

# Announcements

MP6 due on Tue, 04/19 @11:59pm, E/C (harder) due on Tue, 04/12 @ 11:59pm.

**Exam 3 results.** MC average: 75/100, Coding average: 65/100. How many 100% on coding? ~80/700

**Exam 4 topics.** MC: AVL trees, b-trees, hash tables, heaps, disjoint sets. Coding: AVL trees, hash tables.

## Hashing - using “hash tables” to implement \_\_\_\_\_

Suppose we have the following info...

Locker Number	Name
103	Jay Hathaway
92	Linda Stencel
330	Bonnie Cook
46	Rick Brown
124	Kim Petersen
...	...

...and we want to be able to retrieve a name, given a locker number.

Now suppose our keys are not so nicely described...

Course Number -> Schedule info

Color -> BMP

Vertex -> Set of incident edges

Flight number -> arrival information

URL -> html page

dice roll -> payoff amt

# Some general vocabulary

*A dictionary* is a structure supporting the following:

```
void insert(kType & k, dType & d)
```

```
void remove(kType & k)
```

```
dType find(kType & k)
```

*An associative array* is a dictionary w a particular interface—

Overloads the [] operator for insert and find:

```
myDictionary[ ] = 22;
```

```
dType d = myDictionary[ ];
```

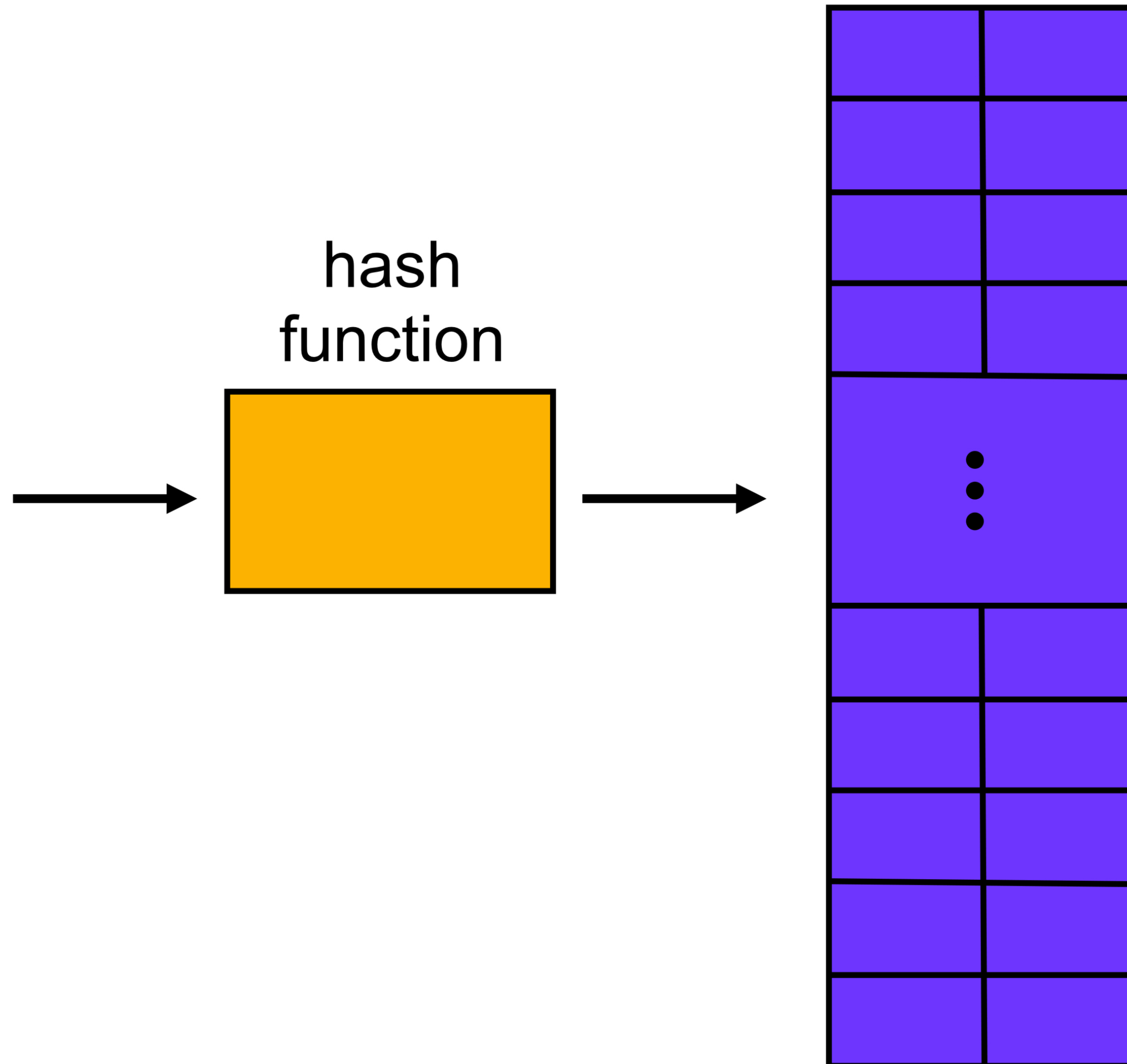

# Hashing:

(defn) **Keyspace** — a (mathematical) description of the keys for a set of data.

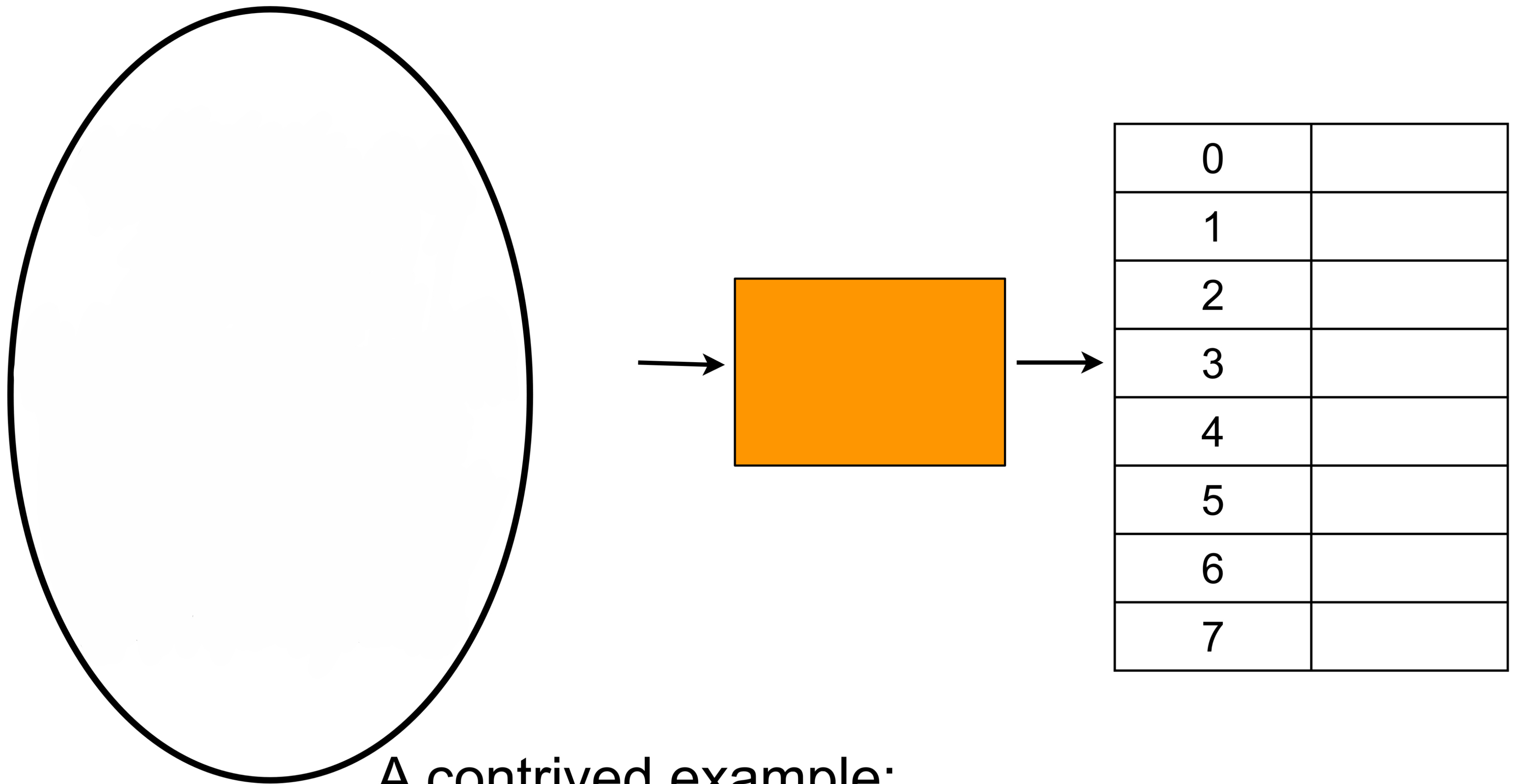
Goal: use a function to map the keyspace into a small set of integers.

