Collections with Uniqueness: Sets

Richard Warburton

@richardwarburto | www.insightfullogic.com

Sets are collections of distinct elements. There are no duplicates.

Outline

`Set` Features

Worked Example

`Set` Implementations

`SortedSet` Features

Set Features

What does a set offer us?

Set Implementations

HashSet, TreeSet and EnumSet

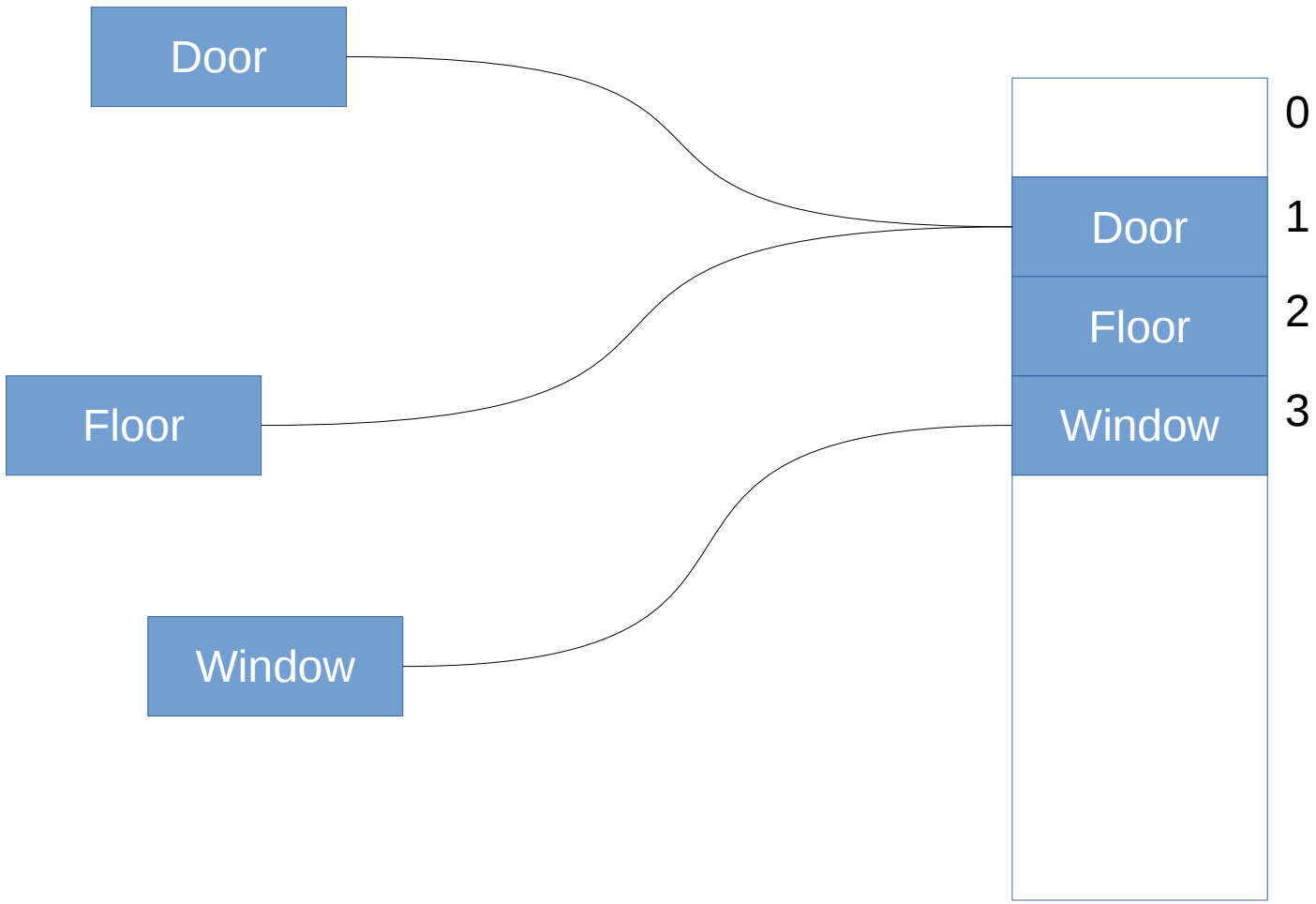
HashSet

Based upon HashMap

Calls `hashCode()` on element and looks up location

Good general purpose implementation

Again resizes when it runs out of space



```
object.equals(other)
```

==>

```
object.hashCode() == other.hashCode()
```

The Equals/Hashcode Contract

NB: One way implication

Correct Hashcodes

- Combine hashcode information from each field
  - `result = 31 * result + obj.hashCode();`
  - `Arrays.hashCode()` for arrays.
  - `(int) (l ^ (l >>> 32))` for longs
  - `Float.floatToIntBits(f);`
- Let your IDE do the heavy lifting
- Or use Objects.hash(...);
- Always use the same fields as equals()

`TreeSet`

Based upon TreeMap

Uses a Binary Tree with a required sort order

Keeps elements in the given order

See `SortedSet` and `NavigableSet`

EnumSet

Specialized implementation for enums

Uses a bitset based upon the ordinal of the enum

Use when storing sets of enums

Algorithmic Performance

	add	contains	next
`HashSet`	O(N), Ω(1)	O(N), Ω(1)	O(Capacity/N)
`TreeSet`	O(log N)	O(log N)	O(log N)
`EnumSet`	O(1)	O(1)	O(1)

SortedSet and NavigableSet

A collection with distinct elements that also have order

```
E first();
E last();

SortedSet<E> tailSet(E fromElement);
SortedSet<E> headSet(E toElement);
SortedSet<E> subSet(E fromElement, E toElement);
```

SortedSet

Defines an order

No indexes, but subset views possible.

```
E lower(E e);
E higher(E e);
E floor(E e);
E ceiling(E e);
E pollFirst();
E pollLast();
```

NavigableSet

Extends `SortedSet`

Provides ways to move through the order

Implemented by `TreeSet`

Conclusions

Summary

- Sets contain distinct elements
- Builtin implementations are based upon hashing or trees
- They can be sorted or navigable


Speed
Cameras