Compilers and Formal Languages

Email: christian.urban at kcl.ac.uk
Slides & Progs: KEATS (also homework is there)

1 Introduction, Languages	6 While-Language
2 Regular Expressions, Derivatives	7 Compilation, JVM
3 Automata, Regular Languages	8 Compiling Functional Languages
4 Lexing, Tokenising	9 Optimisations
5 Grammars, Parsing	10 LLVM

```scala
def fact(n: Int) : Int = {
  if (n == 0) 1 else n * fact(n - 1)
}


def factC(n: Int, ret: Int => Int) : Int = {
  if (n == 0) ret(1)
  else factC(n - 1, x => ret(n * x))
}

fact(10)
factC(10, identity)
```

```
def fibC(n: Int, ret: Int => Int) : Int = {
  if (n == 0 || n == 1) ret(1) else
  fibC(n - 1,
       r1 => fibC(n - 2,
        r2 => ret(r1 + r2)))
}

fibC(10, identity)
```

Are there more strings in
$$L(a^*) \text{ or } L\big((a+b)^*\big)?$$

Can you remember this HW?

- (1) How many basic regular expressions are there to match the string *abcd*?
- (2) How many if they cannot include **1** and **0**?
- (3) How many if they are also not allowed to contain stars?
- (4) How many if they are also not allowed to contain $\_ + \_$?

There are more problems, than there are programs.

There are more problems, than there are programs.

There must be a problem for which there is no program.

Subsets

If $A \subseteq B$ then A has fewer or equal elements than B

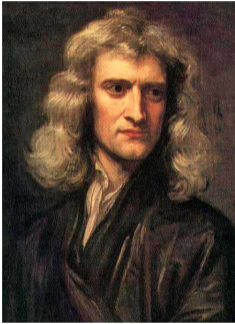$A \subseteq B$ and $B \subseteq A$
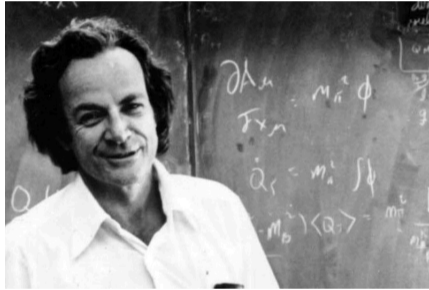
then $A = B$

5 elements

3 elements

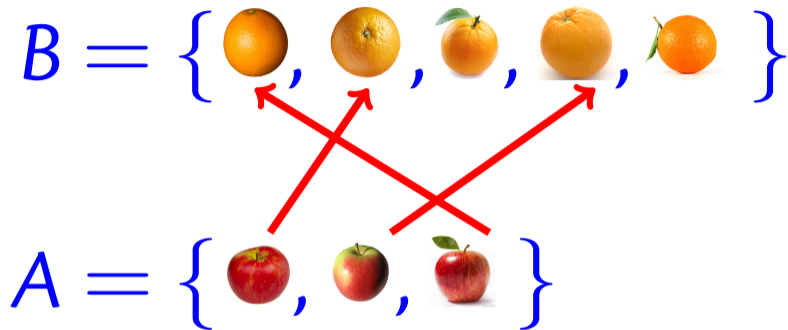# Newton vs Feynman

classical physics

quantum physics

The Goal of the Talk

- show you that something very unintuitive happens with very large sets

- convince you that there are more **problems** than **programs**

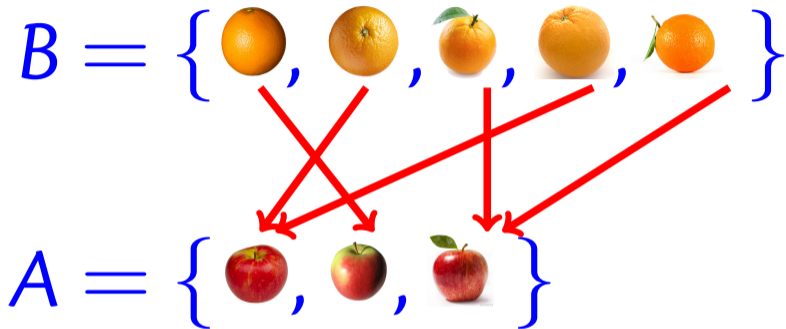$$B = \{\ \bigcirc,\ \bigcirc,\ \bigcirc,\ \bigcirc,\ \bigcirc\ \}$$

