CSE 2001:
Introduction to Theory of Computation

Winter 2013

Suprakash Datta

datta@cse.yorku.ca

Office: CSEB 3043

Phone: 416-736-2100 ext 77875

Course page: http://www.cse.yorku.ca/course/2001

Some of these slides are adapted from Wim van Dam's slides (www.cs.berkeley.edu/~vandam/CS172/ retrieved earlier)

Next

Towards undecidability:

- **The Halting Problem**
- **Countable and uncountable infinities**
- **Diagonalization arguments**

The Halting Problem

The existence of the universal TM U shows that
$A_{TM} = \{\langle M,w \rangle \mid M \text{ is a TM that accepts } w\}$
is TM-recognizable, but can we also *decide* it?

The problem lies with the cases when M does not halt on w. In short: <u>the halting problem</u>.

We will see that this is an insurmountable problem: in general one cannot decide if a TM will halt on w or not, hence A_{TM} is undecidable.

Counting arguments

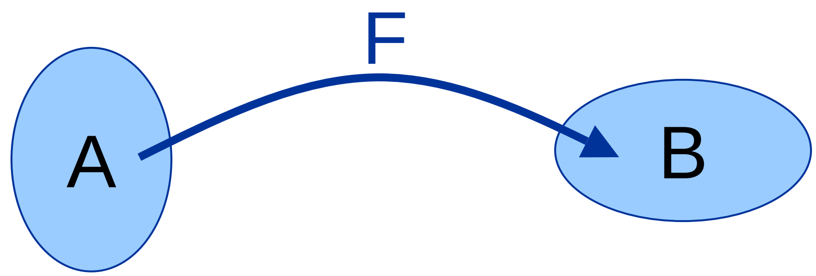
- We need tools to reason about undecidability.
- The basic argument is that there are more languages than Turing machines and so there are languages than Turing machines. Thus some languages cannot be decidable

Baby steps

- What is counting?
  - Labeling with integers
  - Correspondence with integers
- Let us review basic properties of functions

Mappings and Functions

The function $F:A\to B$ maps one set A to another set B:



F is <u>one-to-one</u> (injective) if every $x\in A$ has a unique image $F(x)$: If $F(x)=F(y)$ then $x=y$.

F is <u>onto</u> (surjective) if every $z\in B$ is 'hit' by F: If $z\in B$ then there is an $x\in A$ such that $F(x)=z$.

F is a <u>correspondence</u> (bijection) between A and B if it is both one-to-one and onto.

Cardinality

A set S has k elements if and only if there exists a bijection between S and $\{1,2,\ldots,k\}$.

S and $\{1,\ldots,k\}$ have the same <u>cardinality</u>.

If there is a surjection possible from $\{1,\ldots,n\}$ to S, then $n \geq |S|$.

We can generalize this way of comparing the sizes of sets to infinite ones.

How Many Languages?

For $\Sigma=\{0,1\}$, there are 2^k words of length k.
Hence, there are $2^{(2^k)}$ languages $L \subseteq \Sigma^k$.

Proof: L has two options for every word $\in \Sigma^k$;
L can be represented by a string $\in \{0,1\}^{(2^k)}$.

That's a lot, but finite.

There are infinitely many languages $\subseteq \Sigma^*$.
But we can say more than that…

Georg Cantor defined a way of comparing infinities.

Countably Infinite Sets

A set S is <u>infinite</u> if there exists a surjective function $F:S\rightarrow\mathbb{N}$.
"The set $\mathbb{N}$ has no more elements than S."

A set S is <u>countable</u> if there exists a surjective function $F: \mathbb{N} \rightarrow S$
"The set S has not more elements than ☠."

A set S is <u>countably infinite</u> if there exists a bijective function $F: \mathbb{N} \rightarrow S$.
"The sets $\mathbb{N}$ and S are of equal size."

Counterintuitive facts

- Cardinality of even integers
  - Bijection i $\leftrightarrow$ 2i
  - A proper subset of N has the same cardinality as N !
  - Same holds for odd integers
- What about pairs of natural numbers?
  - Bijection from N to N x N !!
  - Cantor's idea: count by diagonals
  - Implies set of rational numbers is countable

Counterintuitive facts - 2

- Note that the ordering of Q is not in increasing order or decreasing order of value.
- In proofs, you CANNOT assume that an ordering has to be in increasing or decreasing order.
- So cannot use ideas like “between any two real numbers x, y, there exists a real number 0.5(x+y)” to prove uncountability.

More Countably Infinite Sets

One can make bijections between $\mathbb{N}$ and

1. $\{a\}^*$: $i \leftrightarrow a^i$
2. Integers (Z):

1	2	3	4	5	6	7	8	9	10	11
0	+1	-1	+2	-2	+3	-3	+4	-4	+5	-5

Countable sets in language theory

- Σ^* is countable – finitely many strings of length k. Order them lexicographically.
- Set of all Turing machines countable – every TM can be encoded as a string over some Σ.

Summary

A set S is <u>countably infinite</u> if there exists a bijection between {0,1,2,…} and S.

Intuitively: A set S is countable, if you can make a List (numbering) $s_1, s_2, \ldots$ of all the elements of S.

