CS 360

Bartosz Antczak *Instructor: Shalev Ben-David*

- **Alphabet:** a set of characters, using notation $\Sigma = \{a, b, c\}$
- **String:** a finite sequence of characters. Using the alphabet above, a string over Σ can be $aabca$. The empty string ϵ is valid over any language. Things to note:
  - $\{\epsilon\} \neq \emptyset$
  - $|\{\epsilon\}| = 1$
  - $|\epsilon| = 0$
- **Language:** a set of strings over Σ. Languages don't have to be finite.

Concatenation

If A and B are sets of strings, then AB (A concat B) is defined as

$$AB = \{st : s \in A, t \in B\} \qquad |AB| \leq |A||B|$$

When we concatenate a string k, x times, we append k to ϵ x times. For instance, concatenating ab 1 time will yield ab; concatenating it 0 times will yield ϵ.

Concatenation Examples

- $\emptyset\{00, 01\} = \emptyset$, since $|AB| \leq |A||B|$
- $\{\epsilon\}\{00, 01\} = \{00, 01\}$

Star Notation

If A is a set of strings, then A^* is the set defined as

$$\{\epsilon\} \cup A \cup AA \cup AAA \cup ...$$

In other words, A^* is the set of all strings you can form over A. The size of A^* is infinite, but the length of every string in A^* is finite:

- $|\{a, b, c\}^*| = \infty$
- $\forall s \in \{a, b, c\}^*$, $|s| \in \mathbb{N}$

Regular Expressions

Can be used to denote a language. For example, the regex for a language over $\Sigma = \{0, 1\}$ containing an even number of 1's is:

$$(\{1\}\{0\}^*\{1\} \cup \{0\})^*$$

Riddle: write a regex for a language that contains only palindromes, such as 101, 1001001, etc.

Automata

A DFA also determines a language, meaning that DFAs and regexs are connected.