CSCE 355
1/13/2025

Course homepage: ~
https://cse.sc.edu/~fenner/csce355
link to the syllabus

3 sections:

1. Automata theory — regular languages
2. Context-free languages & grammars
3. Universal computation: Turing machines, decidability/undecidability

Automata & regular languages

Def: An alphabet is any nonempty finite set.
Elements of an alphabet are called symbols (or letters or characters).

Ex: $\Sigma = \{a, b, c\}$, $\Sigma =$ ASCII

$\Sigma = \{0, 1\}$ (binary alphabet)

$\Sigma = \{0, \ldots, n-1\}$ for $n \geq 1$, n-ary alphabet

$\Sigma = \{0\}$ — unary alphabet

Def: Fix an alphabet Σ. A string over Σ is any finite sequence of symbols from Σ. (0 or more)

Ex: $\Sigma = \{a, b, c\}$

$\neq \{$ aa, acbca, cacba, ε $\}$

$|aa| = 2$

$|acbca| = 5$

$|\varepsilon| = 0$

If x is a string, then $|x|$ is the length of x.

ε (epsilon) is a metasymbol that denotes the empty string (unique string of length 0 (over any alphabet))

ε will never be a member of any alphabet.

Def: Fix alphabet Σ and let x, y be strings over Σ. The concatenation of x followed by y, written xy is the sequence of symbols from x, followed by those of y. This is a string.

Ex: $x = ab$, $y = cca$

$xy = abcca$

$yx = ccaab$

concat is not commutative i.e., order matters

Concat is associative: $\forall x, y, z$ strings,

$$x(yz) = (xy)z = xyz$$

True for more than 3 strings.

Concat has an identity element: $\forall$ string x,

$$\varepsilon x = x\varepsilon = x$$

ε is the identity element for concatenation.

Concat is length-additive: $|xy| = |x| + |y|$.

Some shorthand: let x be a string.

for $n \geq 0$, $x^n = \underbrace{xx\cdots x}_{n \text{ times}}$

Ex: $x^0 = \varepsilon$

$x^1 = x$

$x^2 = xx$

Notation: For alphabet Σ, let Σ^* denote the set of all strings over Σ.

String induction (induction on string length)

Good for proving facts for all strings & for defining functions on strings.

Ex: Length function:

1. $|\varepsilon| = 0$ in Σ^*
2. For any string $y \in \Sigma^*$ and symbol $a \in \Sigma$,
$|ya| = |y| + 1$.

This uniquely determines the length function on all strings.

[Can identify a symbol with a string of length 1]

Principle of String Induction:

For every string $x \in \Sigma^*$, exactly one of the following holds:

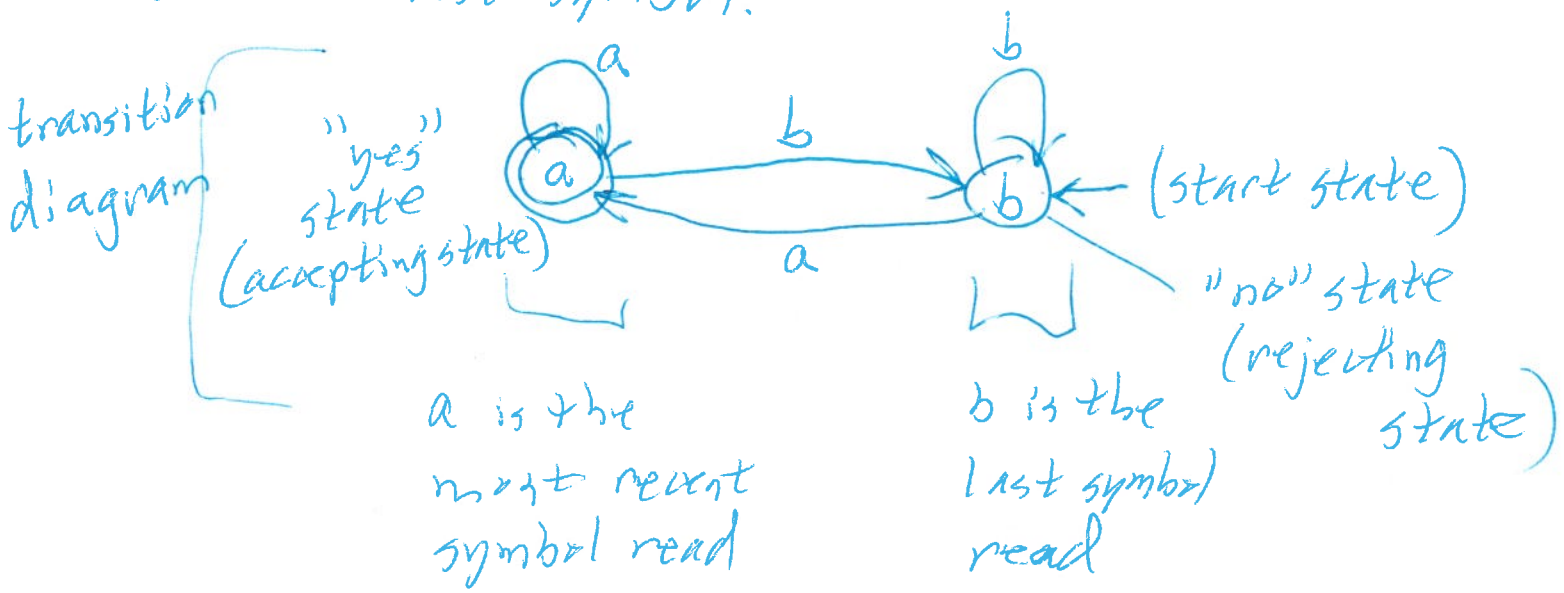
1. $x = \varepsilon$ $(|x| = 0)$
2. There exists a unique string $y \in \Sigma^*$ and a unique symbol $a \in \Sigma$ such that
$x = ya$, and furthermore $|y| = |x| - 1$.

(Here a is the last symbol of x and y is the principal prefix of x.)

This principle allows induction on string length.

Ex: $\Sigma = \{a, b\}$

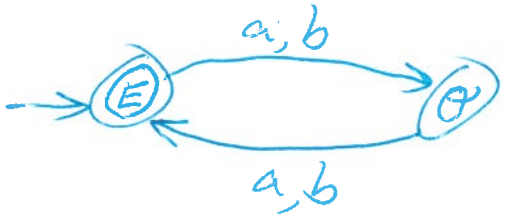
Question: Given a string $x \in \Sigma^*$, does it have a as a last symbol?



Sample input: aaba

b a a b a

Ex: Same Σ. Given $x \in \Sigma^*$, is $|x|$ even?



E - even
O - odd

Ex: Given $x \in \Sigma^*$, is a the first symbol of x?