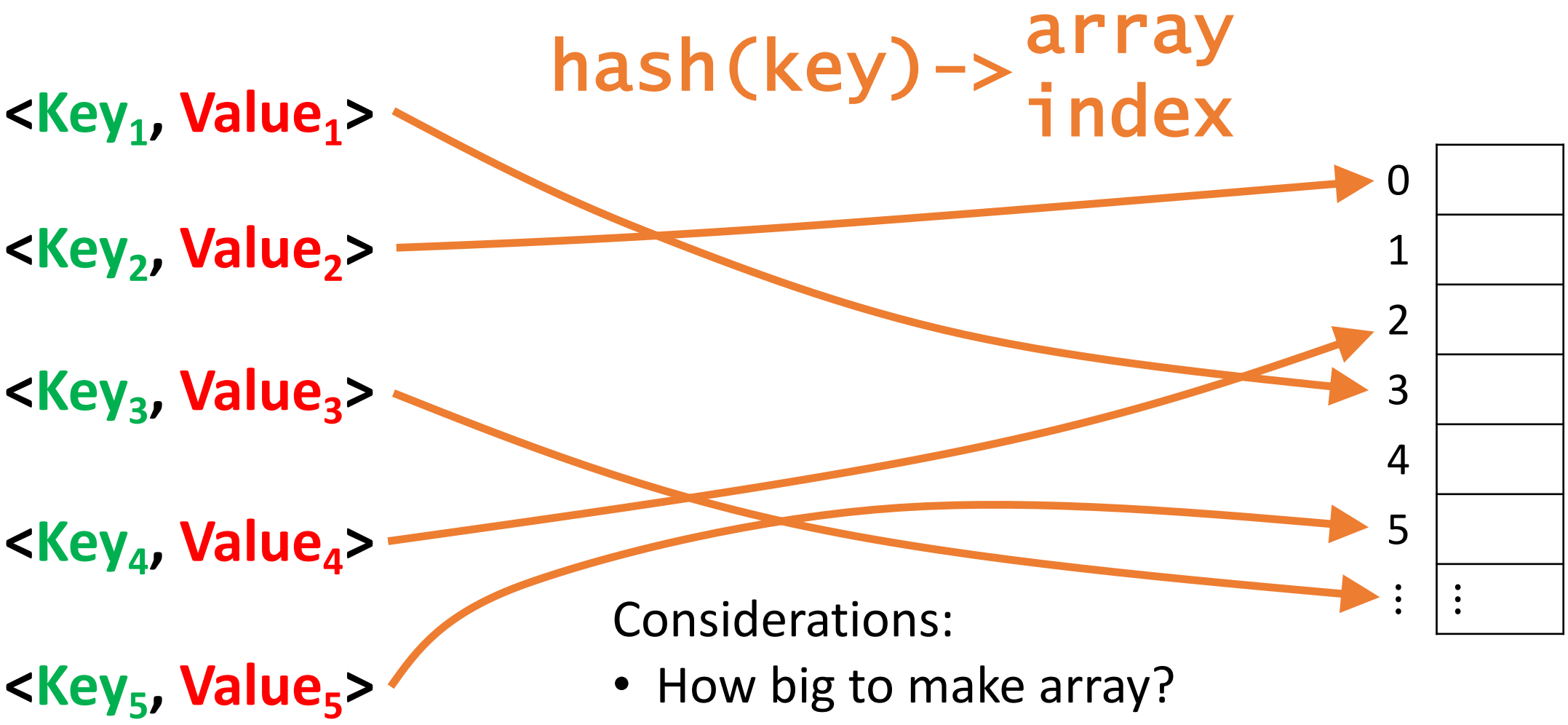
Hash Functions
CSCI 232

Hash Tables 101

Theory.
Hash Functions.
Statistical likelihood.
Expected performance.

Application.
Tools in hand.
Java functionality.

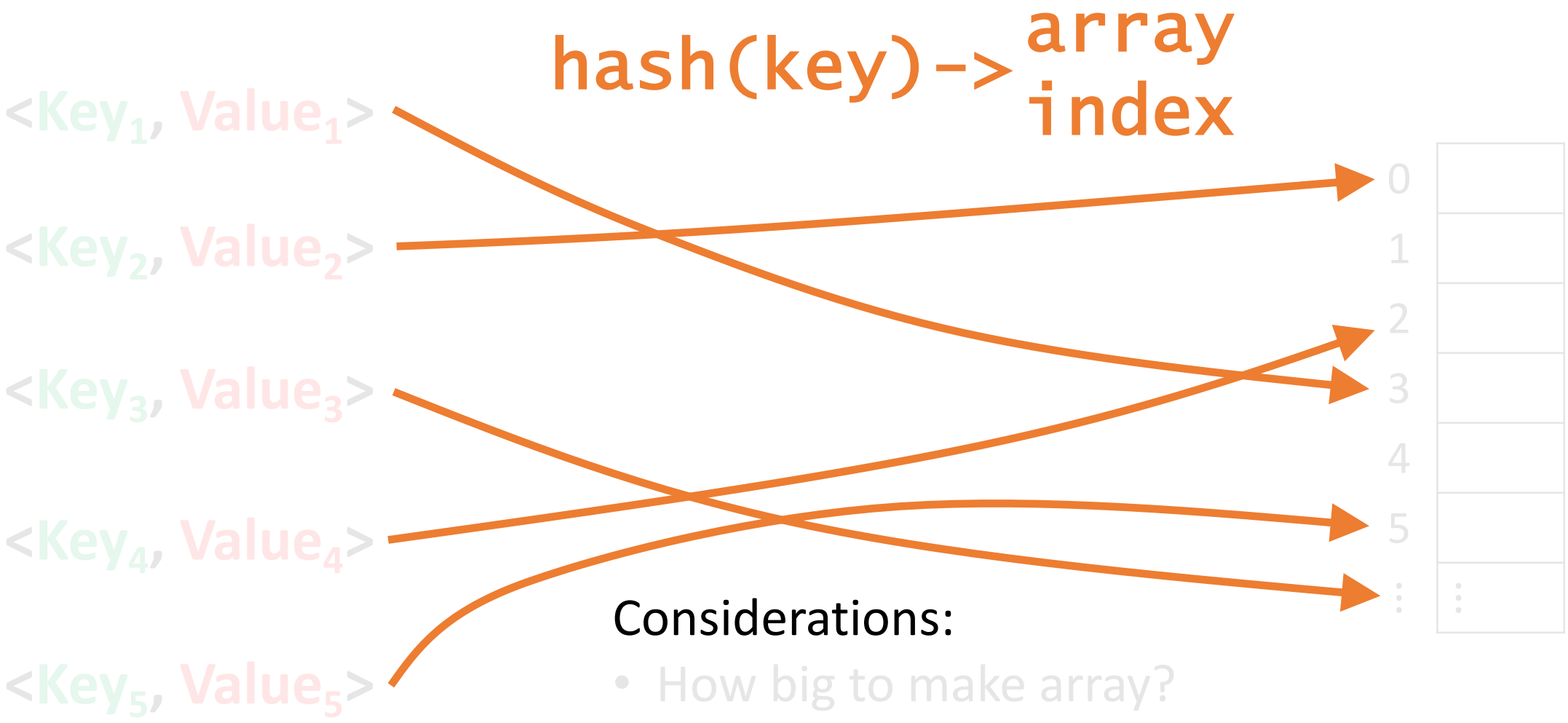
Hash Tables 101



Considerations:

- How big to make array?
- How to avoid collisions?
- How to handle collisions?

Hash Tables 101



Considerations:

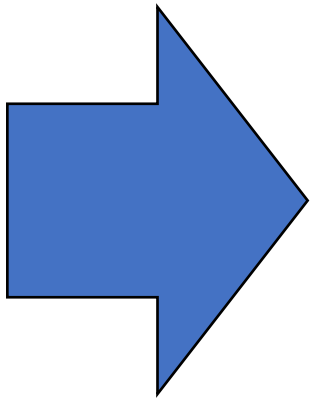
- How big to make array?
- How to avoid collisions?
- How to handle collisions?

Hash Function

Function that translates key values into array indices.

hash(key) = key % 100

Key	Value
100	Flora's Flowers,…
101	The Bank,…
102	Ted's Drug Store,…
103	Joe's Gas Station,…
⋮	⋮
198	Aristocrat Art,…
199	Fire Station,…
200	More Art,…



0	100, Flora's Flowers,…
1	101, The Bank,…
2	102, Ted's Drug Store,…
3	103, Joe's Gas Station,…
⋮	⋮
98	198, Aristocrat Art,…
99	199, Fire Station,…

Hash Function

Hash Function:

Lessons:

Keys	Hash Value
516-07-0854	
516-66-6218	
531-01-7352	
336-82-2121	
517-90-7152	
516-98-8002	
517-45-0907	
531-81-7489	
517-07-7312	
516-24-6185	
669-44-6499	
516-16-4236	
530-92-1795	
611-52-4556	
516-34-3352	
607-86-0812	

Hash Function

Hash Function:

$$\left(\begin{matrix}\text{First three}\\ \text{digits}\end{matrix}\right) \% \ 100$$

Lessons:

Keys	Hash Value
516-07-0854	
516-66-6218	
531-01-7352	
336-82-2121	
517-90-7152	
516-98-8002	
517-45-0907	
531-81-7489	
517-07-7312	
516-24-6185	
669-44-6499	
516-16-4236	
530-92-1795	
611-52-4556	
516-34-3352	
607-86-0812	

Hash Function

Hash Function:

(First three digits) % 100

Called Modular Hashing

Size of the array

Lessons:

Keys	Hash Value
516-07-0854	
516-66-6218	
531-01-7352	
336-82-2121	
517-90-7152	
516-98-8002	
517-45-0907	
531-81-7489	
517-07-7312	
516-24-6185	
669-44-6499	
516-16-4236	
530-92-1795	
611-52-4556	
516-34-3352	
607-86-0812	

Hash Function

Hash Function:

$$\left(\begin{matrix}\text{First three}\\ \text{digits}\end{matrix}\right) \% \ 100$$

Lessons:

Keys	Hash Value
516-07-0854	??
516-66-6218	
531-01-7352	
336-82-2121	
517-90-7152	
516-98-8002	
517-45-0907	
531-81-7489	
517-07-7312	
516-24-6185	
669-44-6499	
516-16-4236	
530-92-1795	
611-52-4556	
516-34-3352	
607-86-0812	

Hash Function

Hash Function:

$$\left(\begin{matrix}\text{First three}\\ \text{digits}\end{matrix}\right) \% \ 100$$

Lessons:

Keys	Hash Value
516-07-0854	16
516-66-6218	??
531-01-7352	
336-82-2121	
517-90-7152	
516-98-8002	
517-45-0907	
531-81-7489	
517-07-7312	
516-24-6185	
669-44-6499	
516-16-4236	
530-92-1795	
611-52-4556	
516-34-3352	
607-86-0812	

Hash Function

Hash Function:

$$\left(\begin{matrix}\text{First three}\\ \text{digits}\end{matrix}\right) \% \ 100$$

Lessons:

Keys	Hash Value
516-07-0854	16
516-66-6218	16
531-01-7352	31
336-82-2121	36
517-90-7152	17
516-98-8002	16
517-45-0907	17
531-81-7489	31
517-07-7312	17
516-24-6185	16
669-44-6499	69
516-16-4236	16
530-92-1795	30
611-52-4556	11
516-34-3352	16
607-86-0812	07

Hash Function

Hash Function:

$$\left(\begin{matrix}\text{First three}\\ \text{digits}\end{matrix}\right) \% \ 100$$

Lessons:

What is the problem?
Why is the problem occurring?
How can we fix it?

