Problem of the Day:

Describe using set descriptor notation the complements of

(a) $\{ \varepsilon, a, aa, aaa\}$ over $\Sigma = \{a\}$
(b) $\{ \varepsilon, a, aa, aaa\}$ over $\Sigma = \{a,b\}$
(c) $\{0,1\}^* \{0010\} \{0,1\}^*$ over $\Sigma=\{0,1\}$
(d) $\{001\} \{0,1\}^*$ over $\Sigma=\{0,1\}$
(e) $\{0,1\}^* \{1101\}$ over $\Sigma=\{0,1\}$

Announcements

Tutorial #2 is posted for your tutorial this week.

Assignment #1 is due at the beginning of class this Fri. Sept. 24. Hand it in on paper (not electronically).

Make sure you sign the class attendance sheet each class.

If you want more time for our proof of the day, the notes are usually posted some time the night before or come to class early.

Assignment #1 is due on Friday.

Any questions?

Regular Languages over Alphabet Σ:

[Basis] 1. Φ and $\{\sigma\}$ for each $\sigma \in \Sigma$ are regular languages.

[Inductive step] If L_1 and L_2 are regular languages, then so are:

2. $L_1 \cdot L_2$,

3. $L_1 \cup L_2$, and

4. $L_1{}^*$.

Regular expressions over Σ:

[Basis] 1. Φ and σ for each $\sigma \in \Sigma$ are regular expressions.

[Inductive step] If α and β are regular expressions, then so are:

2. $(\alpha\beta)$
3. $(\alpha \cup \beta)$ and
4. α^*

Note: Regular expressions are strings over

$\Sigma \cup \{ (,) , \Phi , \cup , * \}$

for some alphabet Σ.

Precedence of Operators

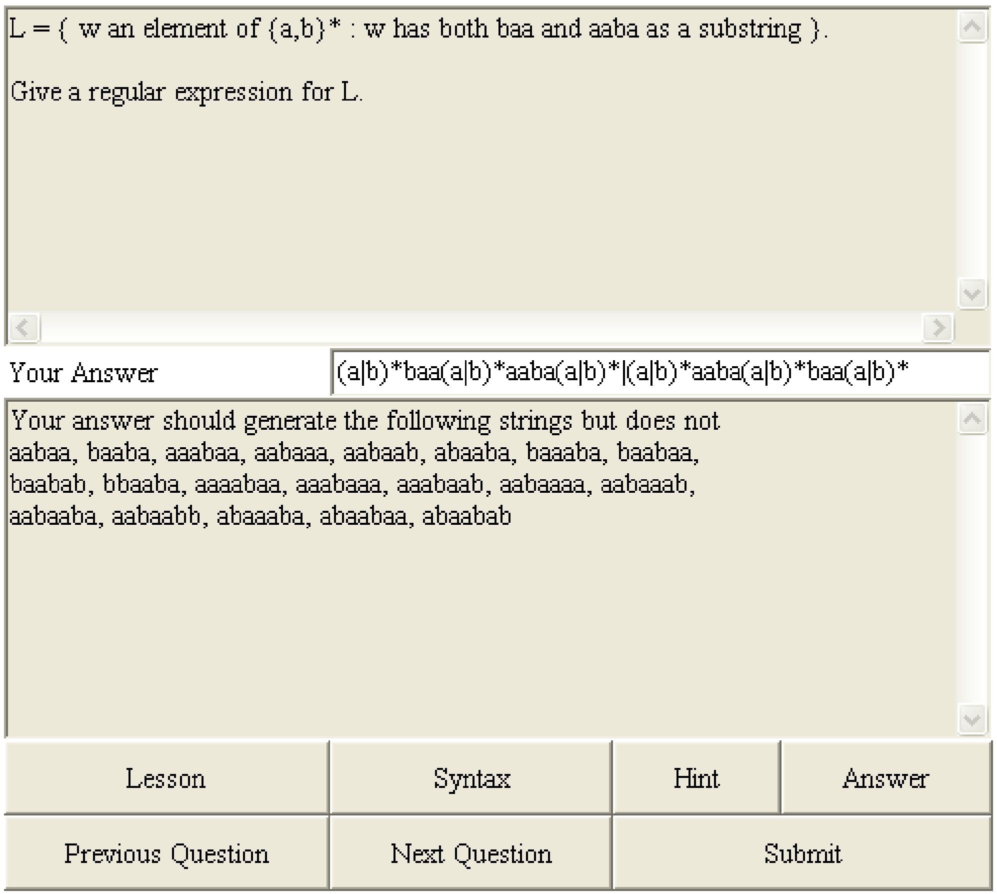
Exponents	highest	Kleene star
Multiplication	$\Downarrow$	Concatenation
Addition	lowest	Union

