Halting Problem

One of the most philosophically important theorems of the theory of computation.

Computers (and computation) are limited in a very fundamental way.

Common, every-day problems are unsolvable.

- does a program sort an array of integers?
- both program and specification are precise mathematical objects.
- proving program $\cong$ specification should be just like proving that triangle 1 $\cong$ triangle 2 . . .

No.

Halting Problem

Does a Turing machine accept a string?

$$A_{\text{TM}} = \{\langle M, w\rangle | M \text{ is a TM that accepts } w\}$$

Theorem: A_{TM} is undecidable.

Recall that A_{DFA}, A_{NFA}, and A_{CFG} are decidable.

Halting Problem

$$A_{\mathsf{TM}} = \{\langle M, w\rangle | M \text{ is a TM that accepts } w\}$$

But, first note that

Theorem: A_{TM} is enumerable.

Define the machine U.

On input $\langle M, w\rangle$, where M is a TM and w a string

1. Simulate M on input w.
2. If M ever enters its accept state, *accept*, and if M ever enters its reject state, *reject*.

Remarks

- if M loops, U loops.
- U is called a *universal* Turing machine.
- inspired the stored-program computer

Diagonalization

Question: What does it mean to say that two infinite sets are the *same size*?

Answered by Georg Cantor in 1873: pair them off.

Correspondence

Question: What does it mean to say that sets A and B are the *same size*?

A map $f : A \to B$ is a *correspondence* if f is

- one-to-one: $f(a) \neq f(b)$ if $a \neq b$.
- onto: for every $b \in B$, there is an a such that $f(a) = b$.

Answer: A and B are the *same size* if there is a correspondence from A to B.

Correspondence

Question: In a crowded room, how can we tell if there are more people than chairs, or more chairs than people?

Answer: Establish a correspondence: ask everyone to sit down.

(C.f., Mathematician's trick for counting a herd of cows . . .)

Correspondence

Claim: The set $\mathcal{N}$ of natural numbers has the same size as the set $\mathcal{E}$ of even numbers

Proof: Let $f(i) = 2i$.

Remark: A proper subset of A can be the same size as A!

Definition: A set A is *countable* if

- either it is finite, or
- it has the same size as $\mathcal{N}$, the natural numbers.

We have just seen that $\mathcal{E}$ is countable.

A countable set is sometimes said to have size $\aleph_0$.

Claim: The set $\mathcal{Z}$ of integers is countable.

Proof: Define $f : \mathcal{N} \to \mathcal{Z}$ by

$$f(i) = \begin{cases} i/2 & \text{if } i \text{ is even} \\ -(\lfloor i/2 \rfloor + 1) & \text{if } i \text{ is odd} \end{cases}$$

Pop Quiz!

In Heaven, there is a hotel with a countable number of rooms.

One day, the society of Prophets, Oracles, and AI Researchers holds a convention that books every room in the hotel.

Then one more guest arrives, angrily demanding a room.

You are the manager. What do you do?

Pop Quiz!

One more guest arrives, angrily demanding a room.

Answer: Ask the guest in room i to move to room $i+1$, and put the newcomer in room 1.

Pop Quiz #2!

Then a countable number of guests arrive, all angrily demanding rooms. (What a noise!)

Now what do you do?

Pop Quiz #2!

Then a countable number of additional guests arrive, all angrily demanding rooms.

Answer: Ask the guest in room i to move to room $2i$, and put the newcomers in the odd-numbered rooms.

Rational Numbers

Let

$$\mathcal{Q} = \left\{ \frac{m}{n} | m, n \in \mathcal{N} \right\}$$

Theorem: $\mathcal{Q}$ is countable.

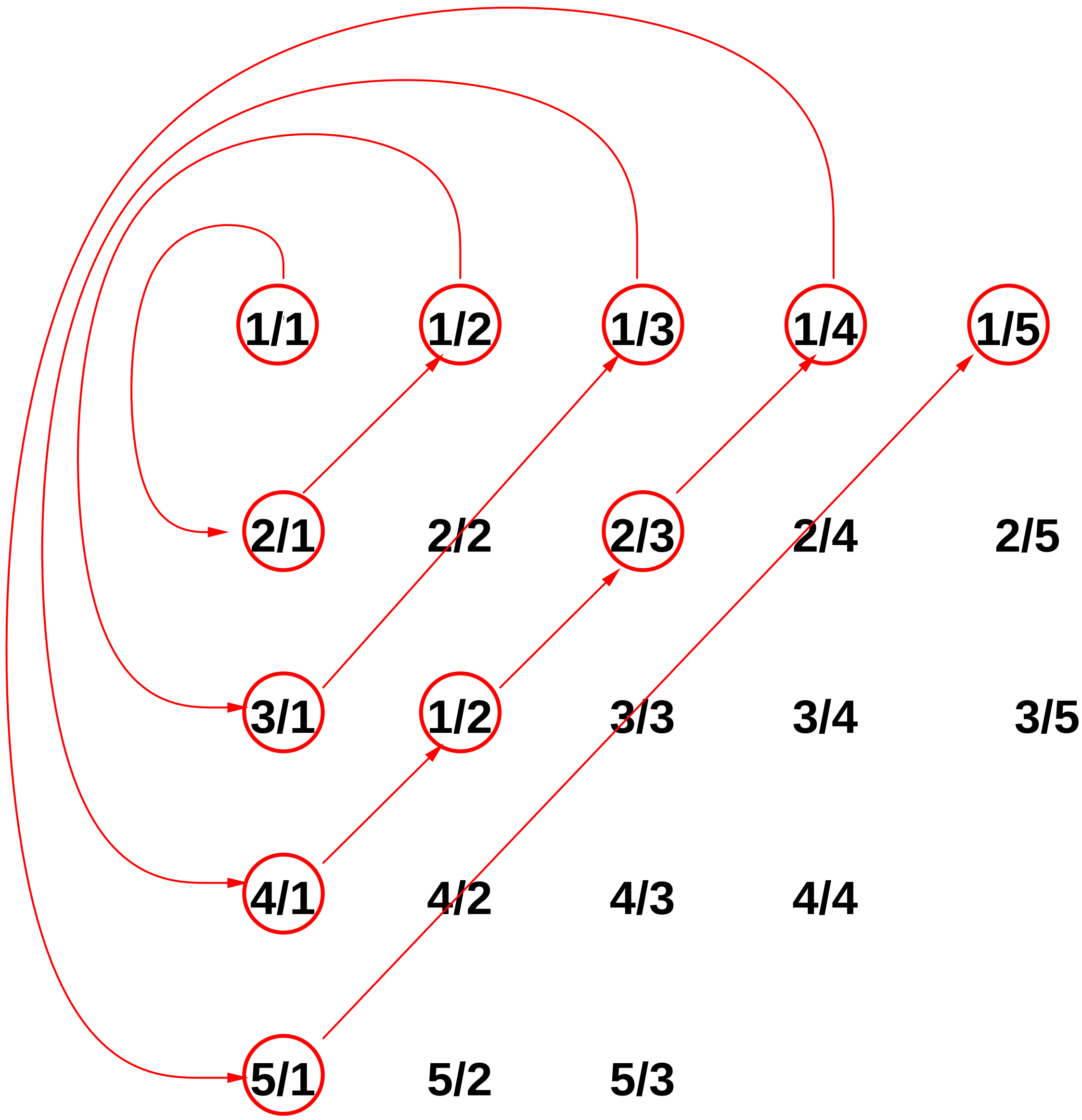
This may seem counterintuitive.