Keys	Hash Value
516-07-0854	16
516-66-6218	16
531-01-7352	31
336-82-2121	36
517-90-7152	17
516-98-8002	16
517-45-0907	17
531-81-7489	31
517-07-7312	17
516-24-6185	16
669-44-6499	69
516-16-4236	16
530-92-1795	30
611-52-4556	11
516-34-3352	16
607-86-0812	07

Hash Function

Hash Function:

$$\left(\begin{matrix}\text{First three}\\ \text{digits}\end{matrix}\right) \% \ 100$$

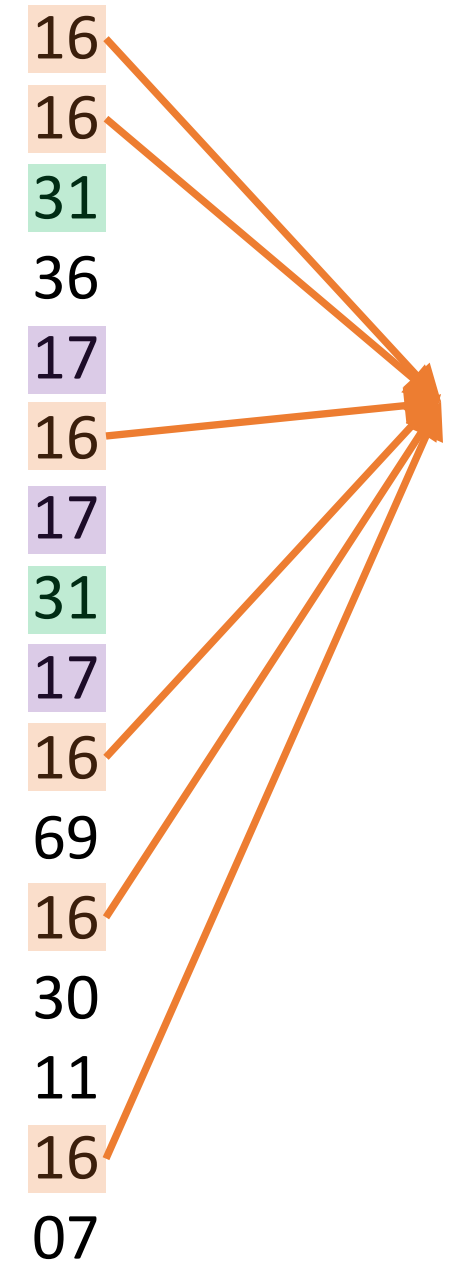
Lessons:

What is the problem?

Why is the problem occurring?

How can we fix it?

Keys	Hash Value
516-07-0854	16
516-66-6218	16
531-01-7352	31
336-82-2121	36
517-90-7152	17
516-98-8002	16
517-45-0907	17
531-81-7489	31
517-07-7312	17
516-24-6185	16
669-44-6499	69
516-16-4236	16
530-92-1795	30
611-52-4556	11
516-34-3352	16
607-86-0812	07



Hash Function

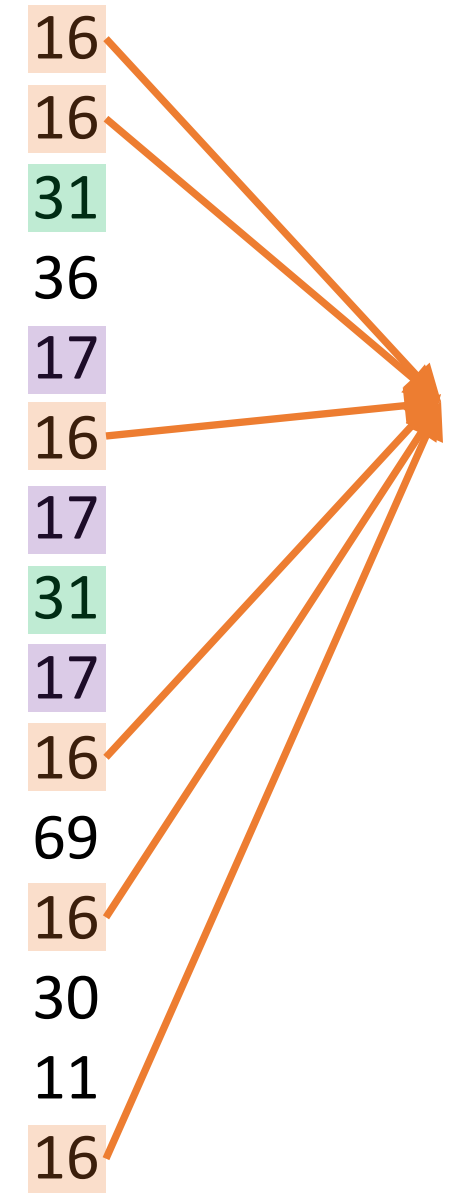
Hash Function:

$$\left(\begin{matrix}\text{First three}\\ \text{digits}\end{matrix}\right) \% \ 100$$

Lessons:

1. Use as much of the key as possible.

Keys	Hash Value
516-07-0854	16
516-66-6218	16
531-01-7352	31
336-82-2121	36
517-90-7152	17
516-98-8002	16
517-45-0907	17
531-81-7489	31
517-07-7312	17
516-24-6185	16
669-44-6499	69
516-16-4236	16
530-92-1795	30
611-52-4556	11
516-34-3352	16
607-86-0812	07



Hash Function

Hash Function:

$$\left(\begin{matrix}\text{Nine-digit}\\ \text{number}\end{matrix}\right) \% \ 5$$

Lessons:

1. Use as much of the key as possible.

Keys	Hash Value
516-07-0854	??
516-66-6218	
531-01-7352	
336-82-2121	
517-90-7152	
516-98-8002	
517-45-0907	
531-81-7489	
517-07-7312	
516-24-6185	
669-44-6499	
516-16-4236	
530-92-1795	
611-52-4556	
516-34-3352	
607-86-0812	

Hash Function

Hash Function:

$$\left(\begin{matrix}\text{Nine-digit}\\ \text{number}\end{matrix}\right) \% \ 5$$

Lessons:

1. Use as much of the key as possible.

Keys	Hash Value
516-07-0854	4
516-66-6218	3
531-01-7352	2
336-82-2121	1
517-90-7152	2
516-98-8002	2
517-45-0907	2
531-81-7489	4
517-07-7312	2
516-24-6185	0
669-44-6499	4
516-16-4236	1
530-92-1795	0
611-52-4556	1
516-34-3352	2
607-86-0812	2

Hash Function

Hash Function:

$$\left(\begin{array}{c}\text{Nine-digit} \\ \text{number}\end{array}\right) \% \ 5$$

Lessons:

1. Use as much of the key as possible.

Keys	Hash Value
516-07-0854	4
516-66-6218	3
531-01-7352	2
336-82-2121	1
517-90-7152	2
516-98-8002	2
517-45-0907	2
531-81-7489	4
517-07-7312	2
516-24-6185	0
669-44-6499	4
516-16-4236	1
530-92-1795	0
611-52-4556	1
516-34-3352	2
607-86-0812	2

Hash Function

Hash Function:

(Nine-digit number) % 5

Lessons:

1. Use as much of the key as possible.
2. Make sure your array is big enough (but not too big).

Keys	Hash Value
516-07-0854	4
516-66-6218	3
531-01-7352	2
336-82-2121	1
517-90-7152	2
516-98-8002	2
517-45-0907	2
531-81-7489	4
517-07-7312	2
516-24-6185	0
669-44-6499	4
516-16-4236	1
530-92-1795	0
611-52-4556	1
516-34-3352	2
607-86-0812	2

Hash Function

Hash Function:

$$\left(\begin{matrix}\text{Nine-digit}\\ \text{number}\end{matrix}\right) \% \ 100$$

Lessons:

1. Use as much of the key as possible.
2. Make sure your array is big enough (but not too big).

Keys	Hash Value
516-07-0854	??
516-66-6218	
531-01-7352	
336-82-2121	
517-90-7152	
516-98-8002	
517-45-0907	
531-81-7489	
517-07-7312	
516-24-6185	
669-44-6499	
516-16-4236	
530-92-1795	
611-52-4556	
516-34-3352	
607-86-0812	

Hash Function

Hash Function:

$$\left(\begin{matrix}\text{Nine-digit}\\ \text{number}\end{matrix}\right) \% \ 100$$

Lessons:

1. Use as much of the key as possible.
2. Make sure your array is big enough (but not too big).

Keys	Hash Value
516-07-0854	54
516-66-6218	??
531-01-7352	
336-82-2121	
517-90-7152	
516-98-8002	
517-45-0907	
531-81-7489	
517-07-7312	
516-24-6185	
669-44-6499	
516-16-4236	
530-92-1795	
611-52-4556	
516-34-3352	
607-86-0812	

Hash Function

Hash Function:

$$\left(\begin{matrix}\text{Nine-digit}\\ \text{number}\end{matrix}\right) \% \ 100$$

Lessons:

1. Use as much of the key as possible.
2. Make sure your array is big enough (but not too big).

Keys	Hash Value
516-07-0854	54
516-66-6218	18
531-01-7352	52
336-82-2121	21
517-90-7152	52
516-98-8002	02
517-45-0907	07
531-81-7489	89
517-07-7312	12
516-24-6185	85
669-44-6499	99
516-16-4236	36
530-92-1795	95
611-52-4556	56
516-34-3352	52
607-86-0812	12

Hash Function

Hash Function:

$$\left(\begin{matrix}\text{Nine-digit}\\ \text{number}\end{matrix}\right) \% \ 100$$

Lessons:

1. Use as much of the key as possible.
2. Make sure your array is big enough (but not too big).

Keys	Hash Value
516-07-0854	54
516-66-6218	18
531-01-7352	52
336-82-2121	21
517-90-7152	52
516-98-8002	02
517-45-0907	07
531-81-7489	89
517-07-7312	12
516-24-6185	85
669-44-6499	99
516-16-4236	36
530-92-1795	95
611-52-4556	56
516-34-3352	52
607-86-0812	12

Hash Function

Hash Function:

$$\left(\begin{matrix}\text{Nine-digit}\\ \text{number}\end{matrix}\right) \% \ 100$$

Lessons:

1. Use as much of the key as possible.
2. Make sure your array is big enough (but not too big).
3. Use a prime number array size.

Keys	Hash Value
516-07-0854	54
516-66-6218	18
531-01-7352	52
336-82-2121	21
517-90-7152	52
516-98-8002	02
517-45-0907	07
531-81-7489	89
517-07-7312	12
516-24-6185	85
669-44-6499	99
516-16-4236	36
530-92-1795	95
611-52-4556	56
516-34-3352	52
607-86-0812	12

Hash Function

Hash Function:

$$\left(\begin{matrix}\text{Nine-digit}\\ \text{number}\end{matrix}\right) \% \ 97$$

Lessons:

1. Use as much of the key as possible.
2. Make sure your array is big enough (but not too big).
3. Use a prime number array size.

Keys	Hash Value
516-07-0854	8
516-66-6218	83
531-01-7352	67
336-82-2121	0
517-90-7152	96
516-98-8002	21
517-45-0907	42
531-81-7489	51
517-07-7312	91
516-24-6185	60
669-44-6499	29
516-16-4236	76
530-92-1795	55
611-52-4556	84
516-34-3352	33
607-86-0812	30

Hash Function

Function that translates key values into array indices.

