Foundations of Computer Science Comp109

University of Liverpool

Boris Konev

konev@liverpool.ac.uk

Olga Anosova

O.Anosova@liverpool.ac.uk

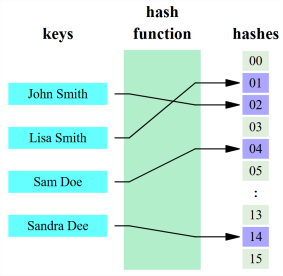
Recap: composition and extended pigeonhole principle.

- **Composition** $g \circ f : X \to Z$ of two functions $f : X \to Y$ and $g : Y \to Z$ is defined by $(g \circ f)(x) = g(f(x))$.
- The **extended pigeonhole principle**: if $|A| > k|B|$ for some $k \in \mathbb{N}$ and a function $f : A \to B$, then there is a value of f which occurs at least $k + 1$ times.
- There are multiple applications of Pigeonhole principle, for example: The adult human body has about 5 million hair follicles (men have a few hundred thousand more than women). Earth population is about 8.2 Billion. What can you deduce?

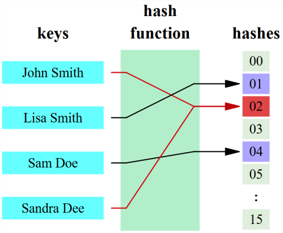
Less fun example: Hash collision

A *hash function* is any function that maps data of arbitrary size to fixed-size values.
Example: password → number of symbols.

A *perfect hash function* maps distinct elements to distinct values, i.e. it is



Problem: function domain needs to be known, hence the perfect hash function either is not dynamic or requires excessive codomain size.
In practice all hash functions are *imperfect* with some additional *hash collision* resolution methods.[1]



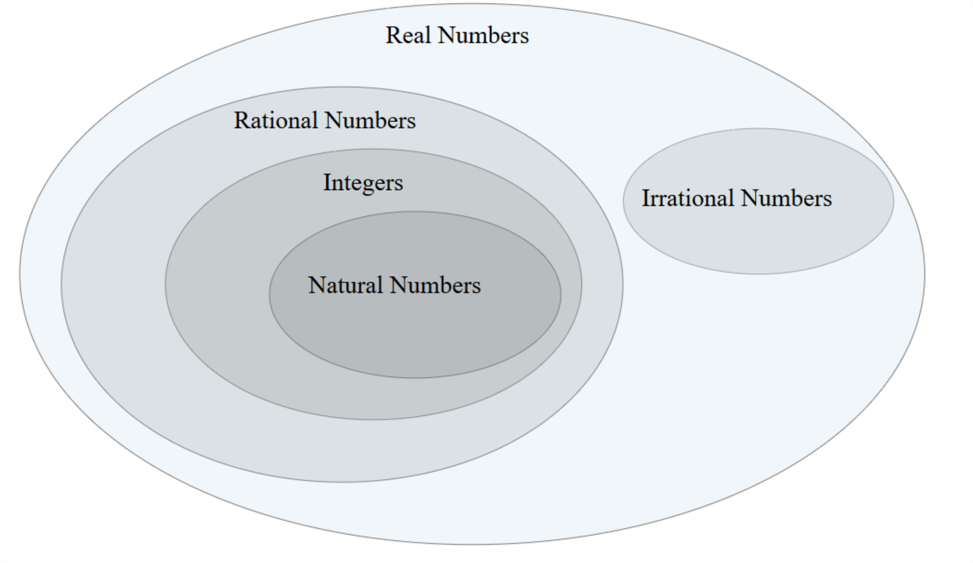
[1] For images and more, see https://en.wikipedia.org/wiki/Hash_collision

The Birthday paradox

The pigeonhole principle guarantees the result with probability$=1$, but the probability is getting $>$50% very quick.

The Birthday paradox states that
- to *guarantee* having two people with the same birthday one needs
how many people?
- the probability to match a *fixed specific birthday* is very low;
- to exceed 50% *probability* of having a birthday match we need
how many people?

Comparing number sets



Hilbert's infinite hotel

A hotel with *infinitely many rooms* (we can number them 1, 2, 3,...) is full.
A new guest is arriving. Will the full hotel be able to locate a room for them?

Bijections and cardinality

We can't **count** infinite cardinalities, but we can **compare** them.

Sets A and B have the same cardinality iff there is a bijection from A to B.

