Trading Space for Time

Chapter Intro

➪ ?Preprocessing

➪ Preconditioning

Differences, one introduces extra inputs to procedure other extra info to original inputs.

➪ Input enhancement

How different from presorting?

Quiz
Can dynamic programming and data compression, both of which rely on extra memory, be viewed as trades for times? (An answer last slide).

➪ Prestructuring: faster access

- ✎ Hash table
- ✎ B-tree

Quiz
What style of trading space for time is used by each application?

➪ Important applications

String matching, hash tables, B-trees

Hashing Basics

Recall, 2 components of ADT: a data arrangement (scheme), and characteristic (defining) ops.

- ⇨ **Dictionary [ADT]**
- ⇨ **Collision**

$K_1, K_2, \cdots, K_n$

A **dictionary** of n records (objects) with keys (property fields), best lookup times.

Reverse access array, i.e., find index given content (a key).

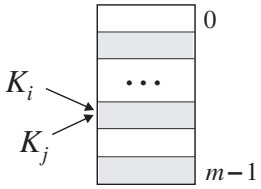
⇨ Motivation

⇨ Hash tables: $H[0..m-1]$

Hash functions have to satisfy confilicting requirements.

⇨ Hash functions (typical)

$$h(K) = K \mod m$$



⇨ Key collisions

Hash Tables
Open Hashing

- **Separate chaining**
- **Key chain**
- **Load factor, α**

A, FOOL, AND, HIS, MONEY, ARE, SOON, PARTED

1 18 5

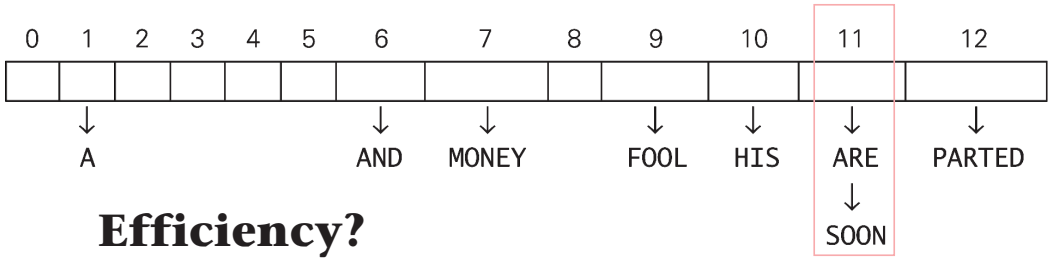
$$h(K) = \sum_{i=0}^{s-1} ord(c_i) \bmod m$$

13

Hash addresses range from 0 to 12.

$n = 8$

keys	A	FOOL	AND	HIS	MONEY	ARE	SOON	PARTED
hash addresses	1	9	6	10	7	11	11	12



Quiz
Define the **load factor** of a hashing scheme (function)? What would be optimal?

Quiz
What is the efficiency of each dictionary operation?

Efficiency?

Hash Tables
Closed Hashing

- Open addressing
- Linear probing

Quiz
Why would this scheme need m ≥ n?

Quiz
When would a key actually be deleted? (Check **lazy deletion** in textbook).

Quiz
Which inputs cause a worst case for closed hashing? What is the solution? Hint: 2.

Quiz
What is the efficiency of dictionary operations? Why is larger m>n preferable?

19 15 14

keys	A	FOOL	AND	HIS	MONEY	ARE	SOON	PARTED
hash addresses	1	9	6	10	7	11	11	12

0	1	2	3	4	5	6	7	8	9	10	11	12
	A											
	A								FOOL			
	A					AND			FOOL			
	A					AND			FOOL	HIS		
	A					AND	MONEY		FOOL	HIS		
	A					AND	MONEY		FOOL	HIS	ARE	
	A					AND	MONEY		FOOL	HIS	ARE	SOON
PARTED	A					AND	MONEY		FOOL	HIS	ARE	SOON

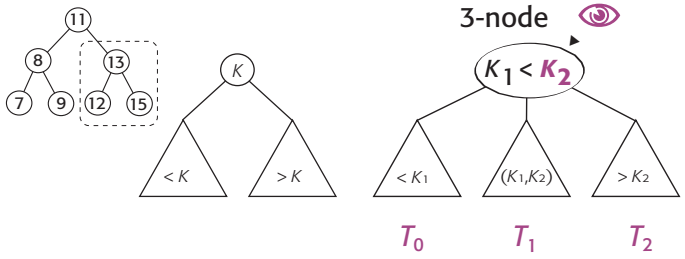
What if H[10] is needed? Determine position.

An Indexing Scheme

➪ n-Node

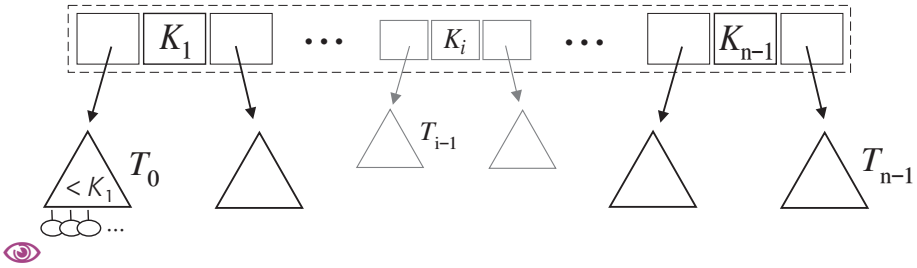
Quiz
What advantage was gained by adding 2nd key?

Recall



Pre-structuring involves "extra info to facilitate faster access." 🕮

Quiz
How many keys are needed to manage m children (example, 3)?