Requirements:

- **Well defined for all input.**
- **Identical keys map to identical indices.**
- **Maps to a specified range.**

Goals:

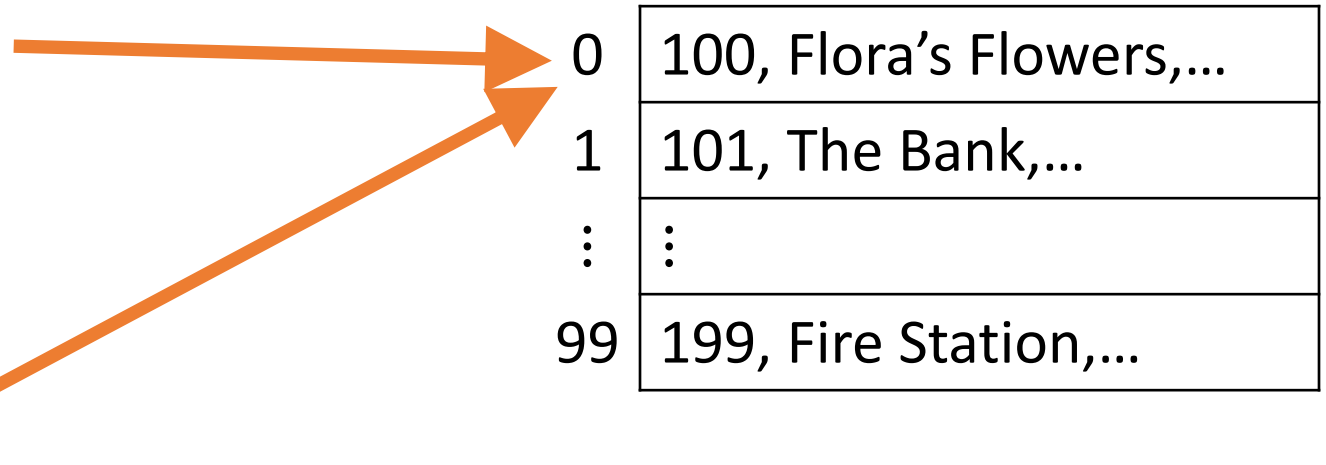
- **Easy to compute.**
- **Uniformly distributes keys across range (collision avoidance).**

Collisions

Collision Avoidance = Hash Function

Collision Resolution = ?

Key	Value
100	Flora’s Flowers,…
101	The Bank,…
⋮	⋮
199	Fire Station,…
200	More Art,…



Announcements

Program 2 due tomorrow @ 11:59 PM

Lab 5 due Wednesday @ 11:59 PM

Quiz 2 due on Thursday @ 11:59

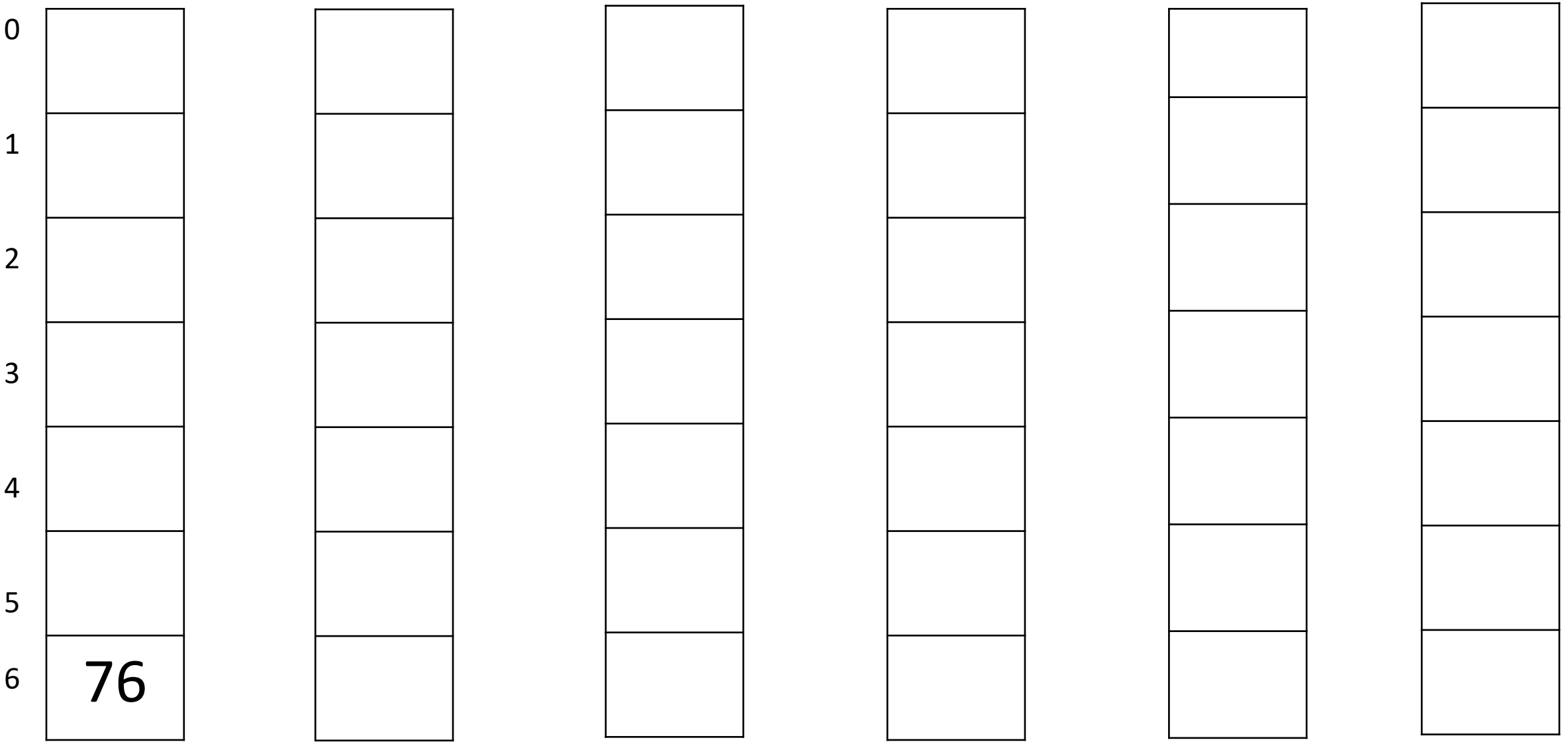
- Will talk about it more on Wednesday
- No Lecture on Thursday

Next weeks lectures will be at an alternative location (location coming soon)

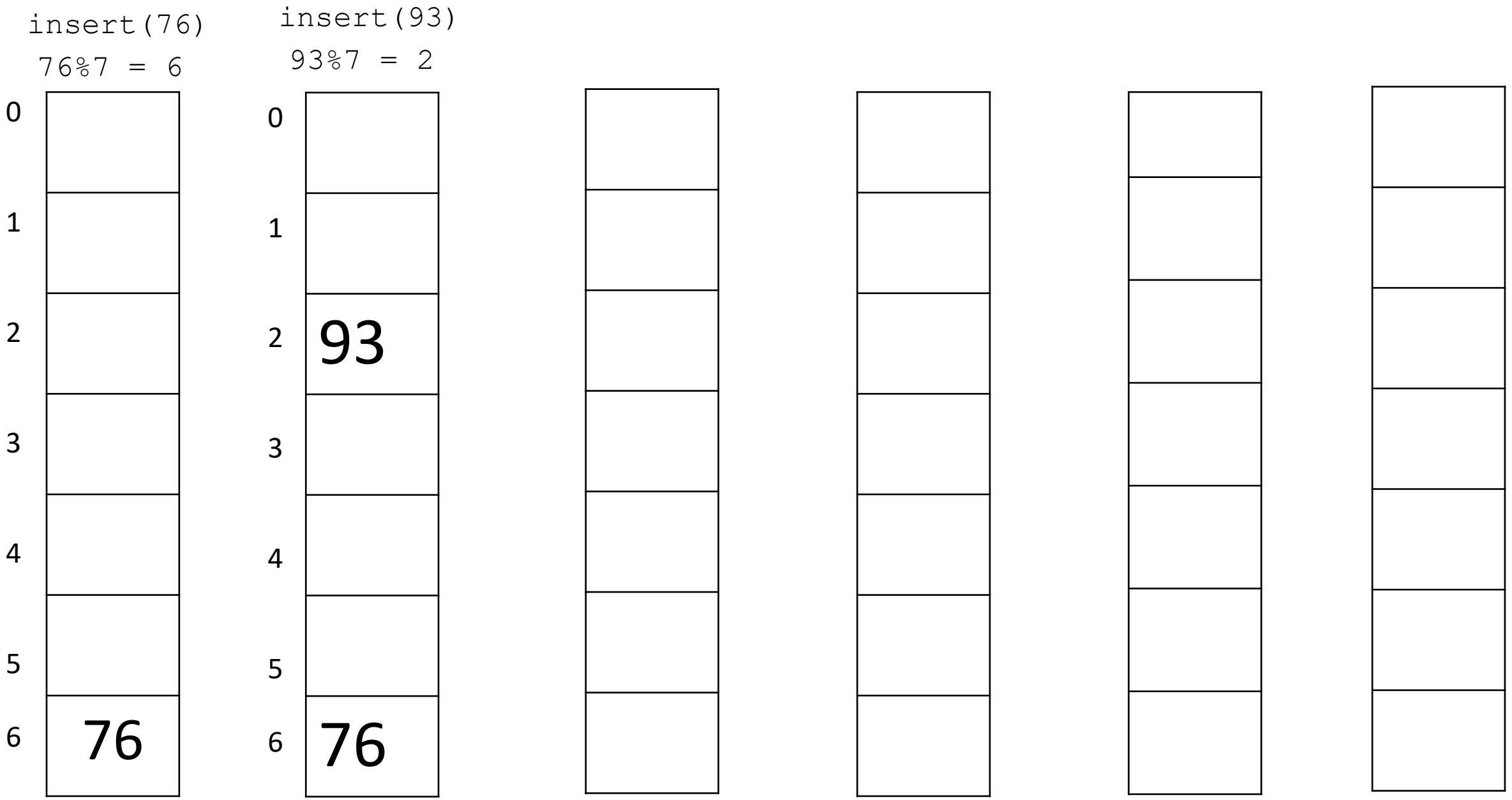
(Online people) If you have not talked to me yet this Summer, you should schedule a time to check-in with me this week or next week ☺

