CS 225
Data Structures

Oct. 30 – Hashing 2
A Hash Table based Dictionary

Client Code:

```
1. Dictionary<KeyType, ValueType> d;
2. d[k] = v;
```

A **Hash Table** consists of three things:

1. 

2. 

3. 
Exam Updates

Current: Exam 8 (Programming)
• Topics: MP4-like (eg: iterators), AVL

Next Week: Exam 9 (Theory)
• Topics: AVL trees
  BTrees
  kD Trees
Collision Handling: Separate Chaining

\[ S = \{ 16, 8, 4, 13, 29, 11, 22 \} \]

\[ h(k) = k \mod 7 \]

\[ |S| = n \]

\[ |Array| = N \]

(Example of open hashing)
Collision Handling: Probe-based Hashing

$S = \{ 16, 8, 4, 13, 29, 11, 22 \}$  \hspace{1cm} |S| = n

$h(k) = k \% 7$  \hspace{1cm} |Array| = N

Try $h(k) = (k + 0) \% 7$, if full...
Try $h(k) = (k + 1) \% 7$, if full...
Try $h(k) = (k + 2) \% 7$, if full...
Try ...
A Problem w/ Linear Probing

Primary clustering:

Description:

Remedy:
Collision Handling: Double hashing

\[ S = \{ 16, 8, 4, 13, 29, 11, 22 \} \quad |S| = n \]

\[ h(k) = k \mod 7 \quad |\text{Array}| = N \]

Try \[ h(k) = (k + 0 \times h_2(k)) \mod 7 \], if full...

Try \[ h(k) = (k + 1 \times h_2(k)) \mod 7 \], if full...

Try \[ h(k) = (k + 2 \times h_2(k)) \mod 7 \], if full...

Try ...

\[ h(k, i) = (h_1(k) + i \times h_2(k)) \mod 7 \]
Running Times

The expected number of probes for find(key) under SUHA

**Linear Probing:**
- Successful: $\frac{1}{2}(1 + 1/(1-\alpha))$
- Unsuccessful: $\frac{1}{2}(1 + 1/(1-\alpha))^2$

**Double Hashing:**
- Successful: $1/\alpha \cdot \ln(1/(1-\alpha))$
- Unsuccessful: $1/(1-\alpha)$

**Separate Chaining:**
- Successful: $1 + \alpha/2$
- Unsuccessful: $1 + \alpha$

*Instead, observe:*

- As $\alpha$ increases:
- If $\alpha$ is constant:

*(Don’t memorize these equations, no need.)*
Running Times

The expected number of probes for find(key) under SUHA

Linear Probing:
• Successful: $\frac{1}{2}(1 + 1/(1-\alpha))$
• Unsuccessful: $\frac{1}{2}(1 + 1/(1-\alpha))^2$

Double Hashing:
• Successful: $1/\alpha \cdot \ln(1/(1-\alpha))$
• Unsuccessful: $1/(1-\alpha)$
ReHasing

What if the array fills?
Which collision resolution strategy is better?

- Big Records:

- Structure Speed:

What structure do hash tables replace?

What constraint exists on hashing that doesn’t exist with BSTs?

Why talk about BSTs at all?
Exam 8 (programming exam, MP4-like and AVL) is live!
More Info: https://courses.engr.illinois.edu/cs225/fa2017/exams/

MP5: EC due tonight!
Extra Credit +7 deadline: Monday, Oct. 30

Lab: lab_hash is released Wednesday
Due Sunday, Dec. 5 at 11:59pm

POTD
Every Monday-Friday – Worth +1 Extra Credit /problem (up to +40 total)