Recall: the powerset and bit vectors

We used this trick to compute the cardinality of the powerset:

Let $S = \{1, 2, \ldots, n\}$ and let B^n be the set of bit strings of length n. The function

$$f : Pow(S) \to B^n$$

that assigns each subset A of S to its *characteristic vector* is a bijection.

Comparing $\mathbb{N}$ and $\mathbb{Z}$

$$\cdots \quad -4 \quad -3 \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad 3 \quad 4 \quad \cdots$$

Countable sets

A set that is either **finite** or has the **same cardinality as** $\mathbb{N}$ is called *countable*.

Cardinality of $\mathbb{N}$ is often denoted as $\aleph_0$ (read: *aleph-nought*, *aleph-zero*, or *aleph-null*).

(DIY) Claim. A set A is countable if and only if there exists an injection from A into $\mathbb{N}$.

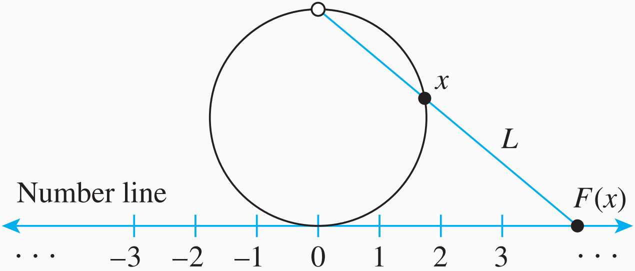
Example: are even integers countable?

Countable Sets: $\mathbb{Q}$

$$
\begin{array}{ccccccc}
\frac{1}{1} & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \cdots \\
\frac{2}{1} & \frac{2}{2} & \frac{2}{3} & \frac{2}{4} & \frac{2}{5} & \frac{2}{6} & \cdots \\
\frac{3}{1} & \frac{3}{2} & \frac{3}{3} & \frac{3}{4} & \frac{3}{5} & \frac{3}{6} & \cdots \\
\frac{4}{1} & \frac{4}{2} & \frac{4}{3} & \frac{4}{4} & \frac{4}{5} & \frac{4}{6} & \cdots \\
\frac{5}{1} & \frac{5}{2} & \frac{5}{3} & \frac{5}{4} & \frac{5}{5} & \frac{5}{6} & \cdots \\
\frac{6}{1} & \frac{6}{2} & \frac{6}{3} & \frac{6}{4} & \frac{6}{5} & \frac{6}{6} & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \ddots
\end{array}
$$

Real numbers: $\{x \in \mathbb{R} \mid 0 < x < 1\}$ and $\mathbb{R}^+$

$\{x \in \mathbb{R} \mid 0 < x < 1\}$ and $\mathbb{R}$



Uncountable sets

A set that is *not countable* is called *uncountable*.

Claim. Set $S = \{x \in \mathbb{R} \mid 0 < x < 1\}$ is uncountable.

Proof: Cantor's diagonal argument

Suppose for a proof by contradiction that there exists a bijection $f : \mathbb{N}^+ \to S$. Consider decimal representations of $f(n)$, for $n \in \mathbb{N}^+$:

$$
\begin{array}{l}
f(1) = 0.a_{11}\ a_{12}\ a_{13} \ldots\ a_{1n} \ldots \\
f(2) = 0.a_{21}\ a_{22}\ a_{23} \ldots\ a_{2n} \ldots \\
f(3) = 0.a_{31}\ a_{32}\ a_{33} \ldots\ a_{3n} \ldots \\
\vdots \qquad\qquad\qquad\qquad \ldots \\
f(n) = 0.a_{n1}\ a_{n2}\ a_{n3} \ldots\ a_{nn} \ldots \\
\vdots
\end{array}
$$

We show that there exists $d \in S$ such that for no $i \in \mathbb{N}^+$ we have $f(i) = d$.

Let $d = 0.d_1\ d_2\ d_3 \ldots d_n \ldots$ where $\qquad d_i = \begin{cases} 2, & \text{if } a_{ii} = 1 \\ 1, & \text{if } a_{ii} \neq 1 \end{cases}$

Then what?

Summary

- Sets A and B have the same cardinality iff there is a bijection from A to B.
- A set that is either **finite** or has the **same cardinality as $\mathbb{N}$** is called **countable**. Otherwise a set is called **uncountable**.
- Sets $\mathbb{Z}, \mathbb{Q}$ are countable, set $\mathbb{R}$ is not.

DIY Problem. Prove that any subset of any countable set is countable.