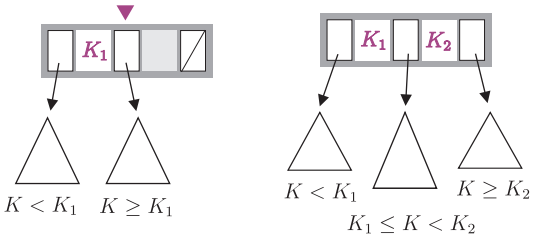
🕮 Structural requirements?

B-Trees

- **[B-tree] Order, m**
- **Parental node, p**

Quiz
What would a tree of order 2 look like?

m = 3 (2-3 tree), 2 keys

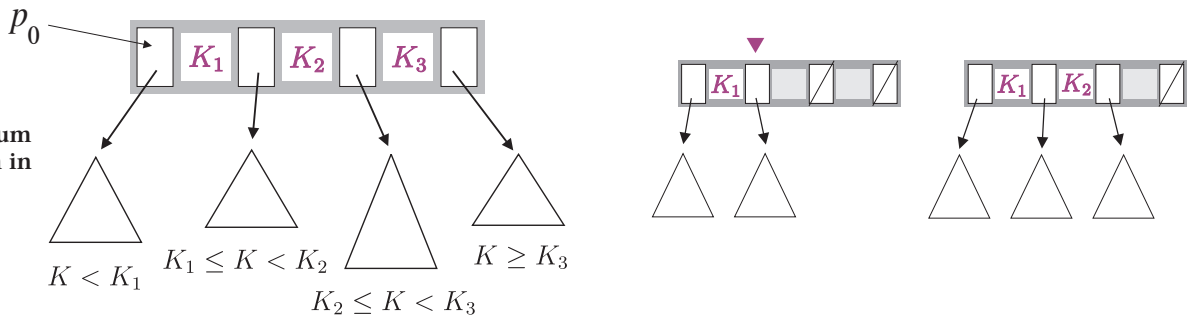


- How many children?
- How many keys?

Exercise
Refer to structural requirements to answer the question above.

m = 4 (2-3-4 tree), 3 keys

Quiz
What is the minimum number of children in a node?

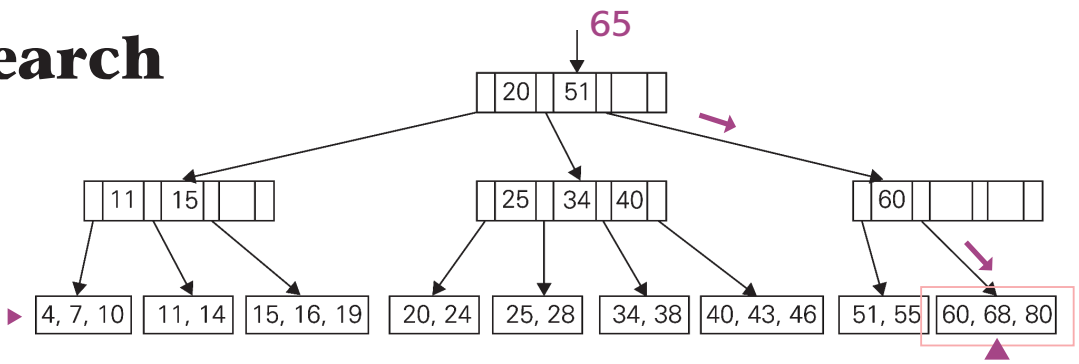


B-Tree Efficiency
Dictionary Ops

Search

Quiz
What is the order of the B-tree?

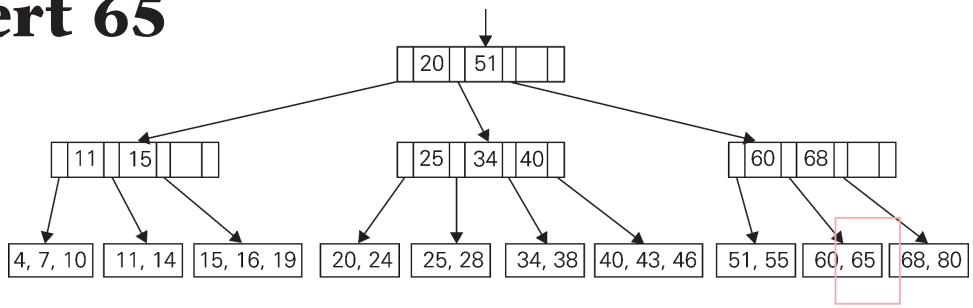
Note a main characteristic in this version of a B-tree.



Similar to 2-3 tree previously, split and promote to insert.

Insert 65

Key promotions expand indices as more records are added.



Trading Time-Space: Prestruct Conclusions

Is dyn prog a trade? No, space waste seems incidental, repeated recording of intermediate results can be avoided in computer memory. How about data compression? Perhaps, can't avoid extra memory even if goal is different, but need to verify that faster algorithms do rely on more extra memory to support notion of time-space trade.

Exercise
Compare hashing and balanced search trees (particularly B-trees) in terms efficiency of **dictionary** operations.

➪ Performance gains, compare

Identify major (signature) efficiency gain

Exercise
List at least 3 common applications of hashing. **Hint**: check textbook first then Google.

➪ Applications

Generating a random graph (hashing)

➪ Do the homework

Exercise
Read the linked article and write a comment in the course discussion group.

https://docs.oracle.com/cd/E11882_01/server.112/e40540/indexiot.htm#CNCPT1895

Assigned Homework Exercises
Third Edition

* Challenge question, don't submit with homework

- ~~Exercise 7.1 ♦ 1, 2, 3, 4, 5~~
- Exercise 7.2 ♦ 1, 2, 3, 4, 5
- Exercise 7.3 ♦ 1, 2
- Exercise 7.4 ♦ 4