CS310

Strings, String Operators, and Languages

Sections:

August 27, 2008
Quick Review

• Sets (Union, Intersection, [Proper] Subset)
  \{ n | rule about n\}
  Cross Product/Power Set

• Sequences/Tuples

• Functions
  \( f : D \rightarrow R \)

• Relation
  \( f : A_1 \times A_2 \times \ldots \times A_n \rightarrow \{\text{TRUE, FALSE}\} \)
  Equivalence Relations: 3 conditions
Strings

• Alphabet: Any finite set, $\Sigma = \{a, b\}$
• String: Any finite sequence of symbols from a given alphabet
  $w = ababaabba$, string over $\Sigma$
  $\epsilon = \text{empty string, zero symbols}$
  length of $w$: $|w| = \text{number of symbols it contains}$
  $|\epsilon| = |w| =$
• Strings are building blocks of computer science
  strings can represent: data sets (DNA), source code, files...
String Operations

• Closure ($\Sigma^*$): set of all strings over $\Sigma$, including $\varepsilon$.

$\Sigma = \{a, b\}$  $\Sigma^* = \{\varepsilon, a, b, ab, ba, aa, bb, \ldots\}$

• Concatenations

If $x, y \in \Sigma^*$, then $xy$ is defined to be the concatenation of strings $x, y$

$x = aba$  $y = bab$  $xy =$  $x^k$ is $k$ copies of $x$ concatenated

$x^2 =$
String Operations

• Prefix/Suffix
  \[ z = xy \] for \( x, y, z \in \sum^* \), \( x \) is a prefix of \( z \)
  \( y \) is a suffix of \( z \)

• Reverse
  \( x \in \sum^* \), \( x^R \) is the reverse of \( x \)
  \( x = ab \), \( x^R = ba \)
Languages

• Language
  Language L over $\Sigma$ is a subset of $\Sigma^*$
  
  $L = \{ x \in \{a,b\}^* \mid |x| \text{ is even} \}$
  
  $= \{\varepsilon, aa, ab, \ldots \}$

• Complement of a language L over $\Sigma$
  $\Sigma^* - L = L'$

• Concatenation of languages
  $L_1$ and $L_2$ over $\Sigma$
  
  $L_1L_2 = \{xy \mid x \in L_1, y \in L_2\}$
  
  $L^2 = LL$
Languages

• Union of languages

$L_1$ and $L_2$ over $\Sigma$

$L_1 \cup L_2 = \{x \mid x \in L_1 \text{ or } x \in L_2\}$

$L_1 = \{0\}^*$

$L_2 = \{1\}^*$

what is in $L_1 \cup L_2$?
what is in $L_1L_2$?
Languages

• Kleene Star

\[ L^* = \text{set of strings formed by concatenating any number of strings from } L \]
\[ L = \{ x \in \{ a, b \}^* \mid |x| \text{ is odd} \} \]

What does L contain:

\{ \}

\[ L^* = \{ \epsilon, \quad , \quad , \quad , \quad , \quad , \} \]
Languages

• Recursive Definitions
  Define $L$ over $\Sigma = \{0, 1\}$ as
  1. $\varepsilon \in L$
  2. If $x \in L$ then $0x1 \in L$
  What is in $L$? $L = \{\}$

• Can we prove that $\{\varepsilon, 01, 0011, 000111, \ldots\}$ is equivalent to $\{0^i1^i \mid i \geq 0\}$?

• Show $L$ is subset of $\{0^i1^i \mid i \geq 0\}$ and the reverse
Proof

• For $x, y \in \Sigma^*$, show $(xy)^R = y^Rx^R$