Lecture 5

Homework #5: 1.7.4a, 1.7.4b, 1.7.5b,c, 1.7.6, 1.8.2a-d, 1.8.3a,b, 1.8.5

Recall from last time that:
the set of all strings over an alphabet $\Sigma$ is denoted $\Sigma^*$.

A language is a set of strings over an alphabet.
  ex. $\Sigma^*$, $\emptyset$, $\Sigma$ are languages

languages are sets - and can be manipulated/combined in the usual (and not quite so usual) ways:

1) $L_1 \cup L_2$

2) $\overline{A}$ = the complement of $A = \Sigma^* - A$.

3) $L_1L_2$ = the concatenation of languages. \{xy : x $\in$ $L_1$ and y $\in$ $L_2$\}
$L^*$ = Kleene star (e is always in $L^*$)
$L^+ = LL^*$ = closure of a language L under concatenation

Q: Is there a good/clear way to represent these possibly infinite languages?

Can certainly use the set notation just discussed:

Example.
\{
  a,b\}^*a\{a\}b+\{a\}\{a,b\}\{a\}\{b\}^*

is there a simpler way of stating this?
(a $\cup$ b)*ab+a(a $\cup$ b)ab*

This is an example of a regular expression

(with which you’re probably familiar if you’ve ever used grep, awk, perl, etc.)

Def. Let $\Sigma$ be an alphabet. The regular expressions over $\Sigma$ and the sets they denote are defined by the following:
1) \( \emptyset \) is a reg expr and denotes the empty set.

2) \( e \) is a reg expr and denotes the set \( \{e\} \)

3) for \( a \in \Sigma \), \( a \) is a reg expr and denotes \( \{a\} \)

4) if \( r \) and \( s \) are reg expr denoting the sets \( R \) and \( S \), then \( (r \cup s), (rs), \) and \( (r^*) \) denote \( R \cup S, R^S, R^* \).

NOTE that the above def is not exactly the definition that's in the text.

Now let's do the reverse: let's find the language (i.e., the set) that is represented by a regular expression.

if \( r \) is a reg expr, \( L(r) \) is the lang denoted by \( r \).

ex. \( (((a^*a)b) \cup b) = L((a^*a)b) \cup L(b) = L(a^*a)L(b) \cup L(b) = L(a^*)L(a)L(b) \cup L(b) = L(a)^*L(a)L(b) \cup L(b) = \{a\}^*\{a\}\{b\} \cup \{b\} \)

A language is a **regular language** iff it can be expressed by a regular expression.

Let’s now do some class exercises to be sure that everyone is comfortable with regular expressions and the languages they express:

I. Let \( \Sigma = \{0, 1\} \). Give the regular expression for each of the following.

\{w | w has exactly a single 1\}
\{w | w has at least one 1\}
\{w | w contains 001 as a substring\}
\{w | w is a string of even length\}
\{w | the length of w is a multiple of 3\}
\{01, 10\}

II. What is the language described by each of the following regular expressions?
(0 ∪ e)1*
∅*
1*∅
(0 ∪ 1)*00(0 ∪ 1)*
(1 ∪ 10)*

III. Give a regular expression for the set of positive odd integers represented in binary. Then give one for the set of positive odd integers represented in decimal.