$$A = \{\ \bullet,\ \bullet,\ \bullet\ \}$$

$$|A| = 5,\ |B| = 3$$

then $|A| \leq |B|$

$B = \{$ 🍊 , 🍊 , 🍊 , 🍊 , 🍊 $\}$

$A = \{$ 🍎 , 🍎 , 🍎 $\}$

for $=$ has to be a **one-to-one** mapping

Cardinality

$|A| \stackrel{\text{def}}{=}$ "how many elements"

$A \subseteq B \Rightarrow |A| \leq |B|$

Cardinality

$|A| \stackrel{\text{def}}{=}$ "how many elements"

$A \subseteq B \Rightarrow |A| \leq |B|$

if there is an injective function
$f : A \rightarrow B$ then $|A| \leq |B|$

$$\forall xy. \, f(x) = f(y) \Rightarrow x = y$$

$A = \{$ 🍊 , 🍊 , 🍊 $\}$

$B = \{$ 🍎 , 🍎 , 🍎 $\}$
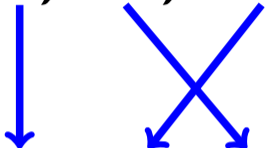
then $|A| = |B|$

Natural Numbers

$$\mathbb{N} \stackrel{\text{def}}{=} \{0, 1, 2, 3, \ldots\ldots\}$$

Natural Numbers

$$\mathbb{N} \stackrel{\text{def}}{=} \{0, 1, 2, 3, \dots\}$$

A is countable iff $|A| \leq |\mathbb{N}|$

First Question

$$|\mathbb{N} - \{0\}| \quad \textcolor{red}{?} \quad |\mathbb{N}|$$

$\geq$ or $\leq$ or $=$?

First Question

$$|\mathbb{N} - \{0\}| \quad ? \quad |\mathbb{N}|$$

$\geq$ or $\leq$ or $=$?

$x \mapsto x + 1$,
$$|\mathbb{N} - \{0\}| = |\mathbb{N}|$$

$$|\mathbb{N} - \{0, 1\}| \quad ? \quad |\mathbb{N}|$$

$$|\mathbb{N} - \{0, 1\}| \quad ? \quad |\mathbb{N}|$$

$$|\mathbb{N} - \mathbb{O}| \quad ? \quad |\mathbb{N}|$$

$$\mathbb{O} \overset{\text{def}}{=} \text{odd numbers} \quad \{1, 3, 5 \ldots\ldots\}$$

$$|\mathbb{N} - \{0, 1\}| \quad ? \quad |\mathbb{N}|$$

$$|\mathbb{N} - \mathbb{O}| \quad ? \quad |\mathbb{N}|$$

$\mathbb{O} \stackrel{\text{def}}{=}$ odd numbers $\quad \{1, 3, 5 \ldots\ldots\}$
$\mathbb{E} \stackrel{\text{def}}{=}$ even numbers $\quad \{0, 2, 4 \ldots\ldots\}$

$$\left| \mathbb{N} \cup -\mathbb{N} \right| \quad ? \quad \left| \mathbb{N} \right|$$

$\mathbb{N} \stackrel{\text{def}}{=}$ positive numbers $\quad \{0, 1, 2, 3, \dots\}$

$-\mathbb{N} \stackrel{\text{def}}{=}$ negative numbers $\quad \{0, -1, -2, -3, \dots\}$

A is countable if there exists an injective $f : A \rightarrow \mathbb{N}$

A is uncountable if there does not exist an injective $f : A \rightarrow \mathbb{N}$

countable: $|A| \leq |\mathbb{N}|$
uncountable: $|A| > |\mathbb{N}|$

A is countable if there exists an injective $f : A \rightarrow \mathbb{N}$

A is uncountable if there does not exist an injective $f : A \rightarrow \mathbb{N}$

countable: $|A| \leq |\mathbb{N}|$
uncountable: $|A| > |\mathbb{N}|$

Does there exist such an *A* ?

