

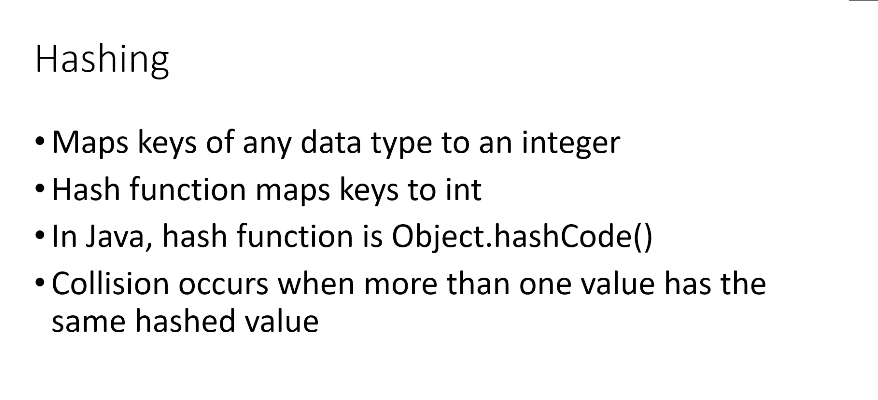
Abstract data type – can be implemented in various ways.

String – common data type used for keys.

Associative arrays are one type of hash tables.

Key data types doesnt have to match.

Map data structures are based on hash tables. Lokup tables...



Keys are converted to integers via hash function. Any data type is converted to integer. Has function is a hash code. Object.hashCode. You can overite it and i tis often hashed.

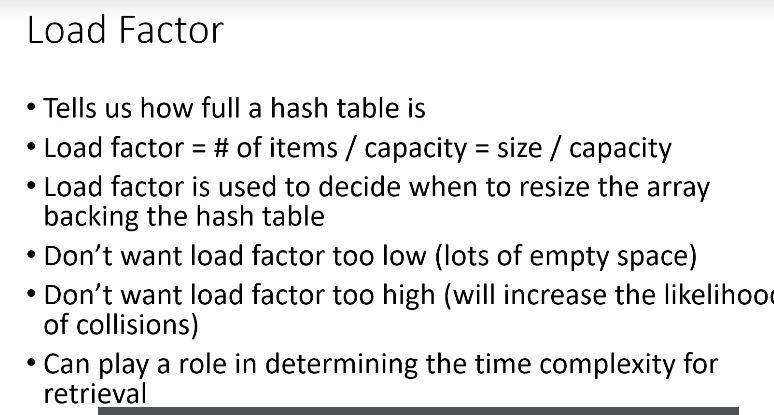
Collisions can happen where the same hashed value is generated for more than one key. The same keys are hashed to same value.

Product number – key, product – value.

