CS 373: Theory of Computation
Sariel Har-Peled and Madhusudan Parthasarathy

Discussion : Enumerators and Diagonalization

14 April 2009

Questions on homework 8?

Any questions? Complaints, etc?

1 Cardinality of a Set

For a finite set X, we denote by $|X|$ the ***cardinality*** of X; that is, the number of elements in A.

Definition 1.1 For two arbitrary sets (maybe infinite) X and Y, we have $|X| \leq |Y|$, iff there exists an injective mapping $f : X \to Y$.

Definition 1.2 Two arbitrary sets (maybe infinite) X and Y, are of the same ***cardinality*** (i.e., same "size") if $|A| = |B|$. Namely, there exists an *injective* and ***onto*** mapping $f : X \to Y$.

Observation 1.3 *For two sets X and Y, if $|X| \leq |Y|$ and $|Y| \leq |X|$ then $|X| = |Y|$.*

For $\mathbb{N}$, the set of all natural numbers, we define $|\mathbb{N}| = \aleph_0$. Any set X, with $|X| \leq \aleph_0$, is referred to as a ***countable*** set.

Claim 1.4 *For any set A, we have $|X| < |\mathbb{P}(X)|$. That is, $|X| \leq |\mathbb{P}(X)|$ and $|\mathbb{P}(X)| \neq |X|$. (Here $\mathbb{P}(X)$ is the power set of X.)*

Proof: It is easy to verify that $|X| \leq |\mathbb{P}(X)|$. Indeed, consider the mapping $h(x) = \{x\} \in \mathbb{P}(X)$, for all $x \in X$.

So, assume for the sake of contradiction that $|X| = |\mathbb{P}(X)|$, and let f be a one-to-one and onto mapping from X onto $\mathbb{P}(X)$. Next, consider the set $B = \left\{ x \in X \;\middle|\; x \notin f(x) \right\}$.

Now, consider element $b = f^{-1}(B)$, and consider the question of whether it is a member of the set B or not. Now if $f^{-1}(B) = b \in B$, then by the definition of B, we hvae $bf^{-1}(B) = b \notin B$. Similarly, if $f^{-1}(B) = b \notin B$, then by definition of B, we have $f^{-1}(B) = b \in B$.

A contradiction. We conclude that our assumption that f exists (since X and $\mathbb{P}(X)$ have the same cardinality) is false. We conclude that $|X| \neq |\mathbb{P}(X)|$. ■

Definition 1.5 An ***enumerator*** `T` for a language L is a Turing Machine that writes out a list of all strings in L. It has no input tape, only an output tape on which it prints the strings, with some separator character (say, #) printed between them.

The strings can be printed in any order and the enumerator is allowed to print duplicates of a string it already printed. However, sooner or later all strings in L must be printed eventually by T. Naturally, all the strings printed by T are in L.

2 Rationals are enumerable

Consider the set of rational numbers

$$\mathbb{Q} = \left\{ a/b \;\middle|\; a \in \mathbb{Z}, b \in \mathbb{N}, b \neq 0, \text{ and } a, b \text{ are relatively prime} \right\}.$$

We remind the reader that two natural numbers a and B are ***relatively prime*** (or ***coprime***) if they have no common factor other than 1 or, equivalently, if their greatest common divisor is 1. Thus 2 and 3 are coprime, but 4 and 6 are not coprime. Thus, although $2/3 = 4/6$, we will consider only the representation $2/3$ to be in the set $\mathbb{Q}$.

We show that this set is enumerable by giving the pseudo-code for an enumerator for it.

```
EnumerateRationals
  for i = 1 ... ∞ do
      for x = 0 ... i − 1 do
          y = i − x
          if x, y are relatively prime then
              print x/y onto the tape followed by #
              print −x/y onto the tape followed by #.
```

It is obvious that every rational number will be enumerated at some point. Any rational number is of the form a/b and as such when $i = a + b$ an d $y = b$ it will enumerate this rational number.

It helps to picture this as travelling along each line $x + y = i$.

3 Counting all words

Consider a finite alphabet Σ, and consider the problem of enumerating all words in Σ^*. That is, we want to come up with a way to be able to compute the ith word in Σ^* (let denote it by w_i). We want to do it in such a way that given a word w we can compute the i such that $w_i = w$, and similarly, given i we can compute w_i.

