

Topic 9

Using Maps

"He's off the map!"

-Stan (Mark Ruffalo) *Eternal Sunshine of the Spotless Mind*



Data Structures

- More than arrays and lists
- Write a program to determine the frequency of all the words in a file.

Performance using ArrayList

Title	Size (kb)	Total Words	Distinct Words	Time (sec)
small sample	0.6	89	25	0.001
2BR02B	34	5,638	1,975	0.051
Alice in Wonderland	120	29,460	6,017	0.741
Adventures of Sherlock Holmes	581	107,533	15,213	4.144
2008 CIA Factbook	10,030	1,330,100	74,042	173.000

Order?

- Express change in size as factor of previous file

Title	Size	Total Words	Distinct Words	Time
small sample	0.6	89	25	0.001
2BR02B	57x	63x	79x	51x
Alice in Wonderland	3.5x	5.2x	3.0x	14.5x
Adventures of Sherlock Holmes	4.8x	3.7x	2.5x	6.0x
2008 CIA Factbook	17x	12.3x	5x	42x

$O(\text{Total Words} * \text{Distinct Words})$??

Clicker 1

- ▶ Given 3 minutes for the 2008 CIA Factbook with 1,330,100 total words and 74,042 distinct words, how long for 1,000x total words and 100x distinct words?

- A. an hour
- B. a day
- C. a week
- D. a month
- E. half a year

Why So Slow??

- ▶ Write a contains method for an array based list

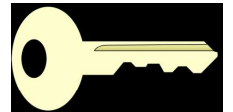
```
public boolean indexOf(Object o) {
```

A Faster Way - Maps

- ▶ Also known as:
 - table, search table, dictionary, associative array, or associative container
- ▶ A data structure optimized for a very specific kind of search / access
- ▶ In a *map* we access by asking "give me the *value* associated with this *key*."

Keys and Values

- ▶ Dictionary Analogy:
 - The *key* in a dictionary is a word:
foo
 - The *value* in a dictionary is the definition:
First on the standard list of metasyntactic variables used in syntax examples
- ▶ A key and its associated value form a pair that is stored in a map
- ▶ To retrieve a value the key for that value must be supplied
 - A List can be viewed as a Map with integer keys



More on Keys and Values

- Keys must be unique, meaning a given key can only represent one value
 - but one value may be represented by multiple keys
 - like synonyms in the dictionary.
- Example:
factor: n. See coefficient of X
- *factor* is a key associated with the same value (definition) as the key *coefficient of X*

Clicker 2

- Is it required that the keys and values of a map be the same data type?
- A. No
B. Yes
C. It Depends

Map <String, List<String>>

Movie	Characters
Wizard of Oz	Dorothy, Toto, Scarecrow, Tin Man, Cowardly Lion
Iron Man	Tony Stark, Pepper Potts, Phil Coulson, Obadiah Stane
Pride and Prejudice	Elizabeth Bennet, Jane Bennet, Mr. Darcy, Mr. Bingley
The Avengers	Tony Stark, Pepper Potts, Steve Rogers, Bruce Banner
Sense and Sensibility	Elinor Dashwood, Marianne Dashwood, Edward Ferrars, John Willoughby, Colonel Brandon

The Map<K, V> Interface in Java

- `void clear()`
 - Removes all mappings from this map (optional operation).
- `boolean containsKey(Object key)`
 - Returns true if this map contains a mapping for the specified key.
- `boolean containsValue(Object value)`
 - Returns true if this map maps one or more keys to the specified value.
- `Set<K> keySet()`
 - Returns a Set view of the keys contained in this map.

The Map Interface Continued

- ▶ `V get(Object key)`
 - Returns the value to which this map maps the specified key. Returns null if key not present.
- ▶ `boolean isEmpty()`
 - Returns true if this map contains no key-value mappings.
- ▶ `V put(K key, V value)`
 - Associates the specified value with the specified key in this map

The Map Interface Continued

- ▶ `V remove(Object key)`
 - Removes the mapping for this key from this map if it is present
- ▶ `int size()`
 - Returns the number of key-value mappings in this map.
- ▶ `Collection<V> values()`
 - Returns a collection view of the values contained in this map.

Results with HashMap

Title	Size (kb)	Total Words	Distinct Words	Time List	Time Map
small sample	0.6	89	25	0.001	0.0008
2BR02B	34	5,638	1,975	0.051	0.0140
Alice in Wonderland	120	29,460	6,017	0.741	0.0720
Adventures of Sherlock Holmes	581	107,533	15,213	4.144	0.2500
2008 CIA Factbook	10,030	1,330,100	74,042	173.000	4.0000

Order?

Title	Size	Total Words	Distinct Words	Time List	Time Map
small sample	0.6	89	25	0.001	0.0008
2BR02B	57x	63x	79x	51x	18x
Alice in Wonderland	3.5x	5.2x	3.0x	14.5x	5x
Adventures of Sherlock Holmes	4.8x	3.7x	2.5x	5.6x	3.5x
2008 CIA Factbook	17x	12.3x	5x	42x	16x

$O(\text{Total Words})?$