Linear Probing

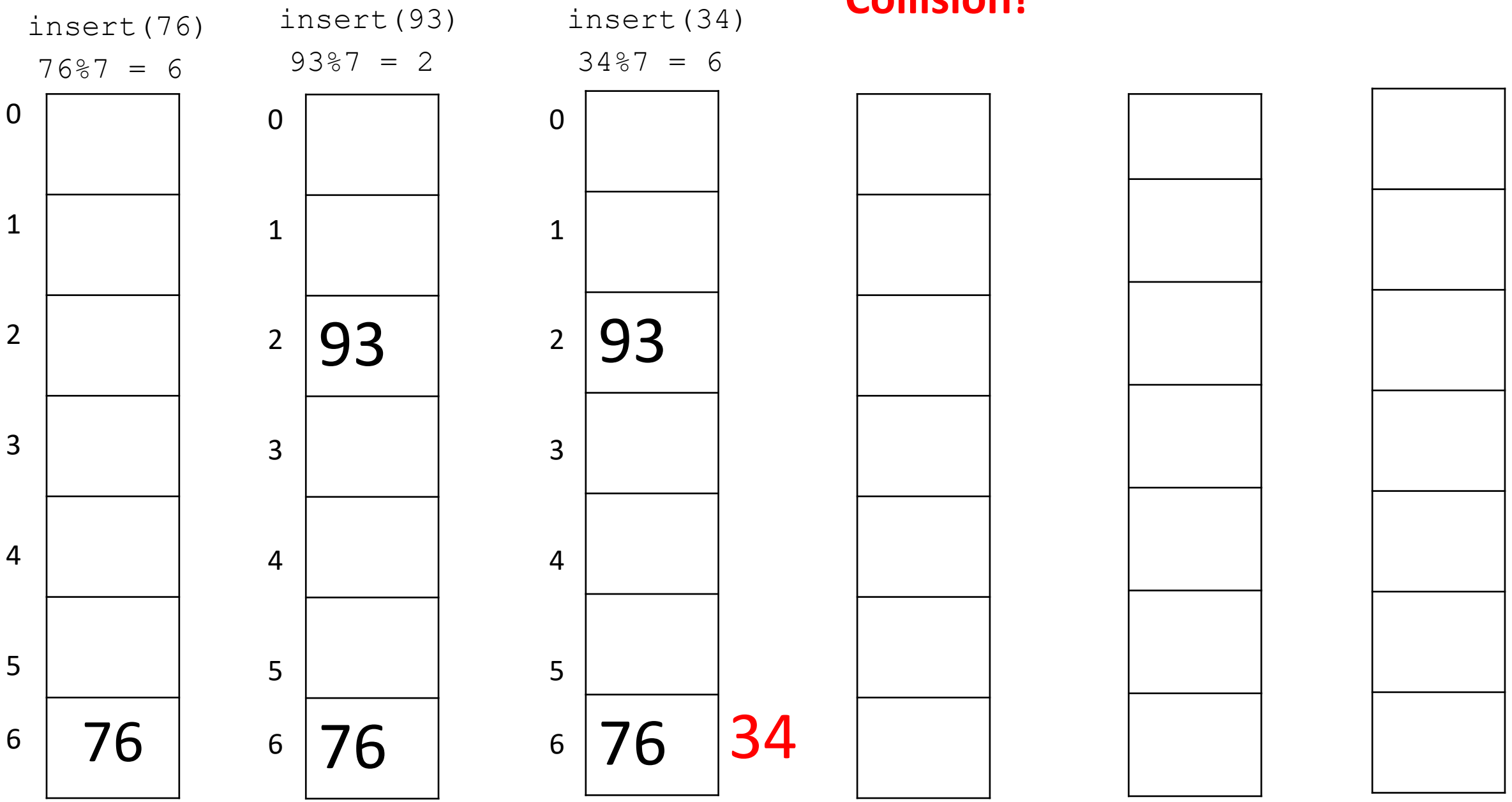
```
insert(76)
 76%7 = 6
```



Linear Probing

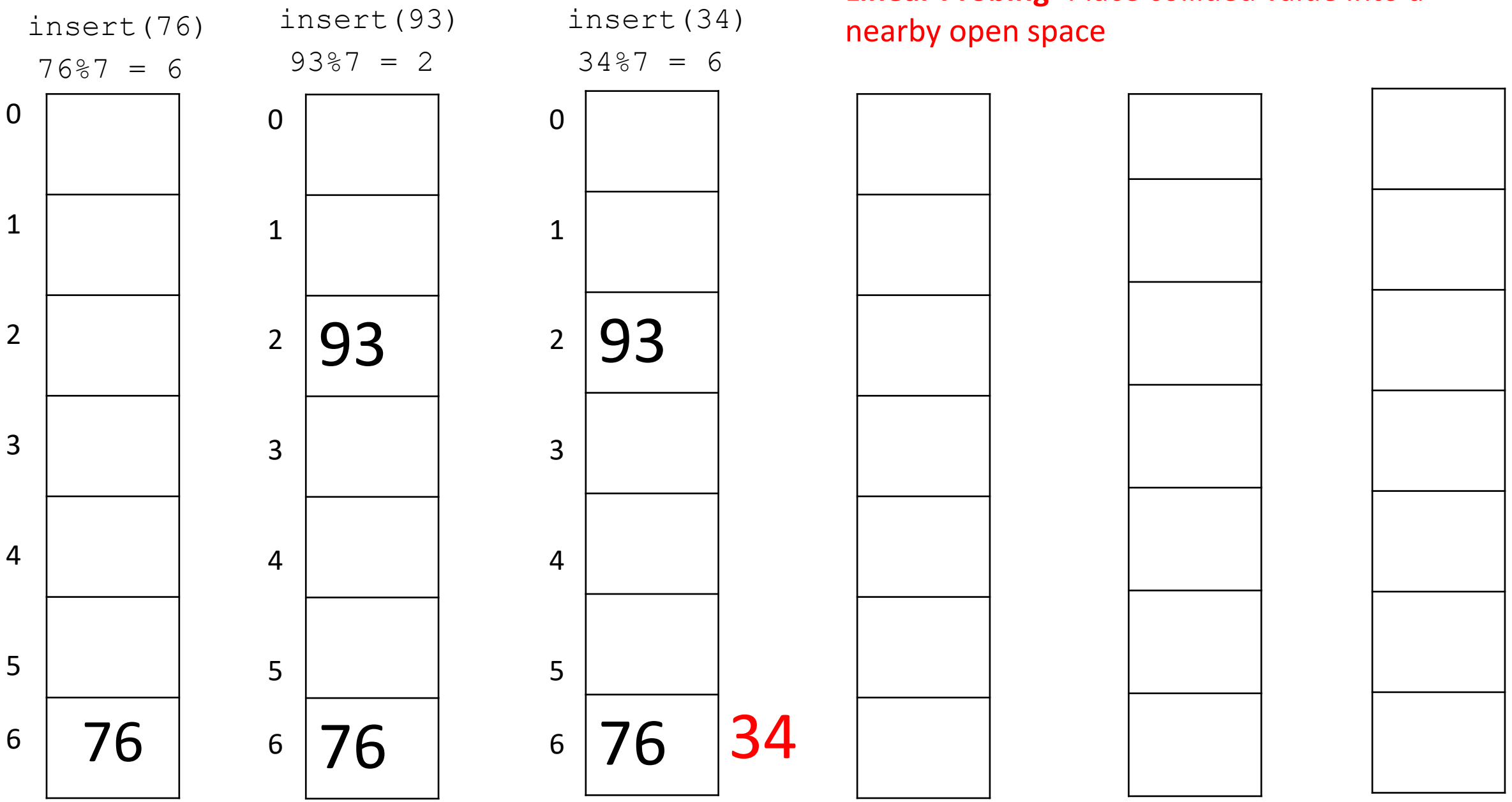


Linear Probing



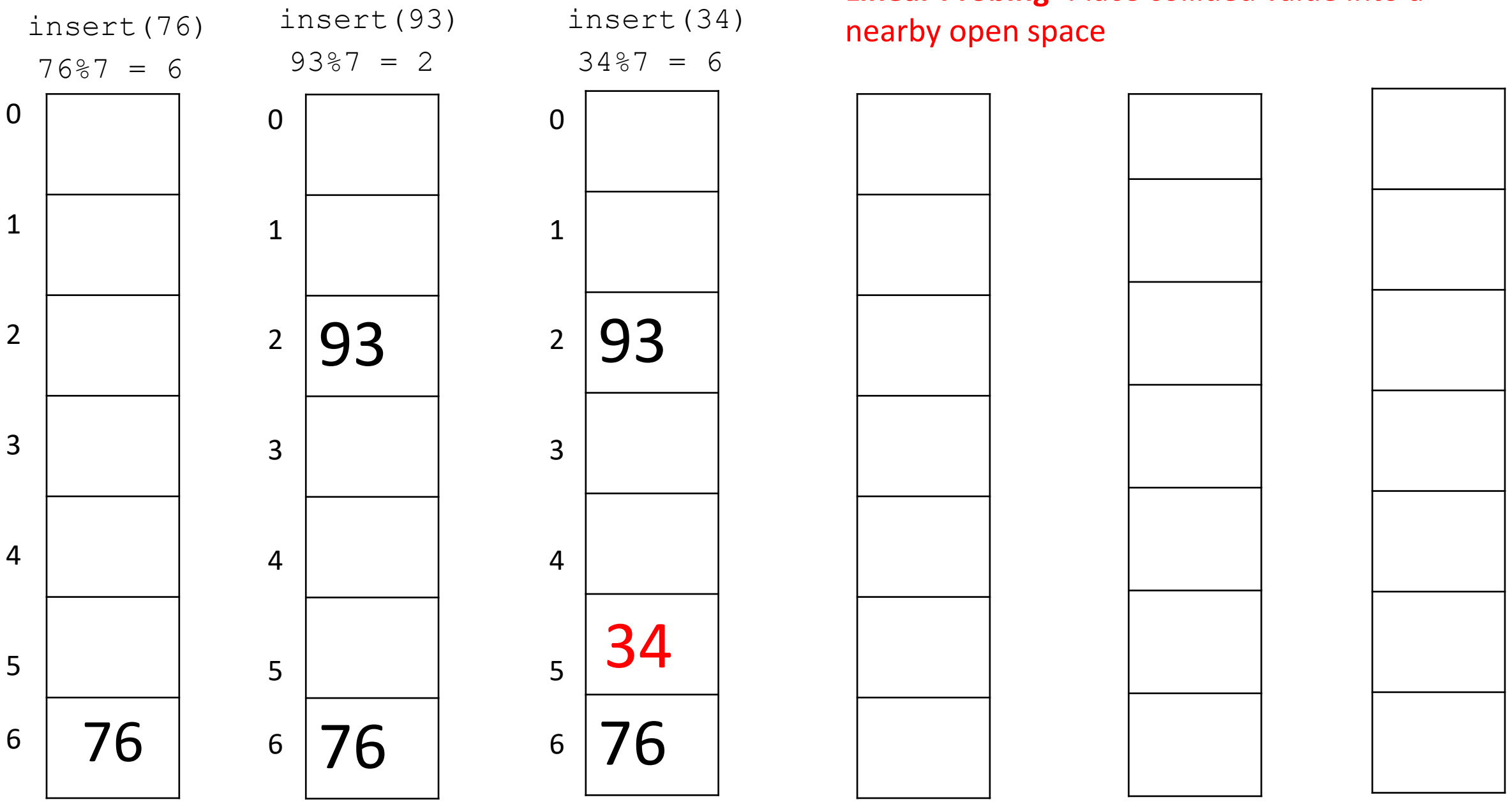
Linear Probing