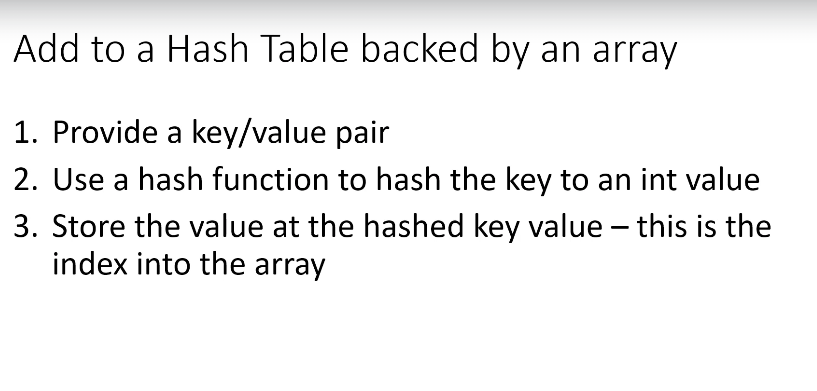
Employee id – employee objects

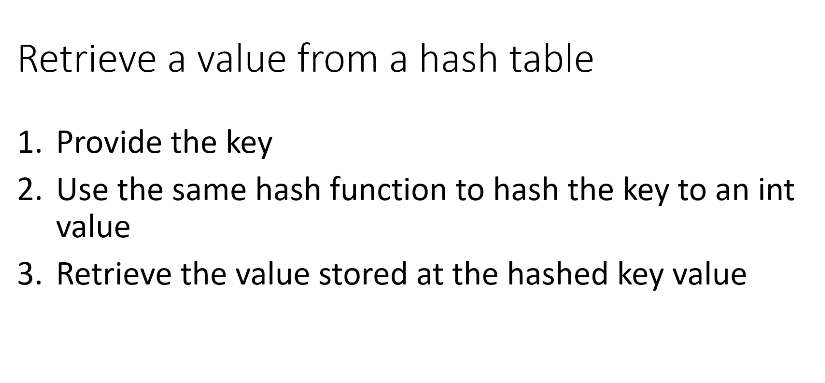
User names – profiles

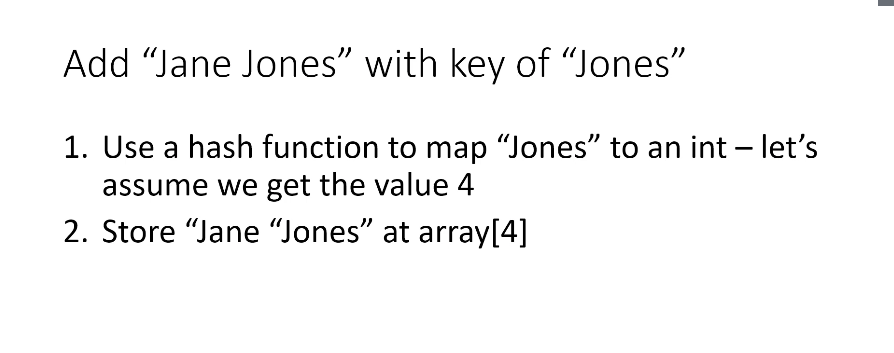
Important value for has table – load factor



We have array of 10 backing our hash table, we have 5 employees in our hash table. 10/ 5 = 0.5 is our load factor. How full is our hash table ? We want to balance this number, we dont want it to be too low (unused) or too high – higher chance of collisions.







Why is array used as a backing structure ? Pragmatic, we can use hashed value as an index of an Array.

Hashing function should very well balance keys inside hashtable.

Long employee id cannot be easily used as indexes to array, because we dont to be constrained by having as large data structure just because the indexes are large numbers.

# Collision resolvance

Open adressing

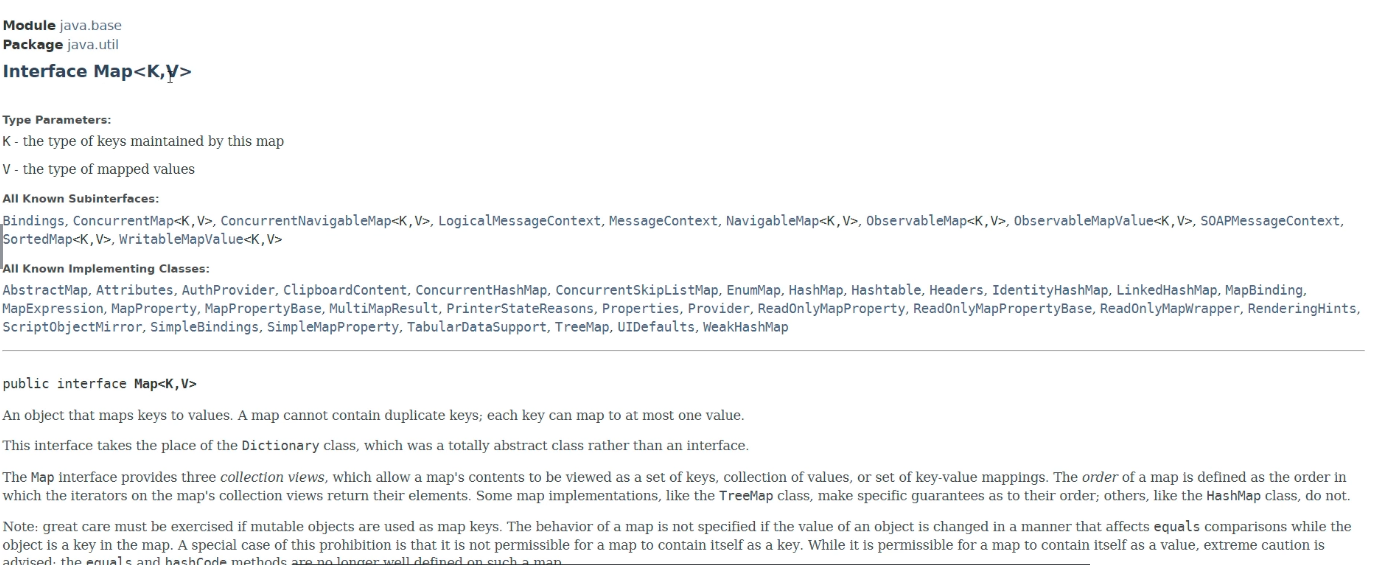
If collision happens we add the key and valllue to next free bucket in the hash table.

Linear probing

IF collision happens, we increase storage index by some for example constant, like + 1 and then we try to store it.

Increment of index is called probe.

# Java SDK



General map in java have only unique keys !!! You cannot have two smiths. It doesnt matter there are no collisions !

# HashMap

Allows storing null values for keys. You can set load factor, if hashmap exceeds it it automatically resizes the map. Its not synchronized, you should use Collections.SynchronizedMap method.

HashTable doesnt allow null values and its thread safe. If we dont need synchronization we use hashmap. Synchronized structures comes with overhead.

There is also an ConcurrentHashMap.

# Bucket Sort