Hilbert's Hotel
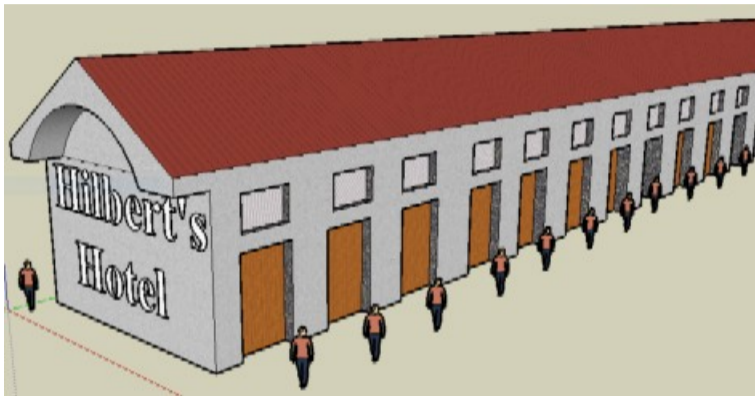
- …has as many rooms as there are natural numbers

Real Numbers between 0 and 1

Real Numbers between 0 and 1

Real Numbers between 0 and 1

1	4	3	3	3	3	3	$\cdots$	$\cdots$
2	1	3	3	4	5	6	7	
3	0	1	0	1	0	$\cdots$		
4	7	8	5	3	9	$\cdots$		

$\cdots$

Real Numbers between 0 and 1

Real Numbers between 0 and 1

1	4	3	3	3	3	3	$\cdots$	$\cdots$
2	1	3	3	4	5	6	7	
3	0	1	1	1	0	$\cdots$		
4	7	8	5	4	9	$\cdots$		

$\cdots$

Real Numbers between 0 and 1

1	4	3	3	3	3	3	$\cdots$	$\cdots$
2	1	3	3	4	5	6	7	
3	0	1	1	1	0	$\cdots$		
4	7	8	5	4	9	$\cdots$		

$\cdots$

$$|\mathbb{N}| < |R|$$

The Set of Problems

$\aleph_0$

	0	1	2	3	4	5	$\cdots$	
1	0	1	0	1	0	1	$\cdots$	$\cdots$
2	0	0	0	1	1	0	0	
3	0	0	0	0	0	$\cdots$		
4	1	1	0	1	1	$\cdots$		

$\cdots$

The Set of Problems

$\aleph_0$

	0	1	2	3	4	5	$\cdots$
1	0	1	0	1	0	1	$\cdots$	$\cdots$
2	0	0	0	1	1	0	0
3	0	0	0	0	0	$\cdots$	
4	1	1	0	1	1	$\cdots$	

$\cdots$

$$|\text{Progs}| = |\mathbb{N}| < |\text{Probs}|$$

Halting Problem

Assume a program H that decides for all programs A and all input data D whether

- $H(A, D) \overset{\text{def}}{=} 1$ iff $A(D)$ terminates
- $H(A, D) \overset{\text{def}}{=} 0$ otherwise

Halting Problem (2)

Given such a program H define the
following program C: for all programs A

- $C(A) \overset{\text{def}}{=} 0$ iff $H(A, A) = 0$
- $C(A) \overset{\text{def}}{=} \text{loops}$ otherwise

Contradiction

$H(C, C)$ is either 0 or 1.

- $H(C, C) = 1 \overset{\text{def } H}{\Rightarrow} C(C) \downarrow \overset{\text{def } C}{\Rightarrow} H(C, C) = 0$

- $H(C, C) = 0 \overset{\text{def } H}{\Rightarrow} C(C) \text{ loops } \overset{\text{def } C}{\Rightarrow}$

$$H(C, C) = 1$$

Contradiction in both cases. So H cannot exist.

Take Home Points

- there are sets that are more infinite than others

- even with the most powerful computer we can imagine, there are problems that cannot be solved by any program

- in CS we actually hit quite often such problems (halting problem)