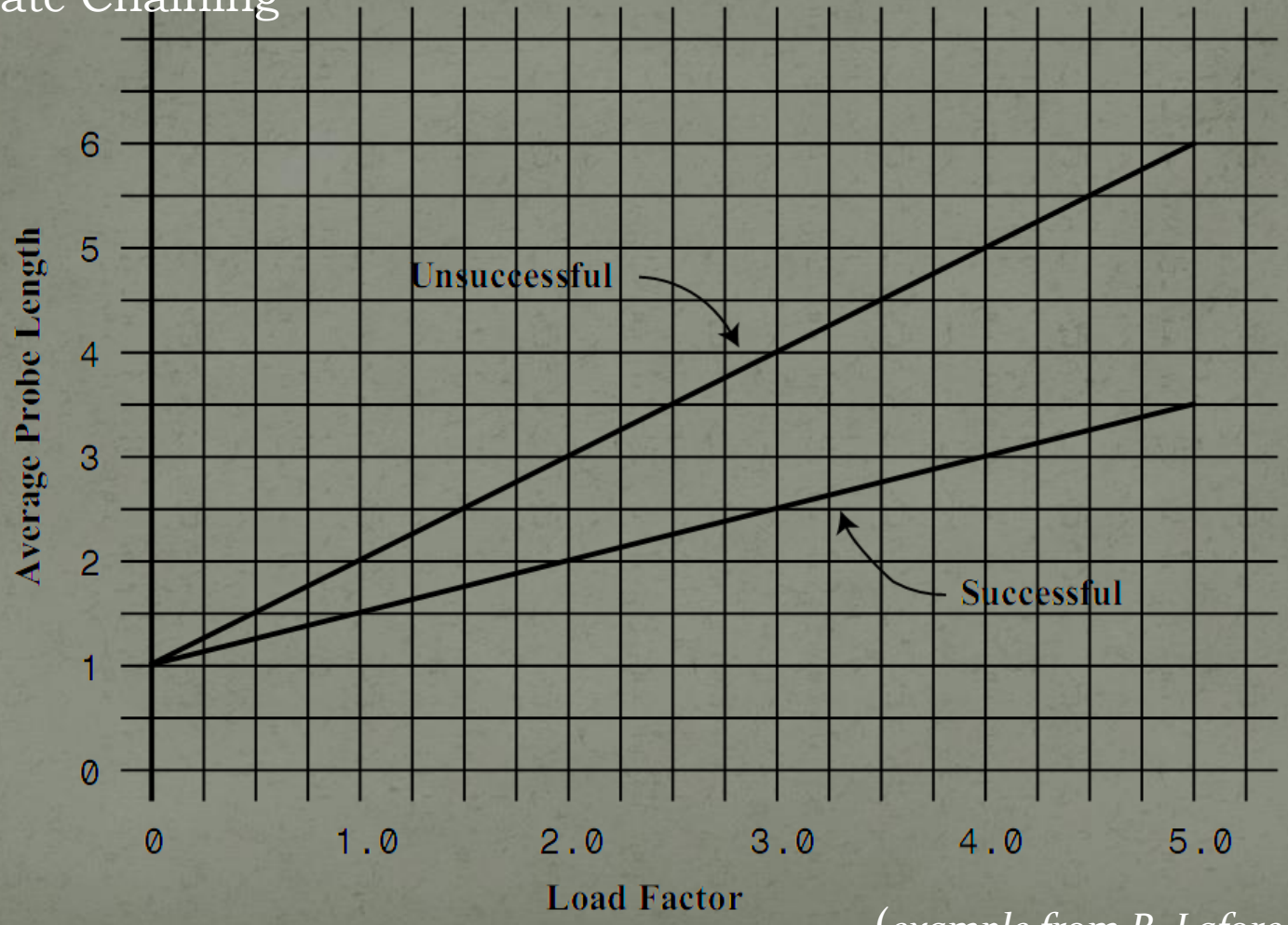
Quadratic Probing
and Double Hashing



(*example from R. Lafore book*)

Efficiency

Separate Chaining



(*example from R. Lafore book*)