Today's announcements:

MP6 available, due 04/14, 11:59p. EC due 04/07, 11:59p

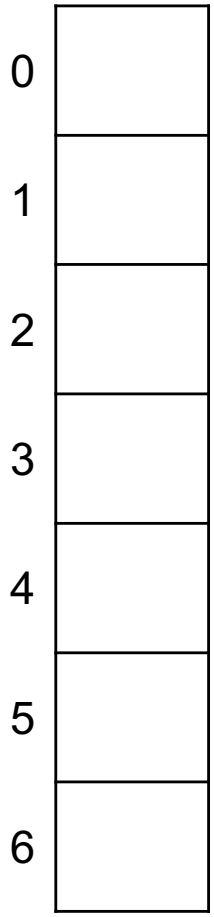
Hash Table consists of the following:

1.
2.
3.

Two different kinds of item 3:

Collision handling - Separate Chaining: (an example of open hashing)

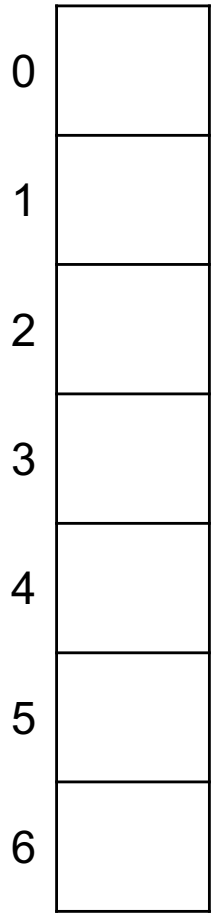
S = {16, 8, 4, 13, 29, 11, 22} |S| = n h(k) = k%7



	Worst case	**Under SUHA**
Insert		
Remove/find		

Collision Handling - Probe based hashing: (example of closed hashing)

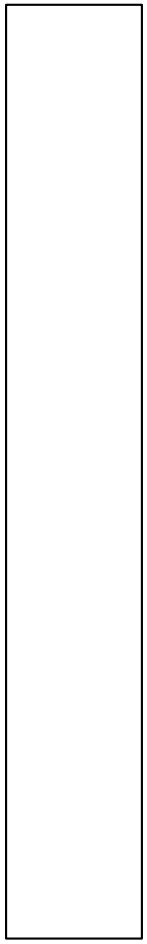
S = {16, 8, 4, 13, 29, 11, 22} |S| = n h(k) = k%7



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...

Probe based hashing – a problem…

Primary clustering:



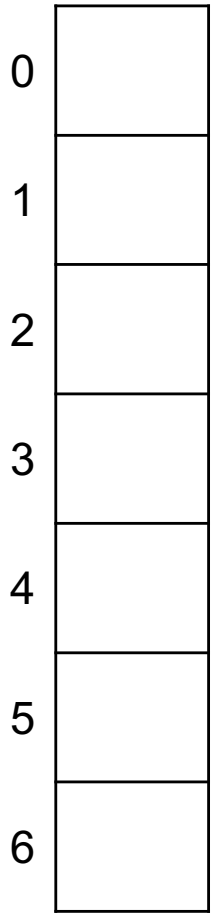
Description:

Remedy:

http://groups.engin.umd.umich.edu/CIS/course.des/cis350/hashing/WEB/HashApplet.htm

Probe based hashing: (double hashing)

S = {16, 8, 4, 13, 29, 11, 22} |S| = n $H(k,i) = h_1(k) + ih_2(k)$



Try $h(k) = (k + 0*h_2(k)) \% 7$. If full...
try $h(k) = (k + 1*h_2(k)) \% 7$. If full...
try $h(k) = (k + 2*h_2(k)) \% 7$. If full...
try...

Hash table performance: expected # 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: $1/\alpha \ln 1/(1-\alpha)$

unsuccessful: $1/(1-\alpha)$

Separate chaining -

successful: $1 + \alpha/2$

unsuccessful: $1 + \alpha$

Do not memorize these!

