Theory of Computation Notes: Countablility and Diagonalization

Arjun Chandrasekhar

The set of nautral numbers is clearly a subset of the set of integers, which is clearly a subset of the rational numbers, which is clearly a subset of the real numbers. But all of these sets are infinite. Are they all the same size? Are abt of then 'bigger' than the other? What does such a statement even mean, and how do we even prove such a statement?

1 Bijections

Definition 1.1. Let S_1 and S_2 be two sets. A **bijection** is a mapping $f : S_1 \to S_2$ that satisfies the following two properties:

1. **Surjective:** every element in S_2 is mapped to by some element in S_1
2. **Injective:** The mapping is one-to-one; every element in S_1 maps to *exactly* one element of S_2

Example 1.1. Let $\mathbb{N} = \{0, 1, 2, \dots\}$ be the natural numbers, and let $S = \{0, 1, 4, 9, \dots\}$ be the square integers. Then there exists a bijection between $\mathbb{N}$ and S. We simly have $f(n) \to n^s$, i.e.

- $0 \to 0$
- $1 \to 1$
- $2 \to 4$
- $3 \to 9$
- $4 \to 16$
- $\dots$

Every square integer is mapped to by its square root, and every natural number maps to exactly one square.

2 Countably Infinite Sets

Axiom. The natural numbers $\mathbb{N}$ are countable

Definition 2.1. A set S is **countably infinite** if there exists a bijection $f : \mathbb{N} \to S$

Remark. If it is obvious that a set is infinite then we will usually just say the set is **countable**.

Example 2.1. The set of square numbers $S = \{0, 1, 4, 9, \dots\}$ is countably infinite

Example 2.2. The set of integers $\mathbb{Z} = \{\dots, -3, -2, -10, 1, 2, 3, \dots\}$ is countable. To show this we give the following bijection:

- $0 \to 0$
- $1 \to 1, 2 \to -1$

- $3 \to 2, 4 \to -2$
- $5 \to 3, 6 \to -3$
- ...

Theorem 2.1. The set of rational numbers $\mathbb{Q} = \{\,{}^{a}/_{b} | \, a, b \in \mathbb{N}, b \neq 0\}$ is countable.

Proof. A naive approach would be to list all of the rationals with 1 on the denominator, then list all of the rationals with 2 on the denominator, etc. But this will use up all of our resources on the rationals with 1 on top. Instead, we use the trick of going diagonally: for every integer $n = 1, 2, 3, \ldots$ we list all rationals whose numerator and denominator add up to k.

- $0 \to 0$
- $n = 2$
  - $1 \to {}^{1}/_{1}$
- $n = 3$
  - $2 \to {}^{1}/_{2}$
  - $3 \to {}^{2}/_{1}$
- $n = 4$
  - $4 \to {}^{3}/_{1}$
  - We skip over ${}^{2}/_{2}$ because we already got 1
  - $5 \to {}^{1}/_{3}$
- $n = 5$
  - $6 \to {}^{4}/_{1}$
  - $7 \to {}^{3}/_{2}$
  - $8 \to {}^{2}/_{3}$
  - $9 \to {}^{1}/_{4}$
- ...

Figure 1 illustrates this technique. □

Proposition 2.1. The set of Turing machines on the alphabet $\Sigma = \{0, 1\}$ is countable.

Proof. We list out the set of TMs as follows:

1. List out all TMs with 1 state
2. List out all TMs with 2 states
3. List out all TMs with 3 states
4. List out all TMs with 4 states
5. ...

□

2.1 Exercises

1. Prove that the set $\mathbb{N}^2 = \{(x, y)|x, y \in \mathbb{N}\}$ is countable.

2. Prove that the set of all <u>finite</u> binary strings is countable

2. Prove that the set of all java programs is countable.

3 Uncountable Sets

Intuitively, it seems unlikely that the real numbers would be countable. If you were to try to list out all of the real numbers, where would you start? How would you know which number to list 'next'? On the other hand, how do we rigorously prove that no such bijection exists? The solution involves using a brilliant technique developed by the mathematician Georg Kantor.

