

Harvard CS 121 and CSCI E-121

Lecture 6:

Countability and Uncountability

Harry Lewis

September 19, 2013

Reading: Sipser, §1.3 and “The Diagonalization Method,” pages 174–178 (from just before Definition 4.12 until just before Corollary 4.18).

3 Characterizations Of Regular Language

A language is regular iff it is accepted by a DFA

iff it is accepted by an NFA

iff it is represented by a regular expression.

All these equivalences are CONSTRUCTIVE

- And we know that this class of languages is closed under union, concatenation, Kleene *, complement, intersection
- So we can mix and match our methods to our problems

Examples of Regular Languages

- $\{w \in \{a, b\}^* : |w| \text{ even \& every 3rd symbol is an } a\}$
- $\{w \in \{a, b\}^* : \text{There are not 7 } a\text{'s or 7 } b\text{'s in a row}\}$
- $\{w \in \{a, b\}^* : w \text{ has both an even number of } a\text{'s and an even number of } b\text{'s}\}$
- Are there non-regular languages???

Goal: Existence of Non-Regular Languages

Intuition:

- Every regular language can be described by a finite string (namely a regular expression).
- To specify an arbitrary language requires an infinite amount of information.

For example, an infinite sequence of bits would suffice: Σ^* has a lexicographic ordering, and the i 'th bit of an infinite sequence specifying a language would say whether or not the i 'th string is in the language.

⇒ Some language must not be regular.

How to formalize?

Cardinality

A set S is

- finite if there is a bijection $\{1, \dots, n\} \leftrightarrow S$ for some $n \geq 0$

In that case, we say $|S| = n$

($|S|$ is the size or cardinality of S)

\rightsquigarrow Is the empty set finite?

- infinite if it is not finite

So $\mathcal{N} = \{0, 1, 2, \dots\}$ is infinite

\rightsquigarrow What about $\{\mathcal{N}\}$?

Countability

A set S is

- countably infinite if there is a bijection $f : \mathcal{N} \leftrightarrow S$

This means that S can be “enumerated,” i.e. listed as $\{s_0, s_1, s_2, \dots\}$ where $s_i = f(i)$ for $i = 0, 1, 2, 3, \dots$

So \mathcal{N} itself is countably infinite

So is \mathcal{Z} (integers) since $\mathcal{Z} = \{0, -1, 1, -2, 2, \dots\}$

Q: What is f ?

$$f(i) = \left\{ \begin{array}{l} \frac{i}{2} \quad \text{if } i \text{ even} \\ \frac{-(i+1)}{2} \quad \text{if } i \text{ odd} \end{array} \right\}$$

- countable if S is finite or countably infinite
- uncountable if it is not countable

Facts about Infinite Sets

- **Proposition:** The union of 2 countably infinite sets is countably infinite.

$$\text{If } A = \{a_0, a_1, \dots\}, B = \{b_0, b_1, \dots\}$$

$$\text{Then } A \cup B = C = \{c_0, c_1, \dots\}$$

$$\text{where } c_i = \begin{cases} a_{i/2} & \text{if } i \text{ is even} \\ b_{(i-1)/2} & \text{if } i \text{ is odd} \end{cases}$$