Linear Probing- Place collided value into a nearby open space

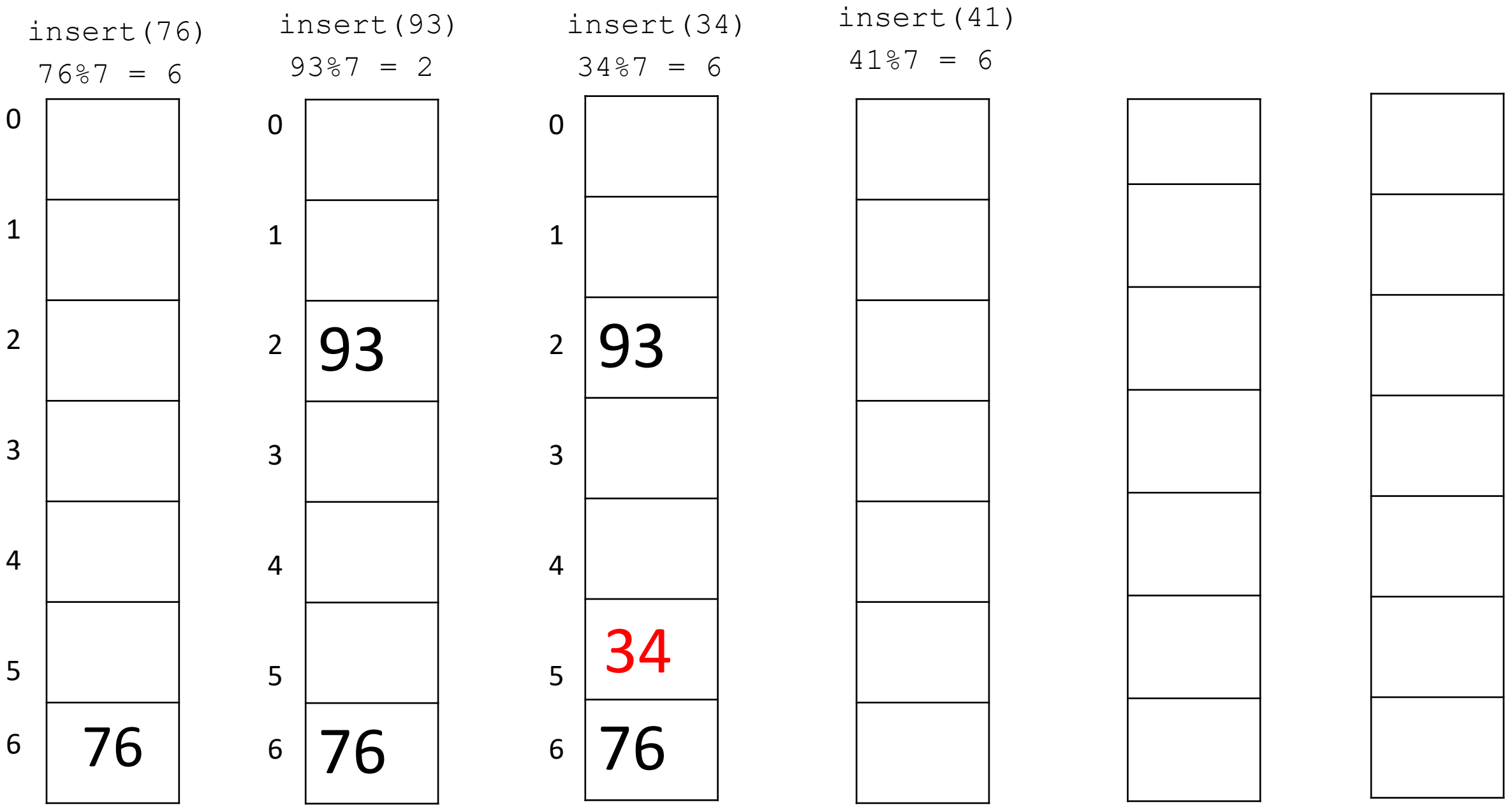


Linear Probing

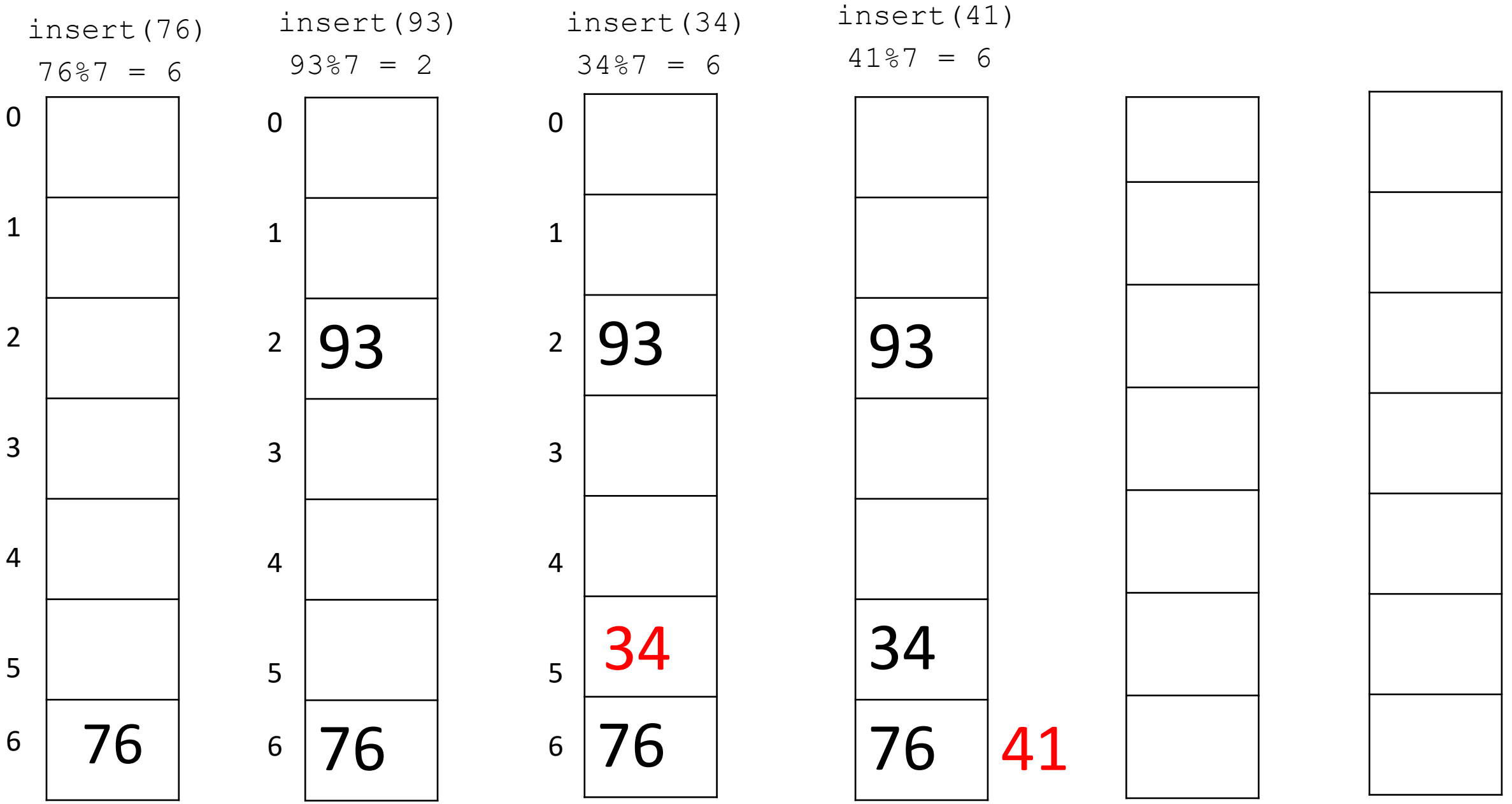
Linear Probing- Place collided value into a nearby open space