Theorem 3.1 (Cantor)**.** The set of real numbers $\mathbb{R}$ is uncountable.

Proof. To prove this we use Cantor's technique of <u>**diagonalization**</u>. The idea is as follows:

1. AFSOC there is a bijection $f : \mathbb{N} \to \mathbb{R}$
2. We construct a real number r^* that disagrees with every other real number at one digit
3. Either r^* disagree. Either way we have reached a contradiction.

Formally, we prove it as follows:

1. AFSOC there is a bijection $f : \mathbb{N} \to \mathbb{R}$

$$0 \to r_0, 1 \to r_1, 2 \to r_2, 3 \to r_3, \ldots$$

2. We construct a real number r^* as follows: For each r_i, we make it so that the i-th digit of r^* is different from the i-th digit of r_i. This way, r^* must be distinct from each r_i
3. r^* is a valid real number, so it must be part of the bijection. This means $r^* = r_i$ for some $i \in \mathbb{N}$. But we constructed r^* such that it does not match each r_i. Thus r^* disagrees with itself at some digit, which is a contradiction.

Figure 2 illustrates Cantor's diagonalization technique.

□

Lemma 3.1. The set of <u>infinite</u> binary strings is uncountable.

Proof. We once again make use of Cantor's diagonalization argument.

1. AFSOC there exists a bijection between $\mathbb{N}$ and the set of <u>infinite</u> binary strings.

$$0 \to s_0, 1 \to s_1, 2 \to s_2, \ldots$$

2. We construct a new string s^* as follows: for each s_i, we make it so that s^* and s_i disagree at the i-th bit
3. s^* is an infinite binary string and it must be part of our bijection. This means $s^* = s_i$ for some $i \in \mathbb{N}$. But we constructed * uch that it does not match any s_i. Thus s^* disagrees with itself at some bit, which is a contradiction.

□

Lemma 3.2. The set of formal languages (on any alphabet) is uncountable.

Proof. Let $w_1, w_2, w_3, \cdots \in \Sigma^*$ be the set of all possible strings. We can represent a language using an infinite binary string: we set bit i to be 1 if $w_i \in L$ and 0 otherwise.

The set of infinite binary strings is uncountable, therefore the set of formal langauges must also be uncountable. □

Corollary 3.1. There exist languages that are not Turing-recognizable.

Proof. The set of Turing machines is countable, but the set of formal languages is uncountable. Thus there cannot possibly be a Turing machine to recognize every language. □