What's fuzzy about this goal?

Basic Idea: we seek a mapping,  $h(k)$



A perfect hash function:



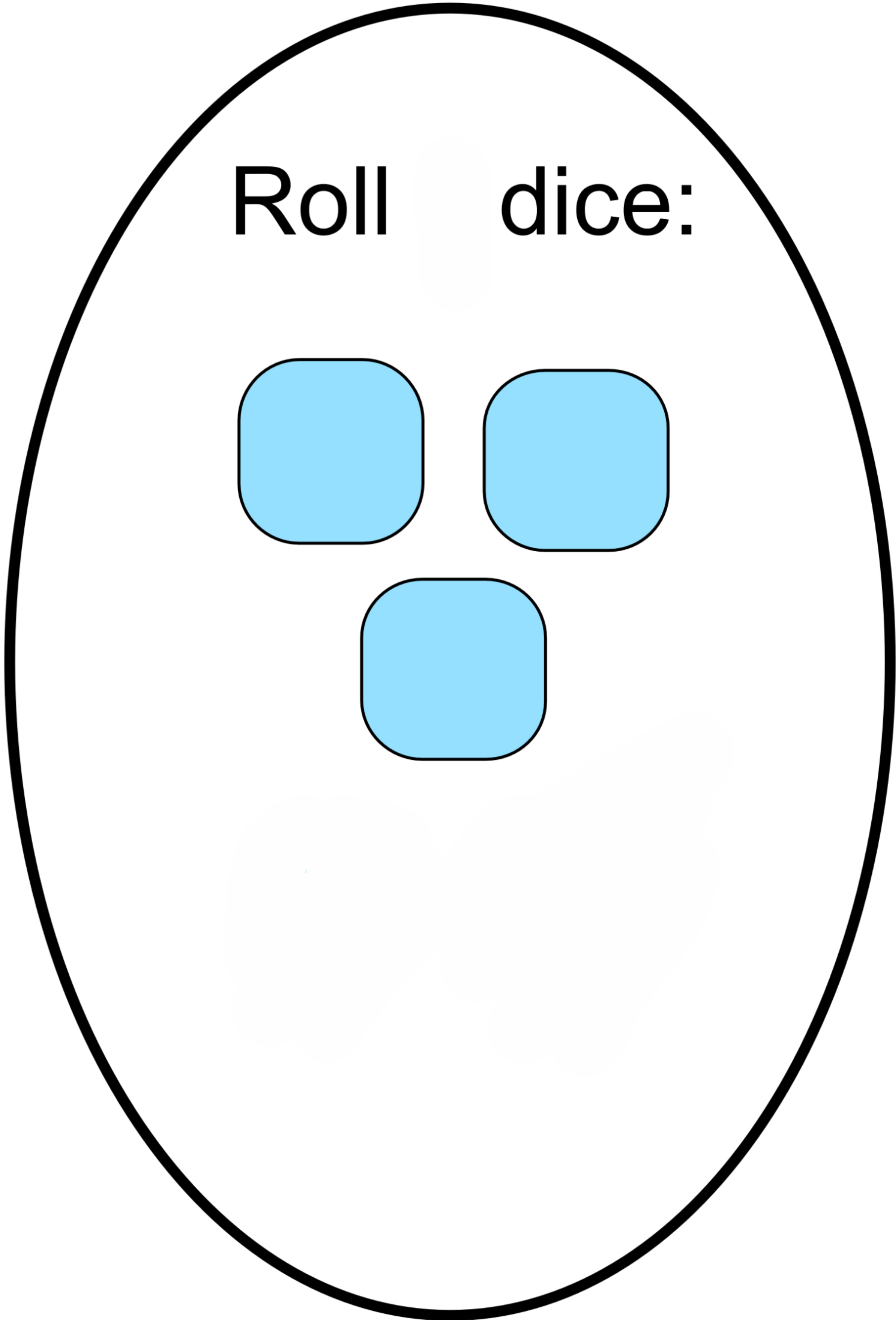
A contrived example:

*these keys have a fabulous hash fn.*

*a. each key hashes to a different int*

*b. collection of keys hash to a seq of ints*

A perfect hash function?



Is this a good hash function?