Prove the following languages over $\Sigma=\{0,1\}$ are regular by giving regular expressions for them:

1. $\{w$: w has odd length$\}$
2. $\{w$: w contains 0011$\}$
3. $\{w$: w does not contain 01$\}$
4. $\{w$: w starts and ends with the same symbol $(|w| \geq 1)\}$

TUTORIAL: L = { w an element of {a,b}* : w has both baa and aaba as a substring }.

(a|b)*baa(a|b)*aaba(a|b)*|(a|b)*aaba(a|b)*baa(a|b)*



MISSING:
aabaa,
baaba,

...

See home page for link to regular expression tutorial

Deterministic Finite Automata

Given a language L, the language recognition problem is:

Given an input string w, is w in L?

Answer: yes or no.

DFA's provide linear time language recognition algorithms for regular languages (languages which can be defined by a regular expression).

This class introduces DFA's, describes how they work, and defines the mathematical notation used to talk about them.

29–th letter in text

Text

Pattern: "events"

Found: at position 29

When in the course of human events it becomes necessary for one people to dissolve the political bands which have connected them with another and to assume among the powers of the earth, the separate and equal station to which the Laws of Nature and of Nature's God entitle them, a decent respect to the opinions of mankind requires that they should declare the causes which impel them to the separation.

We hold these truths to be self–evident,

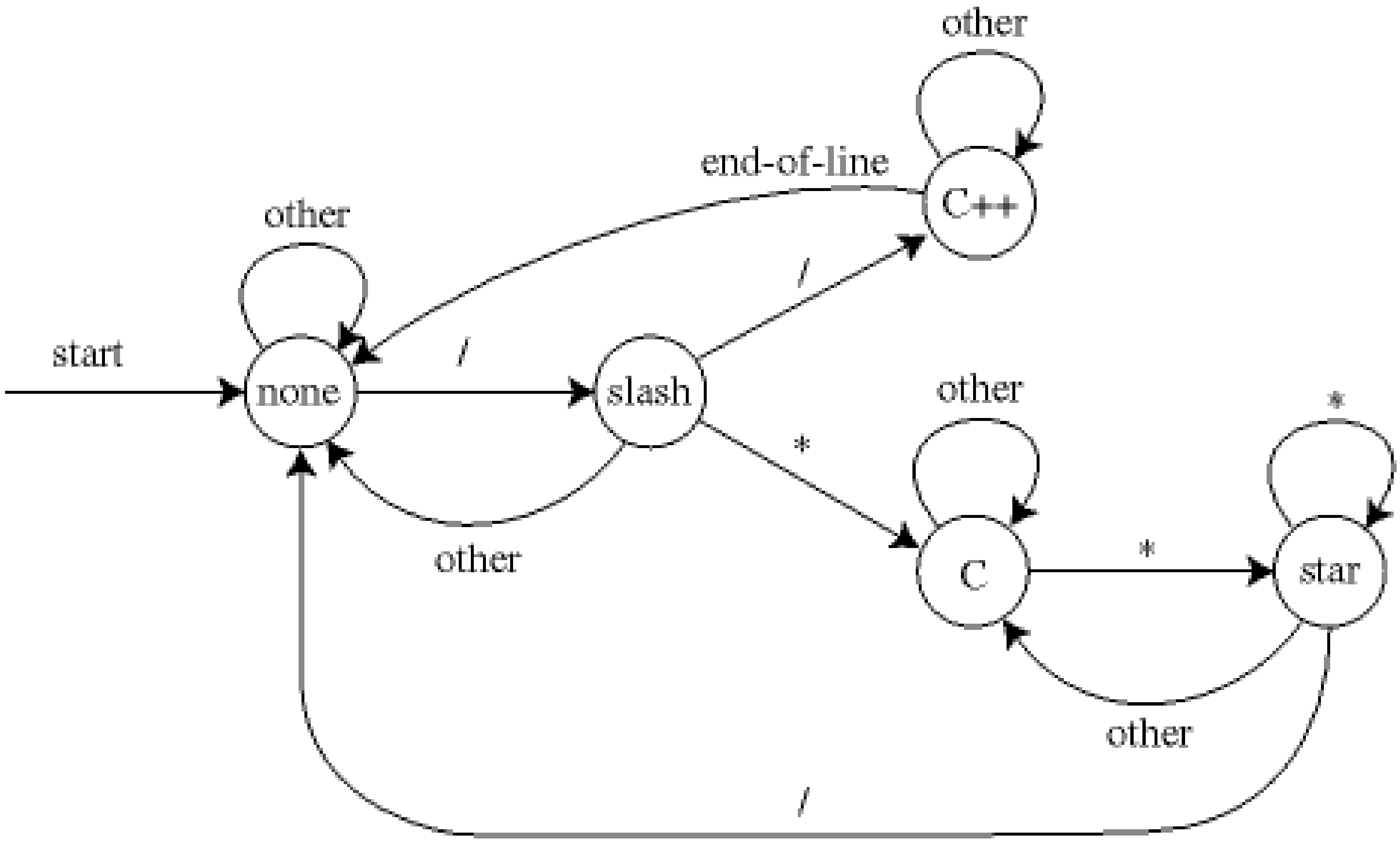
. . .