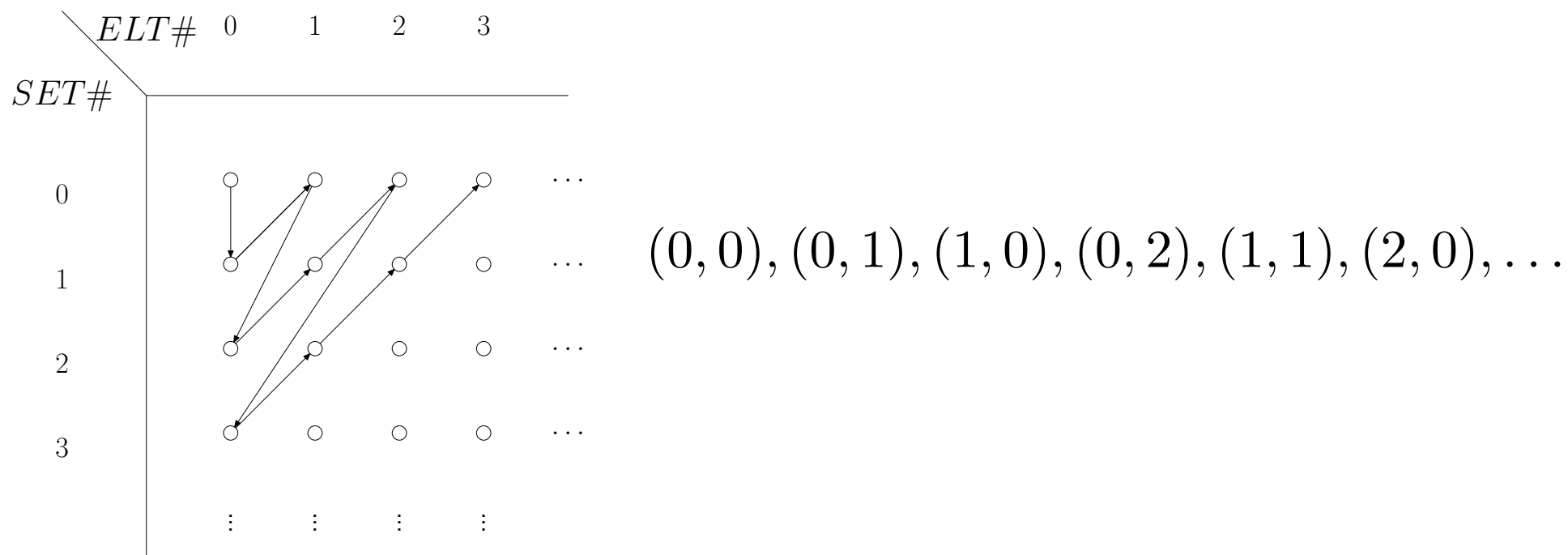
“Hilbert’s Grand Hotel Paradox”

Q: If we are being fussy, there is a small problem with this argument. What is it?

- **Proposition:** If there is an onto function $f : \mathcal{N} \rightarrow S$, then S is countable.

Countable Unions of Countable Sets

- **Proposition:** The union of countably many countably infinite sets is countably infinite



Each element is “reached” eventually in this ordering

Q: What is the bijection $\mathcal{N} \leftrightarrow \mathcal{N} \times \mathcal{N}$?

Are there uncountable sets? (Infinite but not countably infinite)

Theorem: $P(\mathcal{N})$ is uncountable
(The set of all sets of natural numbers)

Proof by contradiction:

(i.e. assume that $P(\mathcal{N})$ is countable and show that this results in a contradiction)

- Suppose that $P(\mathcal{N})$ were countable.
- Then there is an enumeration of all subsets of \mathcal{N} say $P(\mathcal{N}) = \{S_0, S_1, \dots\}$

Diagonalization

S_i	$j =$	0	1	2	3	4	
S_0	Y	N	N	Y	N	...	
S_1	N	N	N	N	N	...	
S_2	Y	Y	N	Y	Y	...	
S_3	N	N	N	Y	N	...	
⋮				D			

“Y” in row i , column j means $j \in S_i$

- Let $D = \{i \in \mathcal{N} : i \in S_i\}$ be the diagonal.
 - $D = YNNY \dots = \{0, 3, \dots\}$
 - Let $\bar{D} = \mathcal{N} - D$ be its complement.
 - $\bar{D} = NYYN \dots = \{1, 2, \dots\}$
 - **Claim:** \bar{D} is omitted from the enumeration, contradicting the assumption that every set of natural numbers is one of the S_i s.
- Pf:** \bar{D} is different from each row because they differ at the diagonal.

The Contradiction:

- Suppose, specifically, that \overline{D} is S_k
- Is $k \in S_k$?
- $k \in S_k$
if and only if
- $k \in D$ (since $D = \{i : i \in S_i\}$)
if and only if
- $k \notin \overline{D}$ (since $\overline{D} = \mathcal{N} - D$)
if and only if
- $k \notin S_k$ (since S_k is supposed to be \overline{D})

This is absurd so something must be wrong:

It can only be the assumption that $P(\mathcal{N})$ was countable. ✓

Cardinality of Languages

- An alphabet Σ is finite by definition
- **Proposition:** Σ^* is countably infinite
- So every language is either finite or countably infinite
- $P(\Sigma^*)$ is uncountable, being the set of subsets of a countable infinite set.
 - i.e. There are uncountably many languages over any alphabet
 - Q:** Even if $|\Sigma| = 1$?

Existence of Non-regular Languages

Theorem: For every alphabet Σ , there exists a non-regular language over Σ .

Proof:

- There are only countably many regular expressions over Σ .
 \Rightarrow There are only countably many regular languages over Σ .
- There are uncountably many languages over Σ .
- Thus at least one language must be non-regular.

\Rightarrow In fact, “almost all” languages must be non-regular.

Q: Could we do this proof using DFAs instead?

Q: Can we get our hands on an *explicit* non-regular language?