The sets Q, {0,1}* are countably infinite.

Example for {0,1}*: the lexicographical ordering:
{0,1}* = {ε,0,1,00,01,10,11,000,…}

Q: Are there bigger sets?

Next

- Chapter 4.2:
  - Uncountable Set of Languages
  - Unrecognizable Languages
  - Halting Problem is Undecidable
  - Non-Halting is not TM-Recognizable

Uncountable Sets

There are infinite sets that are not countable.
Typical examples are R, $\mathcal{P}$(N) and $\mathcal{P}$({0,1}*)

We prove this by a diagonalization argument.
In short, if S is countable, then you can make a list $s_1,s_2,\ldots$ of all elements of S.

Diagonalization shows that given such a list, there will always be an element x of S that does not occur in $s_1,s_2,\ldots$

Uncountability of $\mathcal{P}$(N)

The set $\mathcal{P}$(N) contains all the subsets of {1,2,…}.

Each subset $X \subseteq$ N can be identified by an infinite string of bits x_1x_2... such that x_j=1 iff $j \in X$.

There is a bijection between $\mathcal{P}$(N) and $\{0,1\}^N$.

Proof by contradiction: Assume $\mathcal{P}$(N) countable.

Hence there must exist a surjection F from N to the set of infinite bit strings.

“There is a list of *all* infinite bit strings.”

Diagonalization

Try to list all possible infinite bit strings:

0	0	0	0	0	0	⋯
1	1	1	1	1	1	⋯
2	1	0	0	0	0	⋯
3	0	1	0	1	0	⋯
⋮						⋱

Look at the bit string on the diagonal of this table: 0101… The negation of this string (“1010…”) does not appear in the table.

No Surjection $\mathbb{N} \to \{0,1\}^{\mathbb{N}}$

Let F be a function $N \to \{0,1\}^N$.
F(1),F(2),… are all infinite bit strings.

Define the infinite string $Y=Y_1Y_2\ldots$ by
Y_j = **NOT**(j-th bit of F(j))

On the one hand $Y \in \{0,1\}^N$, but on the other hand: for every $j \in N$ we know that $F(j) \neq Y$ because F(j) and Y differ in the j-th bit.

F cannot be a surjection: $\{0,1\}^N$ is uncountable.

Generalization

- We proved that $\mathcal{P}(\{0,1\}^*)$ is uncountably infinite.
- Can be generalized to $\mathcal{P}(\Sigma^*)$ for any finite Σ.

R is uncountable

- Similar diagonalization proof. We will prove [0,1) uncountable
- Let F be a function $N \to R$
  F(1),F(2),… are all infinite digit strings (padded with zeroes if required).
- Define the infinite string of digits $Y=Y_1Y_2$… by

$$Y_j = \begin{cases} \mathbf{F(i)_i + 1} & \textbf{if } \mathbf{F(i)_i < 8} \\ 7 & \textbf{if } \mathbf{F(i)_i \geq 8} \end{cases}$$

Q: Where does this proof fail on N?

Other infinities

- We proved 2^N uncountable. We can show that this set has the same cardinality as $\mathcal{P}$(N) and R.
- What if we take $\mathcal{P}$(R)?
- Can we build bigger and bigger infinities this way?
- Euler: Continuum hypothesis – YES!

Uncountability

We just showed that there it is impossible to have a surjection from N to the set $\{0,1\}^N$.

What does this have to do with Turing machine computability?

Counting TMs

Observation: Every TM has a finite description; there is only a countable number of different TMs. (A description $\langle M \rangle$ can consist of a finite string of bits, and the set $\{0,1\}^*$ is countable.)

Our definition of Turing recognizable languages is a mapping between the set of TMs $\{M_1, M_2, \ldots\}$ and the set of languages $\{L(M_1), L(M_2), \ldots\} \subseteq \mathcal{P}(\Sigma^*)$.

Question: How many languages are there?

Counting Languages

There are uncountably many different languages over the alphabet Σ={0,1} (the languages $L \subseteq \{0,1\}^*$)**.**
With the lexicographical ordering ε,0,1,00,01,… of Σ^*, every L coincides with an infinite bit string via its <u>characteristic sequence</u> χ_L.

Example for L={0,00,01,000,001,…} with χ_L= 0101100…

Σ^*	ε	0	1	00	01	10	11	000	001	010	⋯
L		X		X	X			X	X	X	⋯
χ_L	0	1	0	1	1	0	0	1	1	1	⋯

Counting TMs and Languages

There is a bijection between the set of languages over the alphabet $\Sigma=\{0,1\}$ and the uncountable set of infinite bit strings $\{0,1\}^N$.

- There are uncountable many different languages $L\subseteq\{0,1\}^*$.
- Hence there is no surjection possible from the countable set of TMs to the set of languages. Specifically, the mapping L(M) is not surjective.

Conclusion: There are languages that are not Turing-recognizable. (A lot of them.)

Is This Really Interesting?

We now know that there are languages that are not Turing recognizable, but we do not know what kind of languages are non-TM-recognizable.

Are there interesting languages for which we can prove that there is no Turing machine that recognizes it?