Oberlin College
CSCI 383
Fall 2026
Lecture 1: Strings and Languages
Kozen - Sections 1 and 2

Computational problems.

- Test primality of n.
- Test if a graph is acyclic.
- Compute a graph's MST.

What datatype are the outputs to these problems?

To standardize: decision problems.

Def: A decision problem is a problem where the answer is True or False.

- First two are already decision problems.
- One version of MST: Given $G = (V, E)$ and integer k, is there an MST of cost $\leq k$?

Observation: This reduces computation to testing set membership. (Is input in the set of "Yes" instances?)

- Recognize all numbers n such that n is prime.
- Recognize all acyclic graphs
- Recognize all pairs (G, k) such that G has MST of cost $\leq k$.

Strings

Idea: Standardized data type for inputs.

- First problem: digits.
  - Binary
  - Decimal
  - Etc.
- Second and third problem: graph.

Question: How could you represent a graph as a string?

Def. An *alphabet* Σ is a finite set of symbols.

- $\{0, 1\}$
- $\{1\}$
- $\{a, b, c, \ldots, z\}$
- $\{\text{square}, \text{circle}, \text{triangle}\}$

Def. A *string* is a finite-length sequence of elements of Σ.

What's your favorite string over $\{a, b, c, \ldots, z\}$?

Mine is the empty string. ε.

Nice things we can do with strings:

- Take their length $|x|$ (What is $|100100|$?)
- Concatenate them xy
- Power them $x^k = x \cdots x$ (k times)

Questions:

- Let $x = 100100$. What is $|x^4|$?
- What should $|\varepsilon|$ be?
- What should x^0 be?

Languages

Idea: Collections of "Yes" instances.

Def: A *language* is a (possibly infinite) set of strings.

- Set of palindromes.
- Set of all binary encodings of prime numbers.
- Set of all 4-touples G, s, t, k such that the shortest $s - t$ path in G has length $\leq k$.

Solving a computational problem amounts to just testing language membership.

Fun things you can do with sets.

- Union $A \cup B = \{x \mid x \in A \text{ or } x \in B\}$
- Intersection $A \cap B = \{x \mid x \in A \text{ and } x \in B\}$
- Complement $\sim A = \{x \mid x \notin A\}$

Complement is alphabet dependent! What is $\sim \{\epsilon, 0, 00, 000, \ldots\}$ when:

- $\Sigma = \{0, 1\}$?
- $\Sigma = \{0\}$?

Funner things you can do with sets.

Concatenate! $AB = \{xy \mid x \in A \text{ and } y \in B\}$

Ex: $\{a, ab\}\{b, ba\} = \{ab, aba, abb, abba\}$.

True or false: $AB = BA$?

Power! A^n. "n copies of A, concatenated"

Recursive definition:

- $A^0 = \{\varepsilon\}$
- $A^{n+1} = AA^n$

Ex: Given alphabet Σ, Σ^n is strings of length n.

Kleene Star: A^* "A concatenated with itself *some* number of times." (The book calls this "asterate.")

Alternatively: $A^* = \{x \mid x \in A^k \text{ for some } k \geq 0\}$.

Formally: $A^* = \bigcup_{k=0}^{\infty} A^k$.

Ex: Given alphabet Σ, Σ^* is all strings.

Exercise: Which of the following are true?

- $A^*A^* \subseteq A^*$
- $A^*A^* \supseteq A^*$

Why?