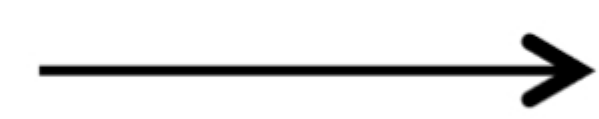
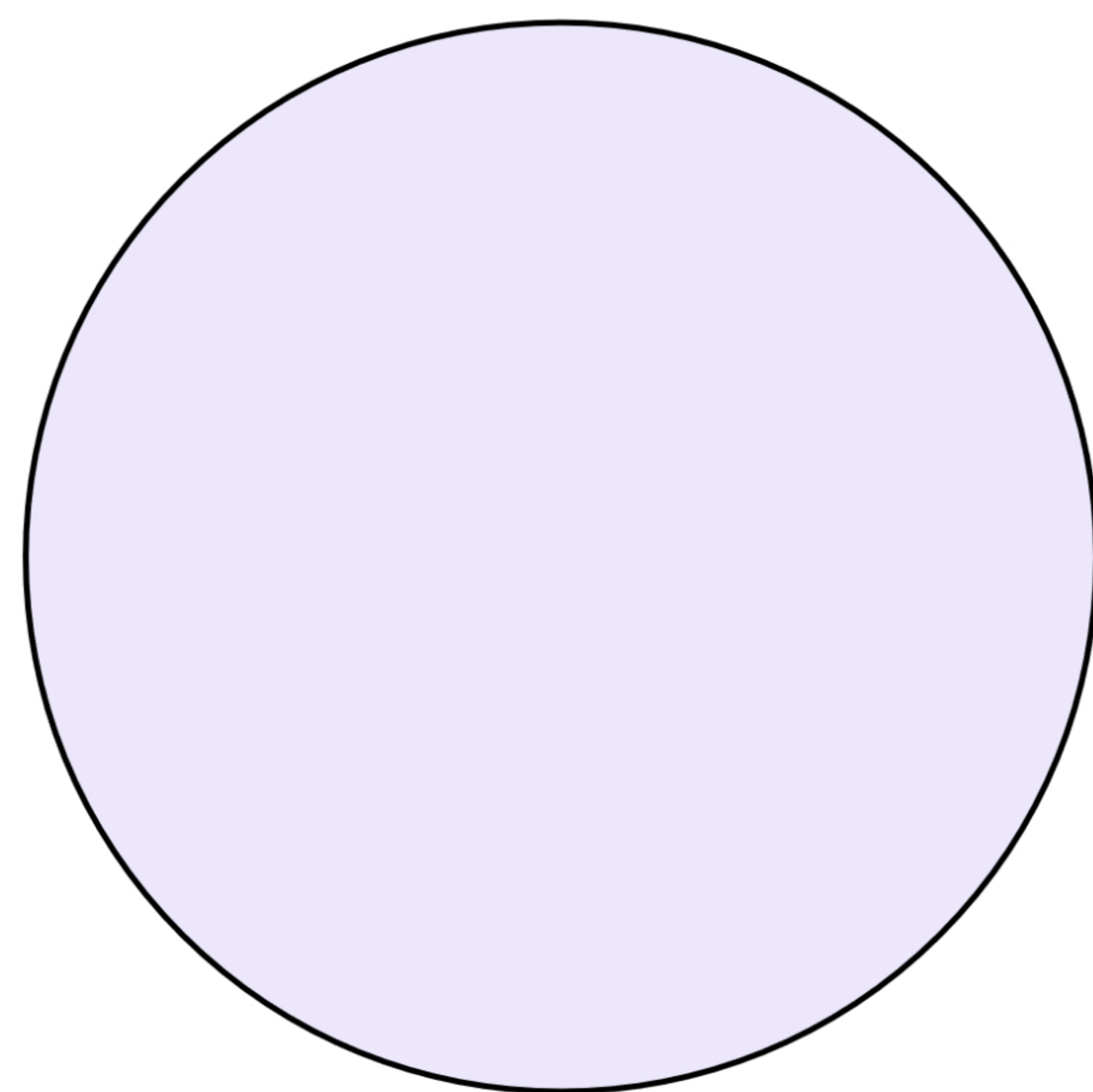


0	
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	

A Hash Table consists of:

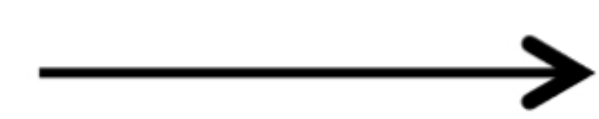
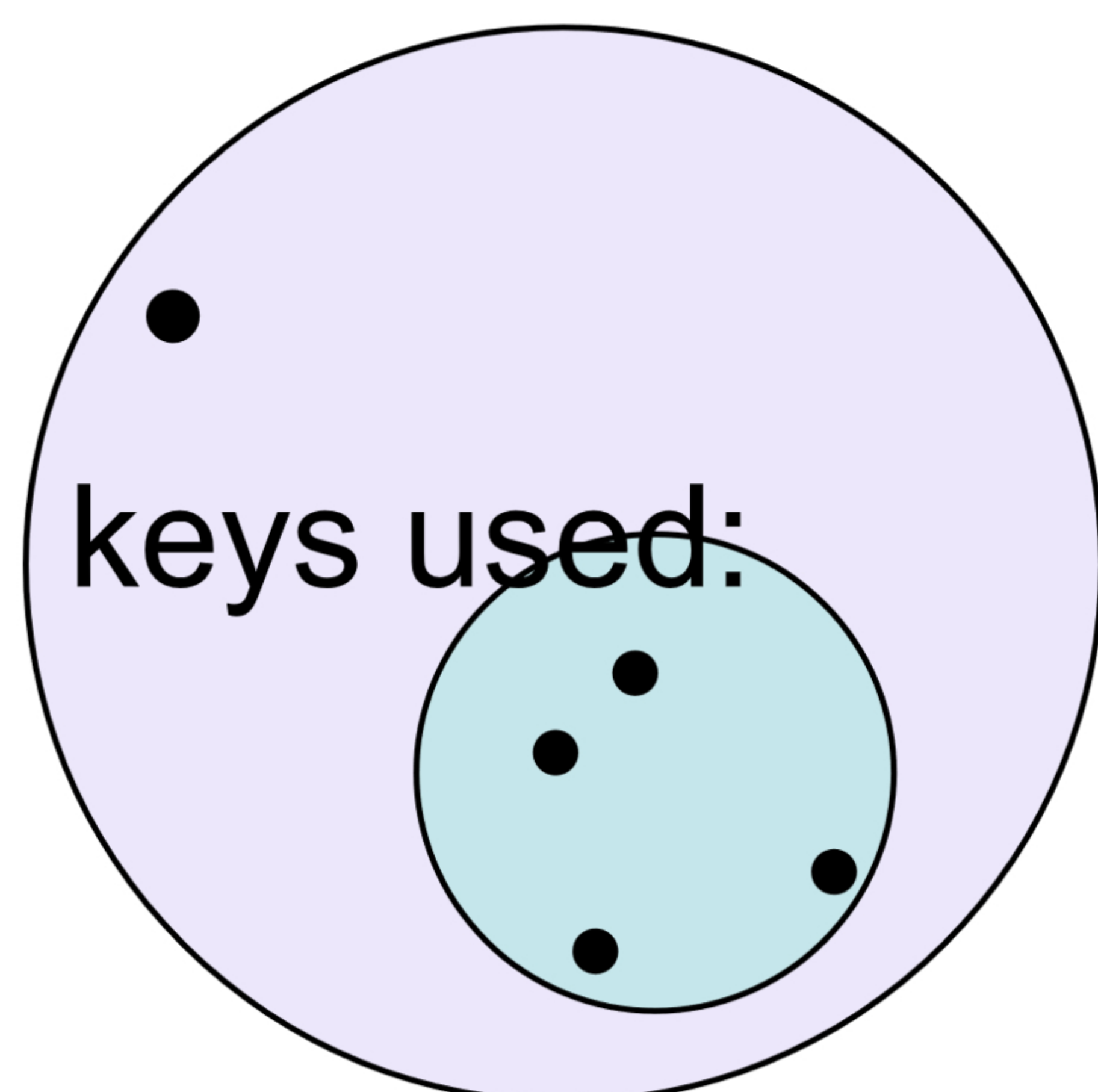
- An array
- A hash function
- A collision resolution strategy

KeySpace



0	1		...				N-1

Easy, if  
 $|\text{KeySpace}| \sim N$



0	1		...				N-1

# Hash Functions:

- Consist of 2 parts:
  - A Hash: Function mapping a key to an integer  $i$
  - A compression: function mapping  $i$  into the array cells 0 to  $N-1$ .
- Choosing a hash function is tricky...
  - Don't create your own (yet)
  - Smart people can produce poor hash functions (what's a bad hash function?)
    - Knuth's multiplicative hash in "the Art of Computer Programming"
- Characteristics:
  - Computed in \_\_\_\_\_ time.
  - Deterministic.
  - Satisfy the SUHA.