Observe:

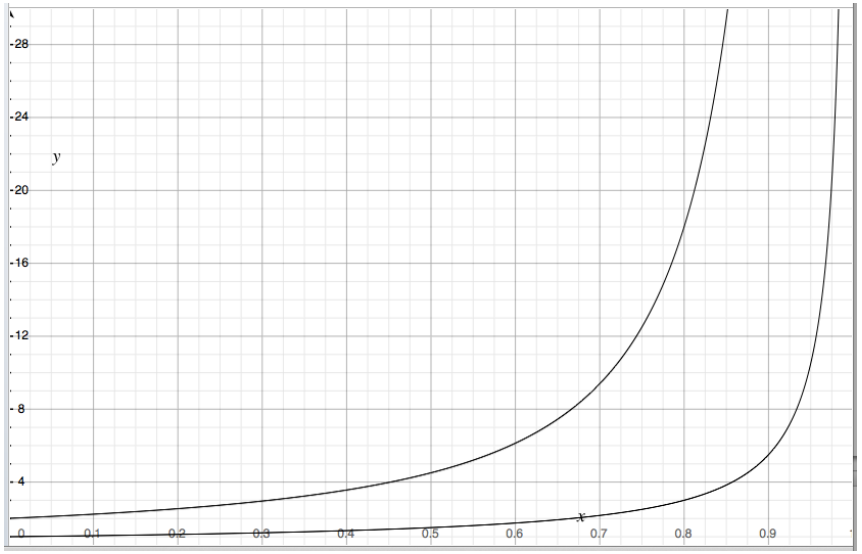
- As α increases, running times increase…
- If α is held constant then running times are constant…

Hash table performance: expected # of probes for Find(key) under SUHA

Linear probing -

successful: $\frac{1}{2}(1 + 1/(1-\alpha))$

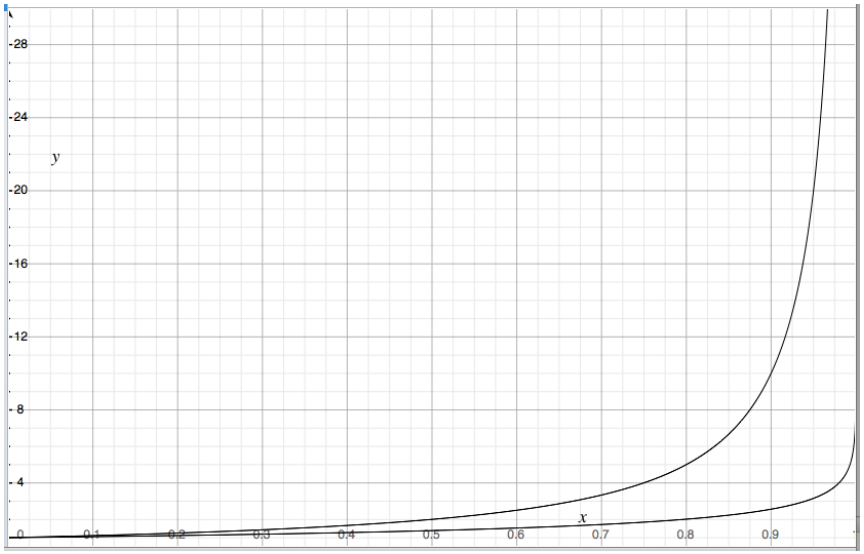
unsuccessful: $\frac{1}{2}(1 + 1/(1-\alpha))^2$



Double hashing -

successful: $1/\alpha \ln 1/(1-\alpha)$

unsuccessful: $1/(1-\alpha)$



ReHashing:

What if the array fills?

Hashing Miscellaneous Discussion –

Which collision resolution strategy is better?

•Big records –

•Structure speed –

What structures do hash tables replace for us?

There is a constraint on Keyspaces for BST that does not affect hashing...

Why do we talk about balanced BST if hashing is so great?

Applications of hashing?

Area of active research in mathematics to develop general purpose hash functions.