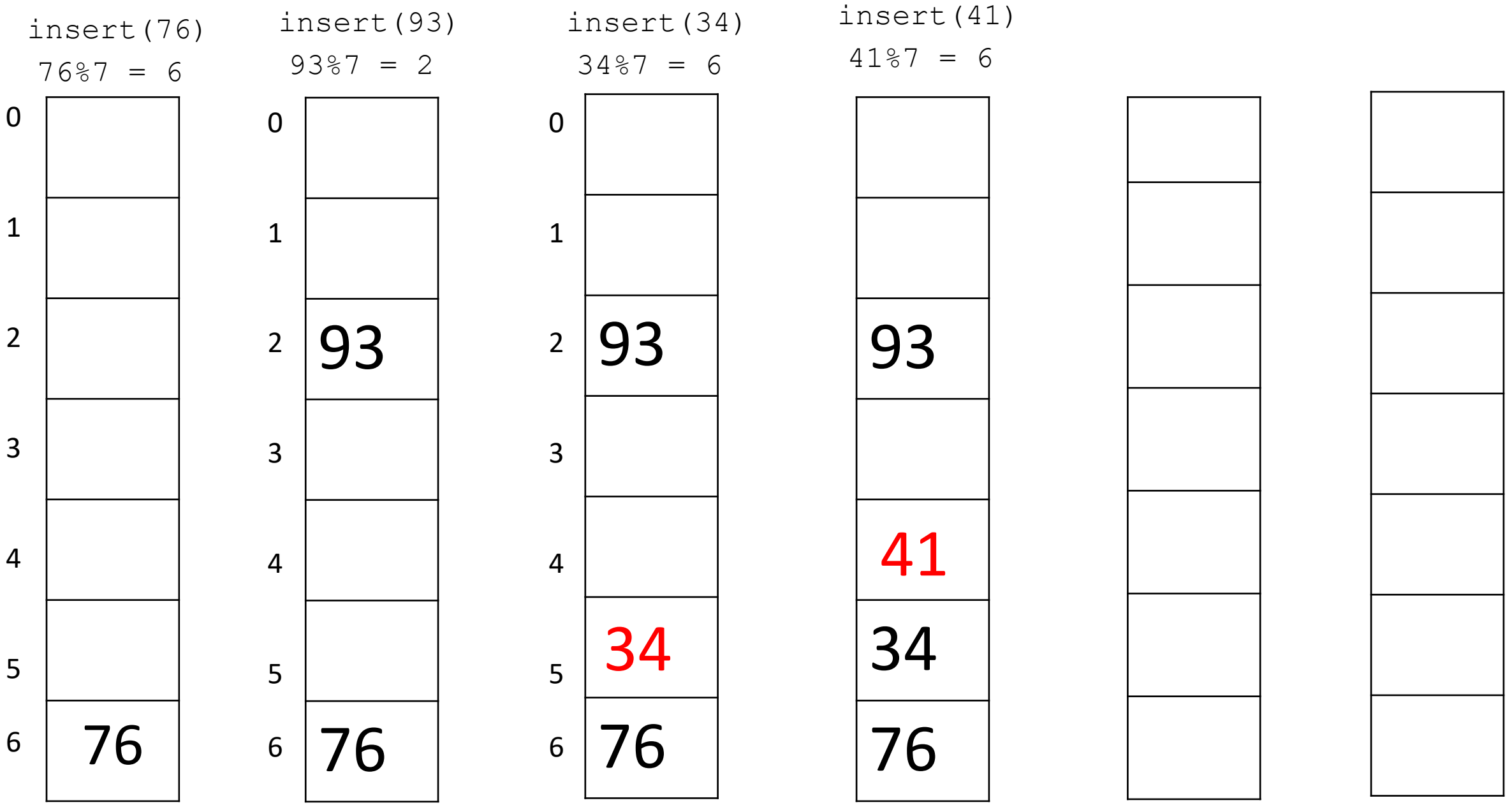
Linear Probing



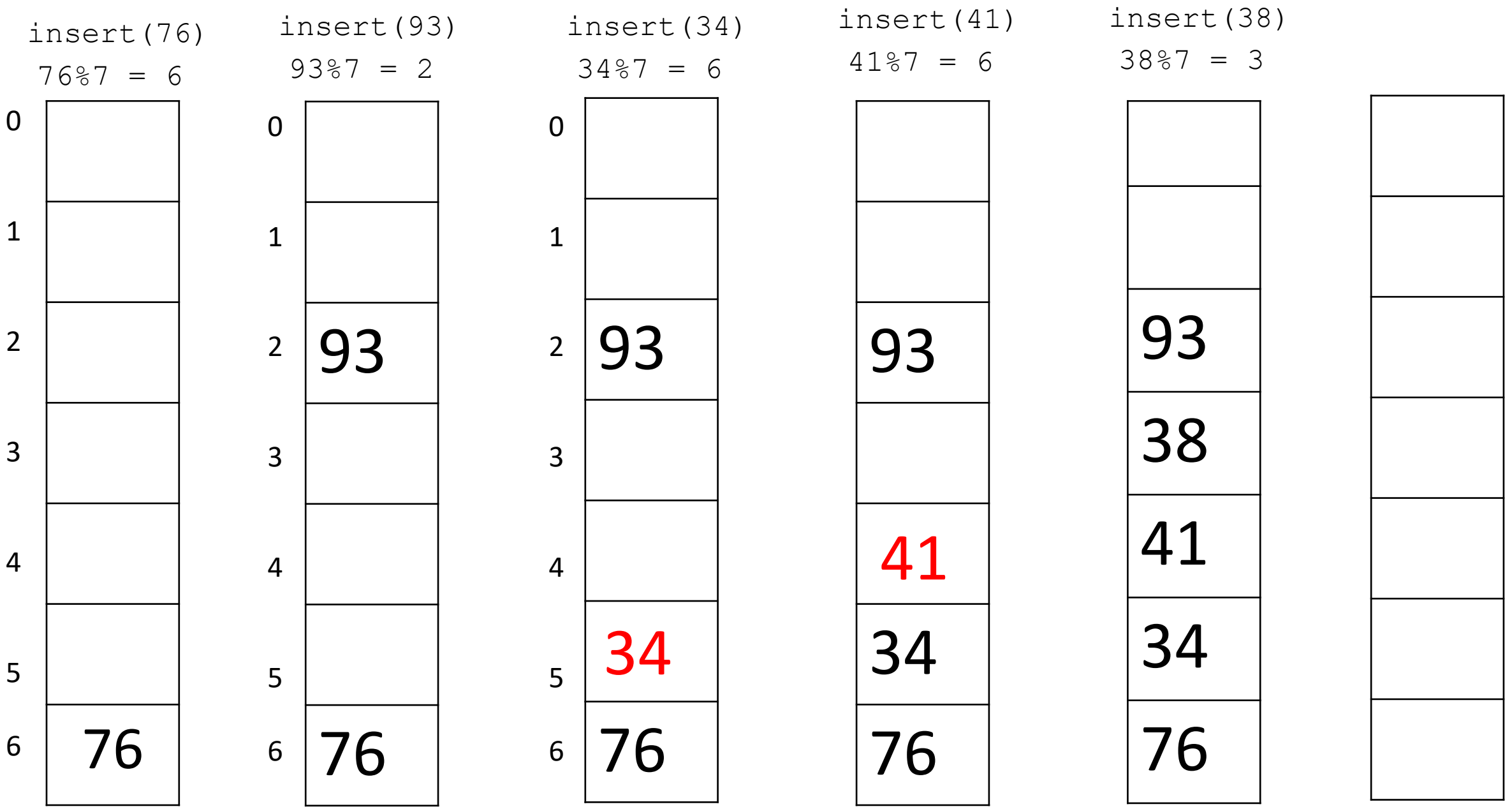
Linear Probing



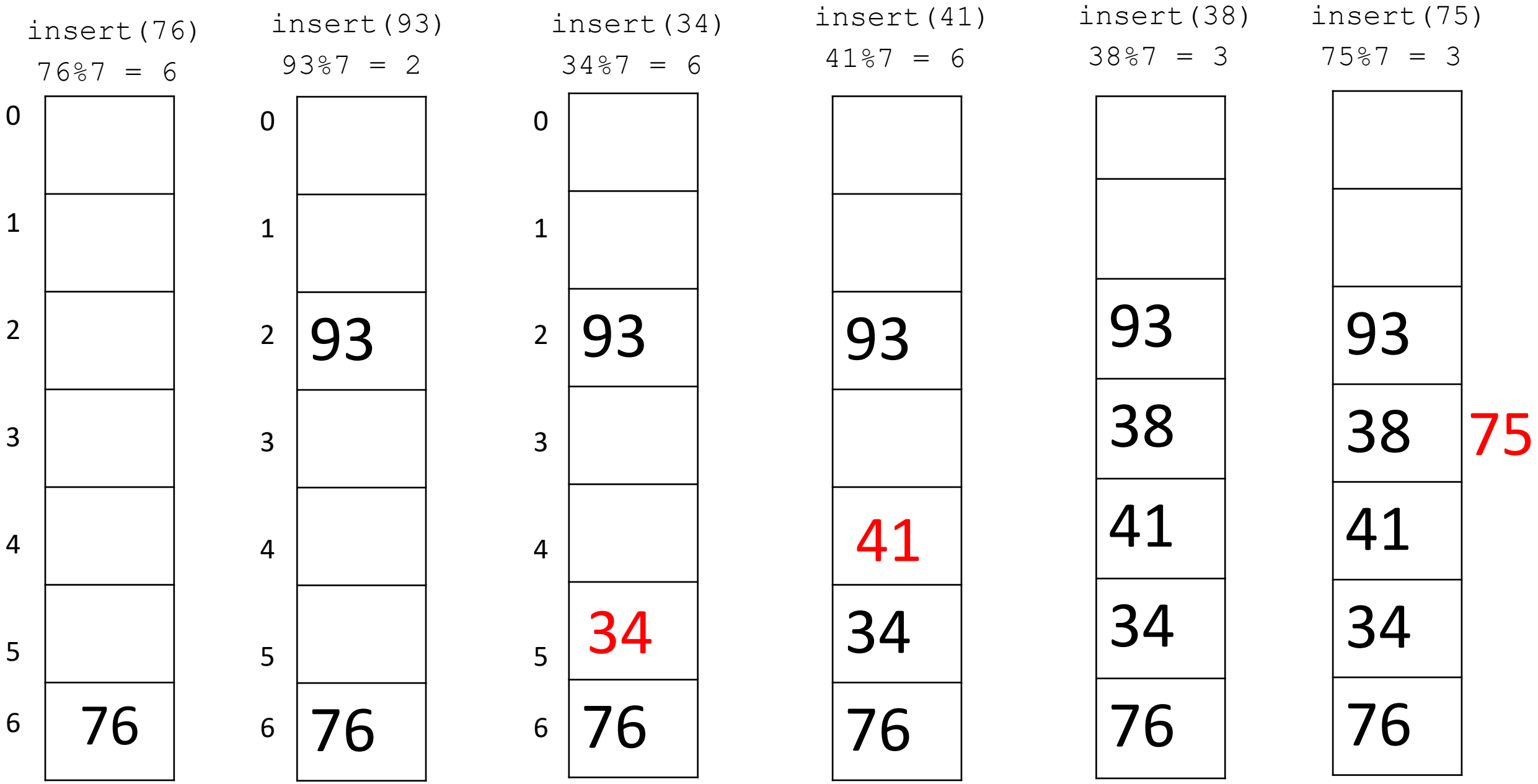
Linear Probing



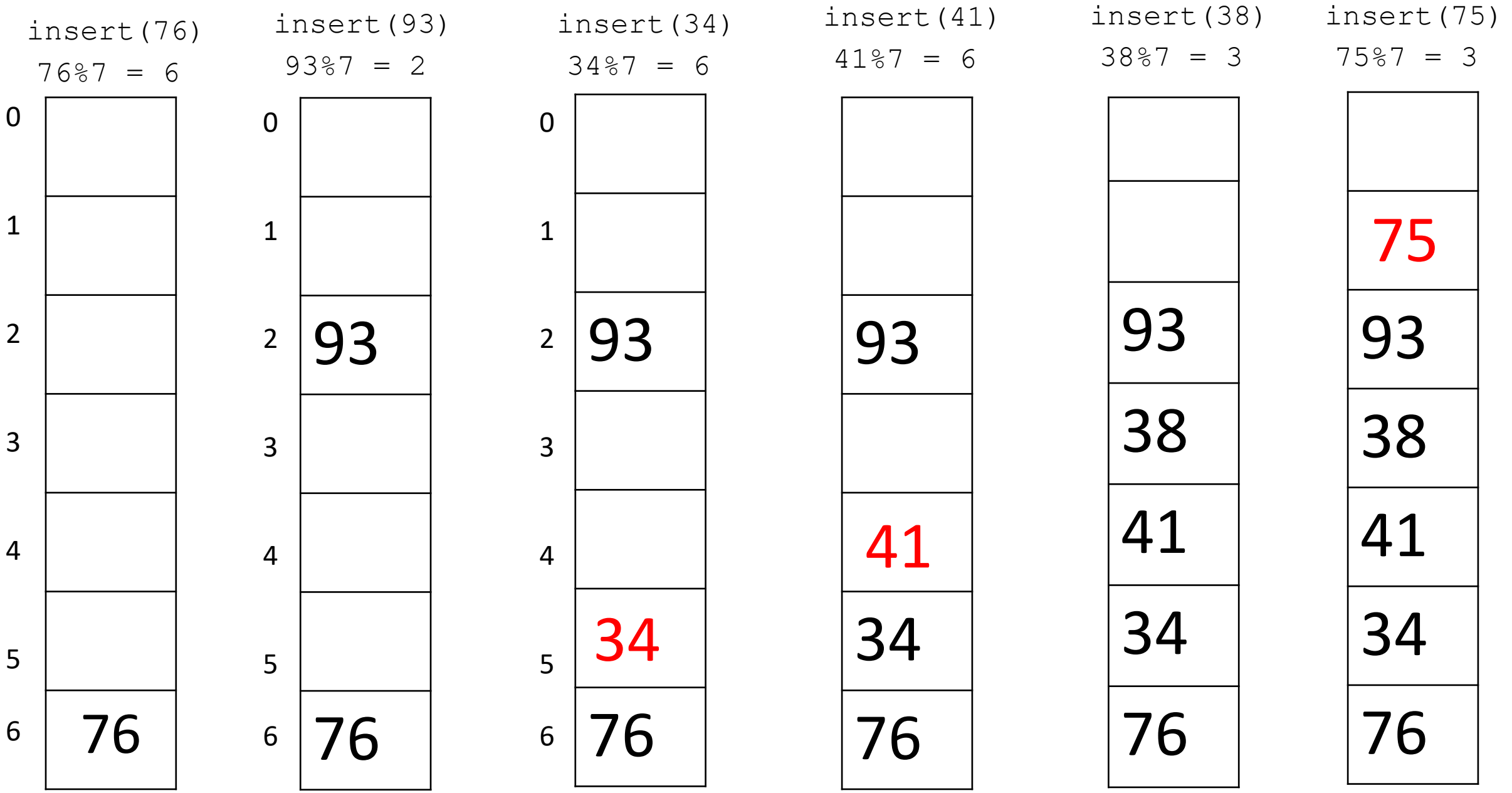
Linear Probing



Linear Probing



Linear Probing



Separate Chaining

Student ID

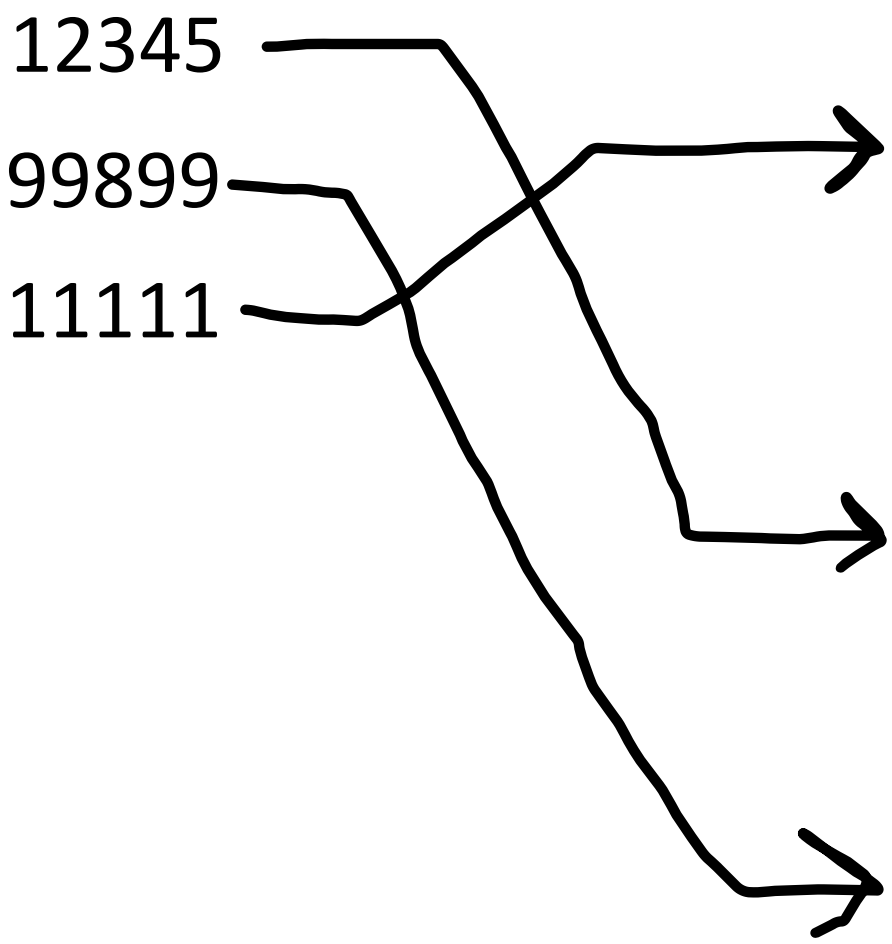
Hash Function

`ID % 100`

12345

99899

11111



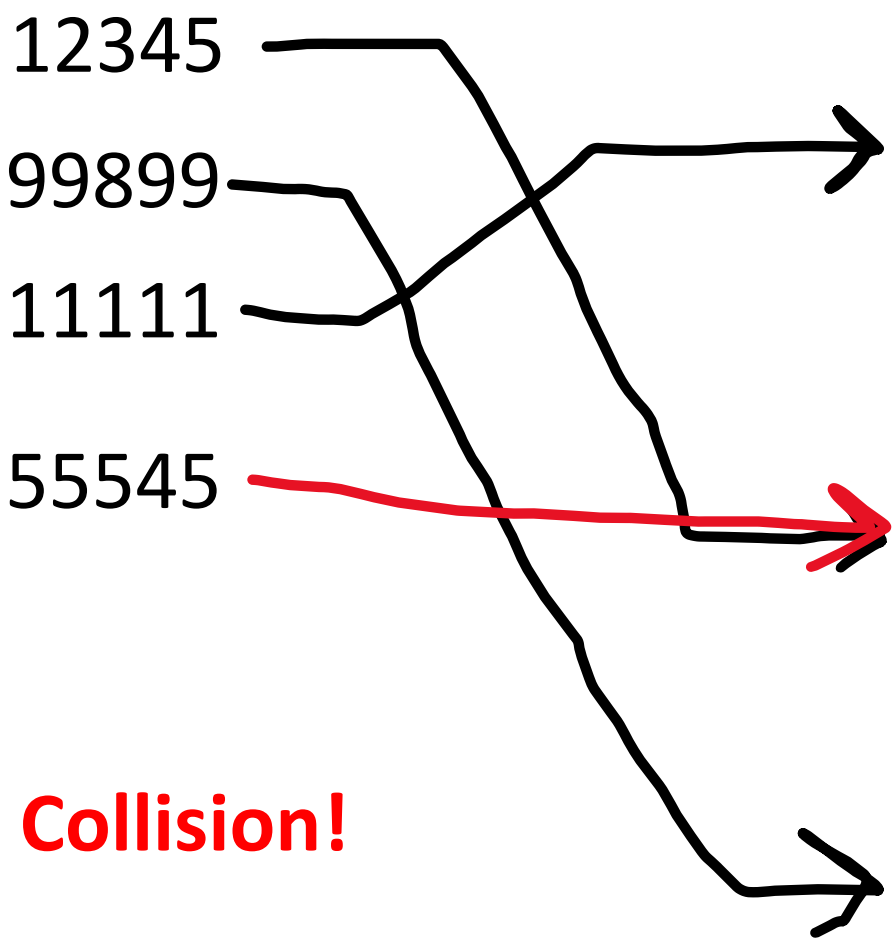
Index	Database
0	
1	
…	
…	
11	Student
…	
…	
…	
45	Student