To this end, let Σ_i be all the words in Σ^* of length exactly i. Clearly, $|\Sigma_i| = |\Sigma|^i$. We sort the words inside Σ_i lexicographically. As such, we can now list all the words in Σ^**, by first listing the words in $\Sigma_0 = \{\epsilon\}$, $\Sigma_1 = \Sigma$, and so on.

For example, for $\Sigma = \{\mathtt{a}, \mathtt{b}\}$, we get the following enumeration of the words of Σ^*:

$$\overbrace{\underbrace{\epsilon}_{=w_1}}^{\Sigma_0}, \overbrace{\underbrace{\mathtt{a}}_{=w_2}, \underbrace{\mathtt{b}}_{=w_3}}^{\Sigma_1}, \overbrace{\underbrace{\mathtt{aa}}_{=w_4}, \underbrace{\mathtt{ab}}_{=w_5}, \underbrace{\mathtt{ba}}_{=w_6}, \underbrace{\mathtt{bb}}_{=w_7}}^{\Sigma_2}, \overbrace{\underbrace{\mathtt{aaa}}_{=w_8}, \underbrace{\mathtt{aab}}_{=w_9}, \underbrace{\mathtt{aba}}_{w_{10}}, \underbrace{\mathtt{abb}}_{w_{11}}, \underbrace{\mathtt{baa}}_{w_{12}}, \underbrace{\mathtt{bab}}_{w_{13}}, \underbrace{\mathtt{bba}}_{w_{14}}, \underbrace{\mathtt{bbb}}_{w_{15}}}^{\Sigma_3}, \ldots$$

It is now easy to verify that given $w \in \Sigma^*$, we can figure out the i such that $w_i = w$. Similarly, given i, we can output the word w_i.

We just demonstrated that there is a one-to-one and onto mapping from $\mathbb{N}$ to Σ^*, and we can conclude the following.

Lemma 3.1 *For a finite alphabet Σ, the set Σ^* is countable.*

4 Languages are not countable

Let

$$\mathsf{L}_{\text{all}} = \left\{ L \;\middle|\; L \text{ is some language, and } L \subset \{\mathtt{a}, \mathtt{b}\}^* \right\}.$$

Claim 4.1 *The set L_{all} is not countable.*

Proof: We show that this set is not countable by using a diagonalization argument. Assume for the sake of contradiction that L_{all} is countable. Then there exists a one-to-one and onto mapping $g : \mathbb{N} \to \mathsf{L}_{\text{all}}$. Let $L_i = g(i)$, for all i.

We can think about this mapping as follows. We create an infinite table where the ith row is the language of L_i. We also enumerate the columns, where the ith column is the ith word in Σ^* (use the above enumeration scheme. We write 1 in the ith row and jth column of this table if w_j is in the language L_i.

Consider the diagonal language of this table:

	w_1	w_2	w_3	w_4	$\ldots$
L_1	**1**	1	0	0	$\ldots$
L_2	0	**1**	0	1	$\ldots$
L_3	1	0	**1**	1	$\ldots$
L_4	0	1	0	**0**	$\ldots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Formally, $L_d = \left\{ w_i \;\middle|\; w_i \in L_i, i \geq 0 \right\}$.

Let $\overline{L_{\text{diag}}}$ be the complement of L_d. Clearly, $\overline{L_{\text{diag}}}$ is well defined and $\overline{L_{\text{diag}}} \in \mathsf{L}_{\text{all}}$. But then, there must exist a k such that $L_k = g(k) = \overline{L_{\text{diag}}}$.

So consider the kth row in this table (i.e., this is the row that corresponds to the language L_k). We consider if the word $w_k \in L_k$. If $w_k \in L_k = \overline{L_{\text{diag}}}$ then the entry (k, k) in the table must be 1. But in that case $w_k \in L_d$, and as such $w_k \notin \overline{L_{\text{diag}}}$. This is impossible.

The other possibility is that $w_k \notin L_k = \overline{L_{\text{diag}}}$, then the entry at position (k, k) in the table must be 0. But in that case, the $w_k \notin L_d$, but then such $w_k \in \overline{L_{\text{diag}}}$, which is again impossible.

A contradiction. We conclude that our assumption that L_{all} is countable is false. ∎