Idea

- list $\mathcal{Q}$ as 2-dim array
- begin counting with the first row . . .

Why doesn't this work?

Rational Numbers



- enumerate numbers along Northeast and Southwest diagonals, and
- skip duplicates.

Does this mean that any infinite set is countable?

The Real Numbers

A *real number* (simplifying somewhat) has a decimal representation.

Such as $\pi = 3.1415926\ldots$, $\sqrt{2} = 1.4142136\ldots$, and $0 = 0.0000000\ldots$.

Let $\mathcal{R}$ be the set of real numbers.

Theorem: $\mathcal{R}$ is uncountable.

$\mathcal{R}$ is sometimes said to have size $\aleph_1$.

- Cantor's historic proof
- introduced the *diagonalization* method.

The Real Numbers

Assume there is a correspondence between $\mathcal{N}$ and $\mathcal{R}$.

Write it down:

n	$f(n)$
1	`3.14159...`
2	`55.55555...`
3	`40.18642...`
4	`15.20601...`

We now show that there is a number x not in this list.

Diagonalization

Pick $0 \leq x \leq 1$, so non-zero digits follow decimal point.

Make sure that x differs from $f(n)$ for all n.

n	$f(n)$
1	3.14159...
2	55.5555...
3	40.1864...
4	15.2060...

- First fractional digit of $f(1)$ is 1
- Pick first fractional digit of x to be something else (say, 2).
- Second fractional digit of $f(2)$ is 5
- Pick second fractional digit of x to be something else (say, 6).
- and so on ...

$$x = 0.2691\ldots$$

Diagonalization

Similarly, there are languages that are not enumerable.

- the set of Turing machines is countable, but
- the set of languages is uncountable!

Ergo,

- there exist languages that are not enumerable why?
- indeed, “most” languages are not enumerable explain

Turing Machines are Countable

Claim: The set of strings in Σ^* is countable.

Proof: List strings of length 1, then length 2, and so on.

Turing Machines are Countable

Claim: The set of Turing machines is countable.

Proof: Each TM M has an encoding as a string $\langle M \rangle$.

We can list all strings, and just omit the ones that are not legal TM encodings.

Languages are Uncountable

Let $\mathcal{B}$ be the set of of infinite binary sequences.

Claim: $\mathcal{B}$ is uncountable.

Proof: Diagonalization argument essentially identical to the proof that $\mathcal{R}$ is uncountable.

(additional helpful clue: think of binary sequence as binary expansion!)

Languages are Uncountable

Let $\mathcal{L}$ be the set of languages over alphabet Σ.

We give a correspondence

$$\chi : \mathcal{L} \to \mathcal{B}$$

called the language's *characteristic sequence*.

- Let $\Sigma^* = \{s_1, s_2, s_3, \ldots\}$ (in lexicographic order).
- Each language $L \in \mathcal{L}$ has a unique sequence $\chi(L) \in \mathcal{B}$:
- the i-th bit of $\chi(L)$ is 1 if and only if $s_i \in L$.

Languages are Uncountable

Each language $L \in \mathcal{L}$ has a unique sequence $\chi(L) \in \mathcal{B}$: the i-th bit of $\chi(L)$ is 1 if and only if $s_i \in L$.

Example:

$$
\begin{array}{llllllllll}
\Sigma^* & \{\varepsilon, & 0, & 1, & 00, & 01, & 10, & 11, & 000 & \ldots\} \\
A & \{ & 0, & & 00, & 01, & & & 000 & \ldots\} \\
\chi(A) & \{0, & 1, & 0 & 1, & 1, & 0 & 0 & 1 & \ldots\}
\end{array}
$$

The map $\chi : \mathcal{L} \to \mathcal{B}$

- is one-to-one, why?
- and onto why?
- and is hence a correspondence.

It follows that $\mathcal{L}$ is uncountable.