…	
…	
…	
99	Student

Separate Chaining

Student ID

Hash Function

`ID % 100`



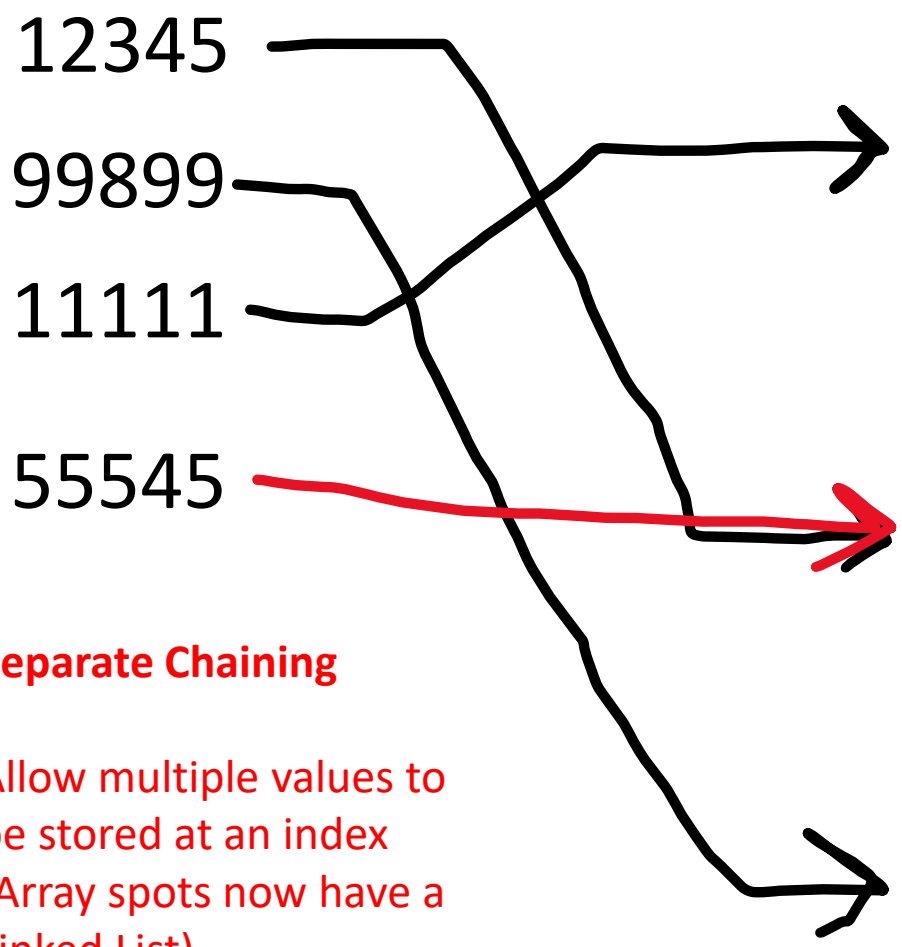
Index	Database
0	
1	
…	
…	
11	Student
…	
…	
…	
45	Student
…	
…	
…	
99	Student

Separate Chaining

Hash Function

Student ID

`ID % 100`



Index	Database
0	
1	
…	
…	
11	Student
…	
…	
…	
45	Student
…	
…	
…	
99	Student

Separate Chaining

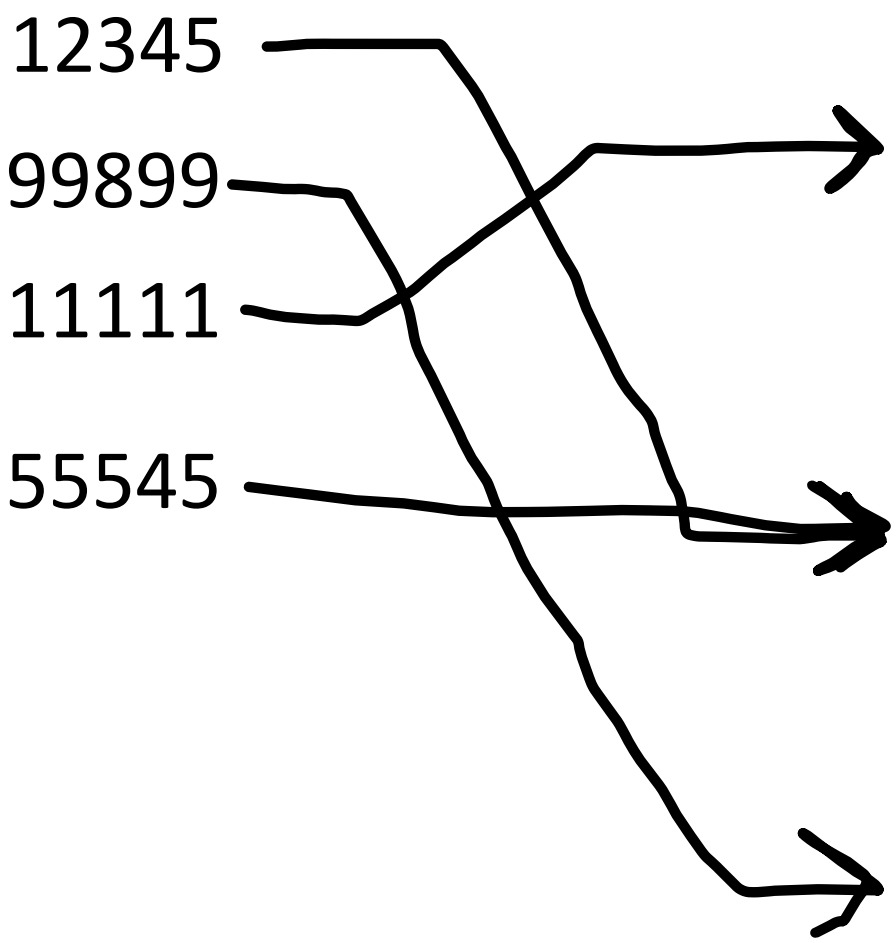
Allow multiple values to be stored at an index (Array spots now have a Linked List)

