Cardinalities related to formal languages

Arno Pauly

March 9, 2021

Cardinality recap

Definition
We write $|X| \leq |Y|$ if there is a surjection from Y to X, and $|X| = |Y|$ if there is a bijection from X to Y.

$|X|$ is called the *cardinality* of the set X.

- For any two sets, $|X| \leq |Y|$ or $|Y| \leq |X|$.
- If $|X| \leq |Y|$ and $|Y| \leq |X|$, then $|X| = |Y|$.
- If $|X| \leq |Y|$ and $|Y| \leq |Z|$, then $|X| \leq |Z|$,

We usually write $|X| = n$ rather than $|X| = |\mathbf{n}|$.

More terminology

- We call X *countable* if there is a surjection $s : \mathbb{N} \to X$, (Otherwise it is called *uncountable*.)
- A set X is *infinite* if there is an injection $\iota : \mathbb{N} \to X$.
- The cardinality of $\mathbb{N}$ is called $\aleph_0$ (read $\aleph_0$).

The continuums hypothesis (detour)

With $\aleph_1$ we denote the second-smallest infinite cardinality.

Theorem (Gödel and Cohen)

The usual axioms for mathematics neither prove that $|\mathbb{R}| = \aleph_1$ *nor do they prove that* $|\mathbb{R}| \neq \aleph_1$.

So it is impossible to tell whether there is a set having more elements than the natural numbers, but fewer than the reals (in terms of cardinality).

How big are formal languages? How many are there?

Proposition

For a non-empty finite alphabet Σ *it holds that* $|\Sigma^*| = |\mathbb{N}|$.

Corollary

Every formal language is countable.

Corollary

For a non-empty finite alphabet Σ *it holds that*
$|\mathcal{P}(\Sigma^*)| = |\mathcal{P}(\mathbb{N})| > |\mathbb{N}|$.

Read: There are uncountably many formal languages.

How many grammars are there?

Proposition

For a finite alphabet Σ, there are countably many grammars (up to renaming of non-terminal symbols).

Corollary

There are (way) more formal languages than there are grammars.

Corollary

There are formal languages not describable by any grammar.