Pattern: "imp * "

Found at position ...

http://www.seas.gwu.edu/~simhaweb/cs151/lectures/module5/module5.html

To strip comments from a program:



Picture by David Eppstein.

Applications of DFA's or DFA-like code:

[*www.fbeedle.com/formallanguage/ch04.pdf*]

Compilers and other language-processing systems- to divide an input program into tokens like identifiers, constants, and keywords and to remove comments and white space.

Typed commands- the command language can be quite complex, almost like a little programming language.

Speech-processing and other signal-processing systems- to transform an incoming signal.

More Applications

[*www.fbeedle.com/formallanguage/ch04.pdf*]

Text processors- to search a text file for strings that match a given pattern. This includes most versions of Unix tools like awk, egrep, and Procmail, along with a number of platform-independent systems such as MySQL.

Controllers- to track the current state of a wide variety of finite-state systems, from industrial processes to video games, implemented in hardware or in software.

A Deterministic Finite Automaton (DFA) M is defined to be a quintuple

$(K, \Sigma, \delta, s, F)$ where

K is a finite set of states,

Σ is an alphabet,

δ is a function from $K \times \Sigma$ to K,

$s \in K$ is the start state, and

$F \subseteq K$ is the set of final states.

Some examples:

1. {w: w has odd length}
2. {w: w contains 0011}
3. {w: w does not contain 01}
4. {w: w starts and ends with the same symbol ($|w| \geq 1$)}

$M = (K, \Sigma, \delta, s, F)$:

A configuration of a M is an element of K x Σ^*.

For $\sigma \in \Sigma$, configuration $(q, \sigma w) \vdash (r, w)$

[$(q, \sigma w)$ yields in one step (r, w)]

if $\delta(q, \sigma) = r$.

The notation $\vdash^*$ means yields in zero or more steps.

$M = (K, \Sigma, \delta, s, F)$:

DFA M accepts input w if and only if

$(s, w) \vdash^* (f, \varepsilon)$ for some $f \in F$.

L(M), the language accepted by M is

$\{ w \in \Sigma^* : (s, w) \vdash^* (f, \varepsilon) \text{ for some } f \in F\}$.

$L_1 = \{ w \in \{a, b\}^* :$ w begins with aab$\}$

$L_2 = \{ w \in \{0, 1\}^* :$ has both 00 and 11 as substrings$\}$

$L_3 = \{ w \in \{a, b\}^* :$ w ends with abab$\}$

$L_4 = \{ w \in \{0, 1\}^* :$ the number of 0's in w is even and the number of 1's is congruent to 1 modulo 3$\}$

$L_5 = \{ w \in \{a, b\}^* :$ w contains abaab$\}$

What language does this DFA accept?