Separate Chaining

Hash Function

Student ID

`ID % 100`



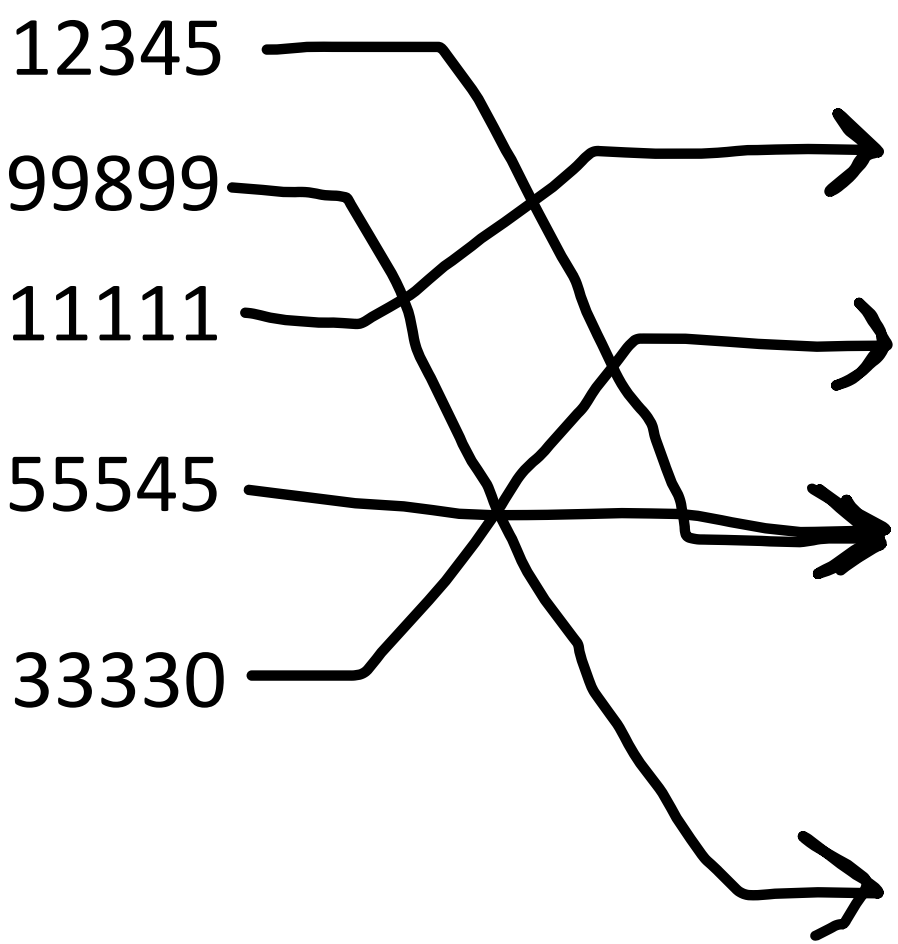
Index	Database
0	<>
1	<>
…	<>
…	<>
11	< Student >
…	<>
…	<>
…	<>
45	< Student → Student >
…	<>
…	<>
…	<>
99	< Student >

Separate Chaining

Hash Function

Student ID

`ID % 100`



Index | Database | < > = Linkedlist

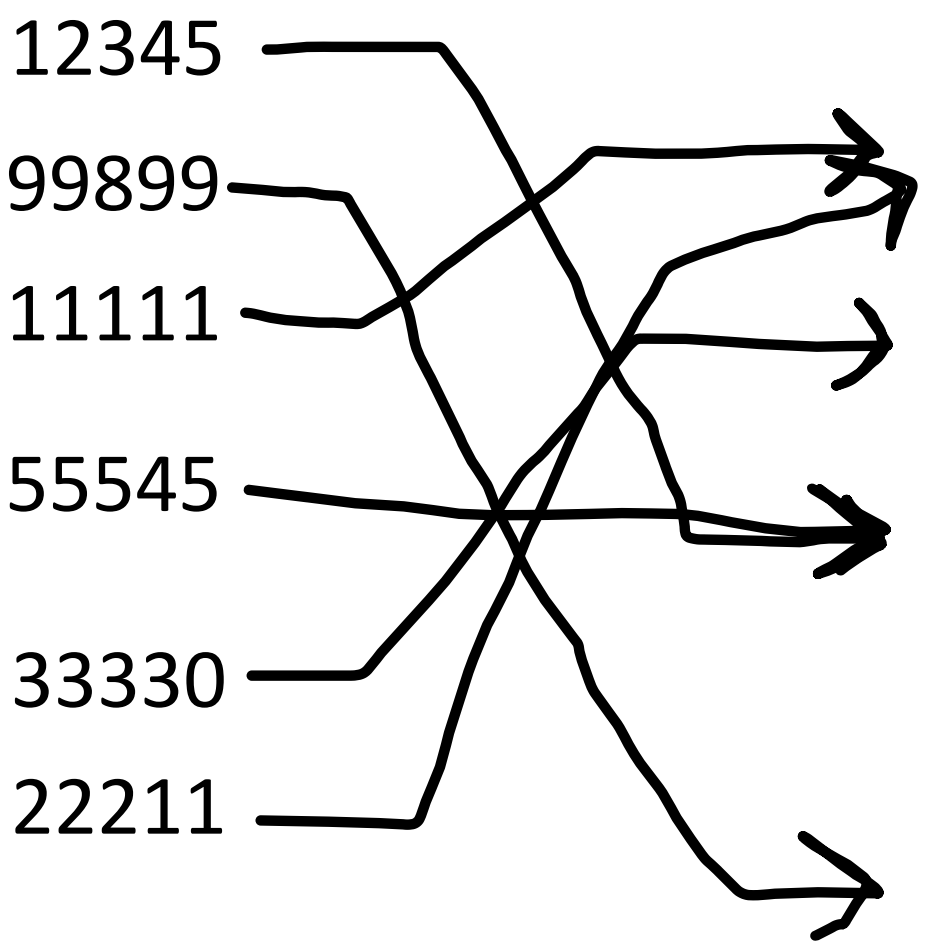
Index	Database
0	<>
1	<>
…	<>
…	<>
11	< Student >
…	<>
30	< Student >
…	<>
45	< Student → Student >
…	<>
…	<>
…	<>
99	< Student >

Separate Chaining

Hash Function

Student ID

`ID % 100`



< > = Linkedlist

Index	Database
0	<>
1	<>
…	<>
…	<>
11	< Student → Student >
…	<>
30	< Student >
…	<>
45	< Student → Student >
…	<>
…	<>
…	<>
99	< Student >

Separate Chaining

Hash Function

Student ID

`ID % 100`

12345

99899

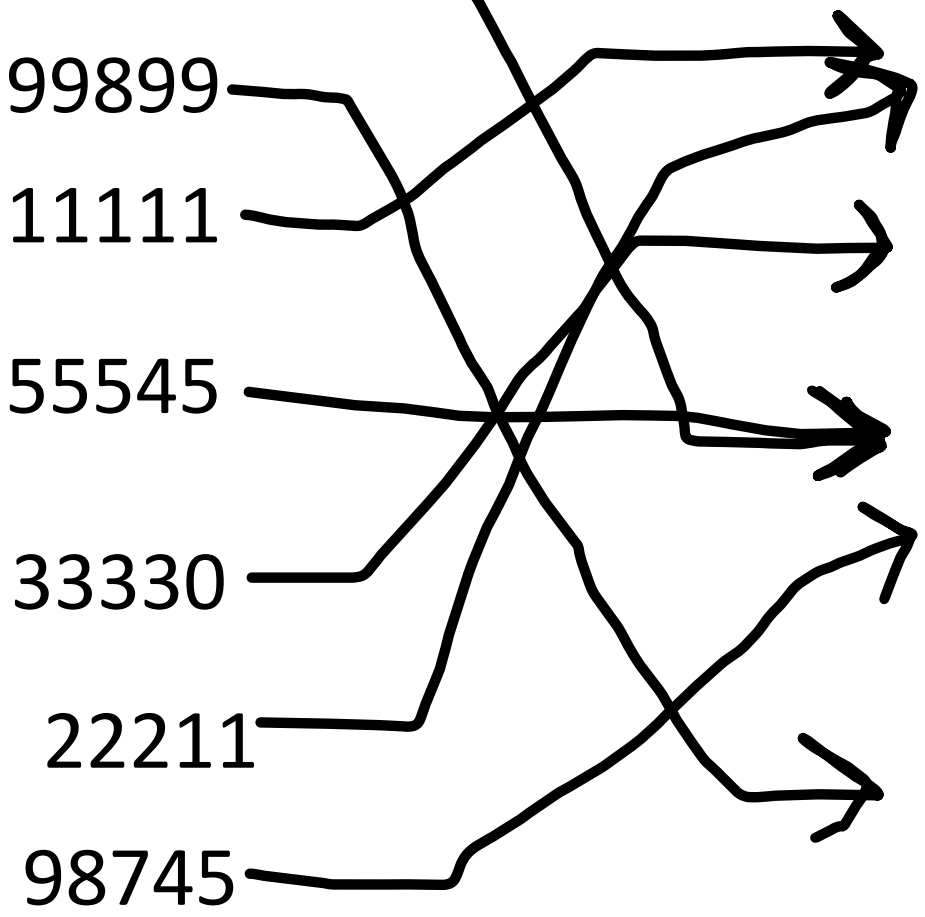
11111

55545

33330

22211

98745



Index | Database | < > = Linkedlist

Index	Database
0	<>
1	<>
…	<>
…	<>
11	< Student → Student >
…	<>
30	< Student >
…	<>
45	< Student → Student → Student >
…	<>
…	<>
…	<>
99	< Student >

Separate Chaining

Hash Function

Student ID

`ID % 100`

12345

55545

get(55545)

98745

Index | Database | < > = Linkedlist

Index	Database
0	<>
1	<>
…	<>
…	<>
11	< Student → Student >
…	<>
30	< Student >
…	<>
45	< Student → Student → Student >
…	<>
…	<>
…	<>
99	< Student >

Separate Chaining

Hash Function

Student ID

`ID % 100`

12345

55545

get(55545)

98745

Index | Database | < > = Linkedlist

Index	Database
0	<>
1	<>
…	<>
…	<>
11	< Student → Student >
…	<>
30	< Student >
…	<>
45	< Student → Student → Student >
…	<>
…	<>
…	<>
99	< Student >

Separate Chaining

Hash Function

Student ID

`ID % 100`

12345

55545

get(55545)

98745

Index | Database

< > = Linkedlist

Index	Database
0	<>
1	<>
…	<>
…	<>
11	< Student → Student >
…	<>
30	< Student >
…	<>
45	< Student → Student → Student >
…	<>
…	<>
…	<>
99	< Student >

Separate Chaining

Hash Function

Student ID

`ID % 100`

12345

55545

get(55545)

98745

Index | Database | < > = Linkedlist

Index	Database
0	<>
1	<>
…	<>
…	<>
11	< Student → Student >
…	<>
30	< Student >
…	<>
45	< Student → Student → Student >
…	<>
…	<>
…	<>
99	< Student >

Separate Chaining

Hash Function

Student ID

`ID % 100`

12345

55545

get(55545)

O(n)

☹

98745

Index | Database | < > = Linkedlist

Index	Database
0	<>
1	<>
…	<>
…	<>
11	<>
…	<>
30	<>
…	<>
45	< Student → Student → Student >
…	<>
…	<>
…	<>
99	<>