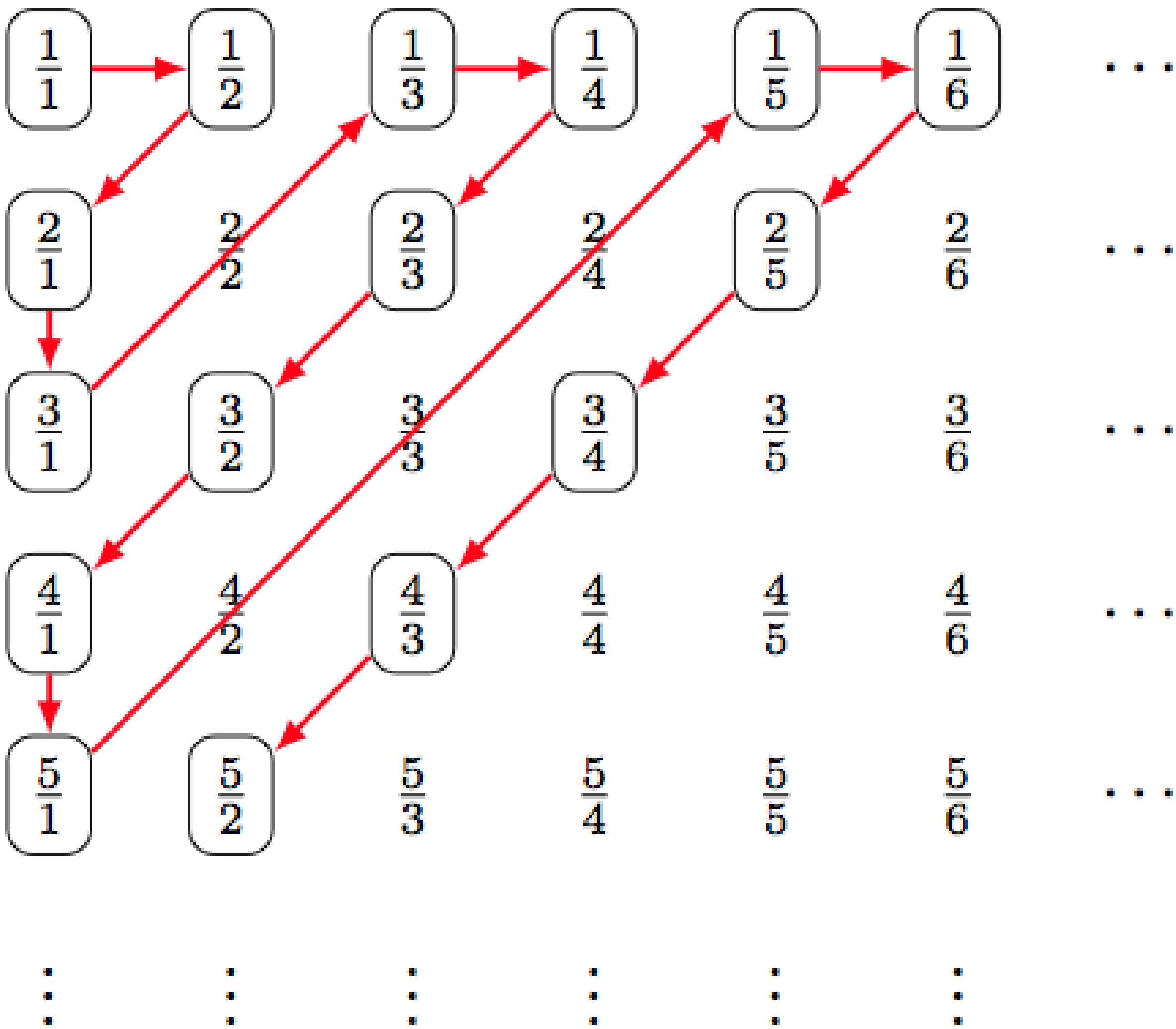

Figure 1: **Counting the rational numbers**

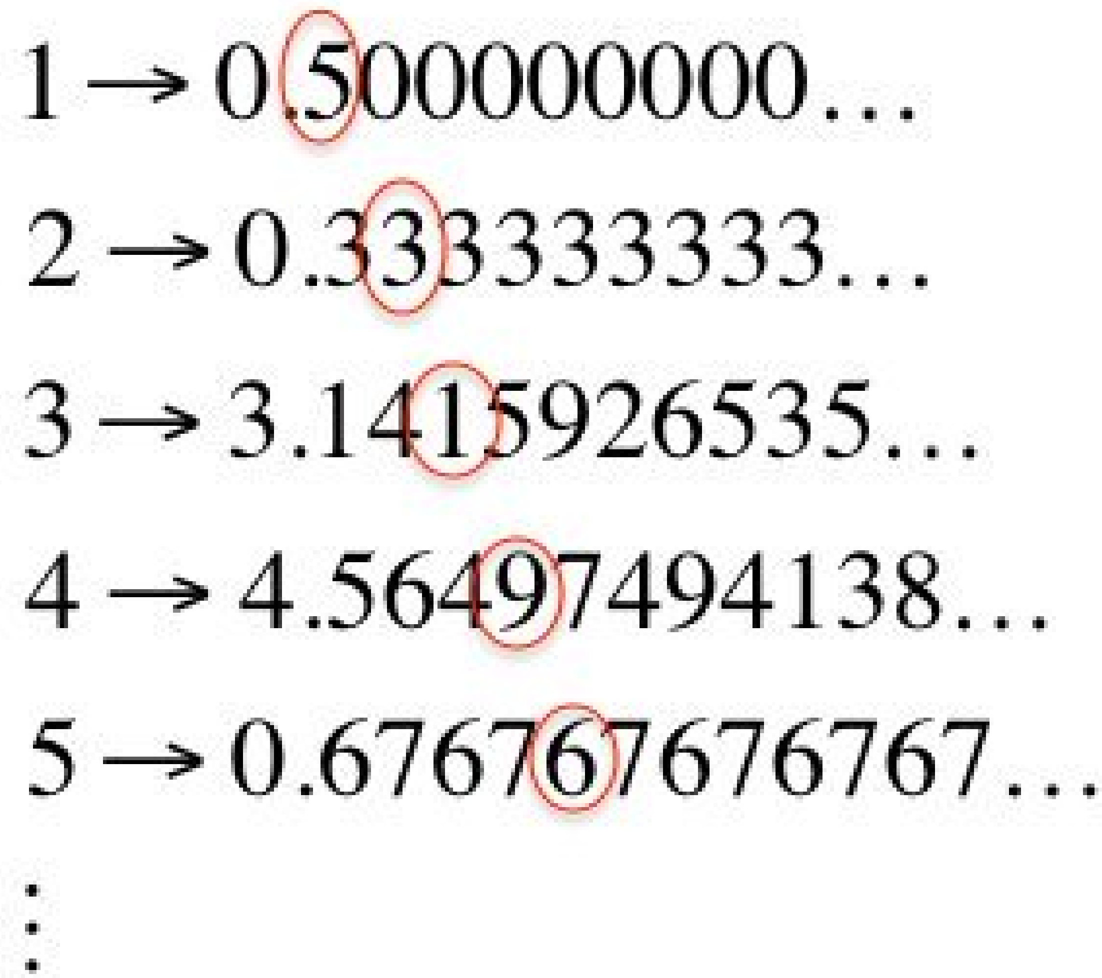

Figure 2: **Cantor's diagonalization techinque:** We first assume we can list out the reals in order, and then construct a new real number r^*. For each real number r_i we make sure r^* and r_i disagree at the i-th digit.

$$
\begin{array}{lcl}
s_1 & = & 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ \ldots \\
s_2 & = & 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ \ldots \\
s_3 & = & 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ \ldots \\
s_4 & = & 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ \ldots \\
s_5 & = & 1\ 1\ 0\ 1\ 0\ 1\ 1\ 0\ 1\ 0\ 1\ \ldots \\
s_6 & = & 0\ 0\ 1\ 1\ 0\ 1\ 1\ 0\ 1\ 1\ 0\ \ldots \\
s_7 & = & 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ \ldots \\
s_8 & = & 0\ 0\ 1\ 1\ 0\ 0\ 1\ 1\ 0\ 0\ 1\ \ldots \\
s_9 & = & 1\ 1\ 0\ 0\ 1\ 1\ 0\ 0\ 1\ 1\ 0\ \ldots \\
s_{10} & = & 1\ 1\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 0\ 1\ \ldots \\
s_{11} & = & 1\ 1\ 0\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ \ldots \\
\vdots & & \vdots\ \vdots\ \vdots\ \vdots\ \vdots\ \vdots\ \vdots\ \vdots\ \vdots\ \vdots\ \vdots\ \ddots
\end{array}
$$

$$
s = 1\ 0\ 1\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 1\ \ldots
$$

Figure 3: **Infinite binary strings are uncountable:** Here we illustrate how to apply Cantor's diagonalization technique to the set of infinite binary strings.