Given these languages, what string w would you select if you wanted to apply the pumping lemma to prove the language is not regular:

$L_1 = \{ a^n b^m : n^2 \leq m \leq n^3 \}$

$L_2 = \{ a^n b^m c^p : n=p \text{ , and } m \text{ is odd} \}$

$L_3 = \{ (01)^n 11 (01)^m : n, m \geq 0 \}$

Announcements

The midterm is in class on Wed. June 21.
There is a tutorial on Tuesday June 13.
No tutorial on Tuesday June 20.
Midterm tutorial: Monday June 19, 6:30pm, ECS 123.
Bring any questions you have about assignments 1-3, old midterms or any other class material.

Assignment 3 is due on Friday at the beginning of class.

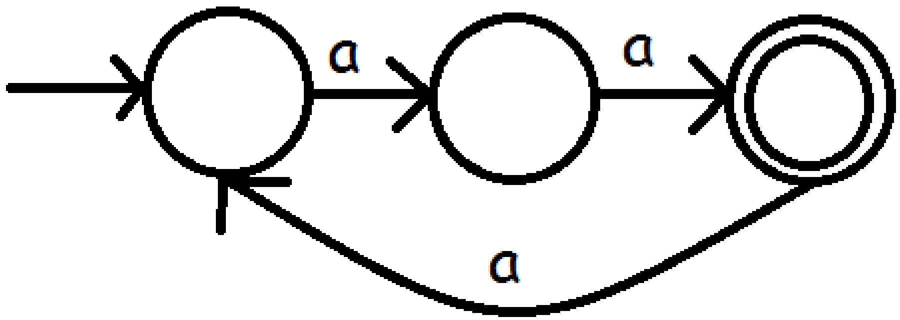
The preliminary final exam schedule has CSC 320 at 2pm on Monday August 14.

A well-parenthesized string is a string with the same number of ('s as)'s which has the property that every prefix of the string has at least as many ('s as)'s.

1. Write down all well-parenthesized strings of length six or less over the alphabet $\Sigma = \{ (,) \}$.

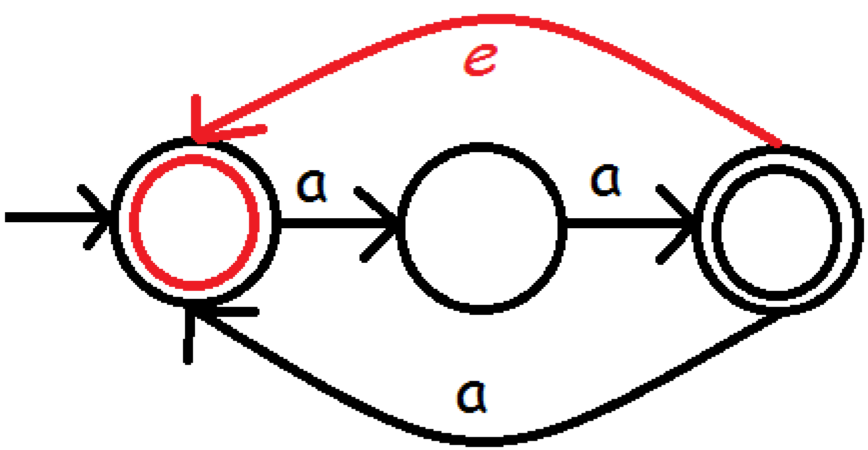
2. Let L= $\{ w \in \{ (,) \}^* :$ w is a well-parenthesized string$\}$. Prove that L is not regular.

Let M_1 be this DFA:



1. Give a regular expression for $L(M_1)$.

Let M_2 be this DFA:



2. Is M_2 a NDFA that accepts $L(M_1)^*$? Justify your answer.

Outline: Chapter 3: Context-free grammars

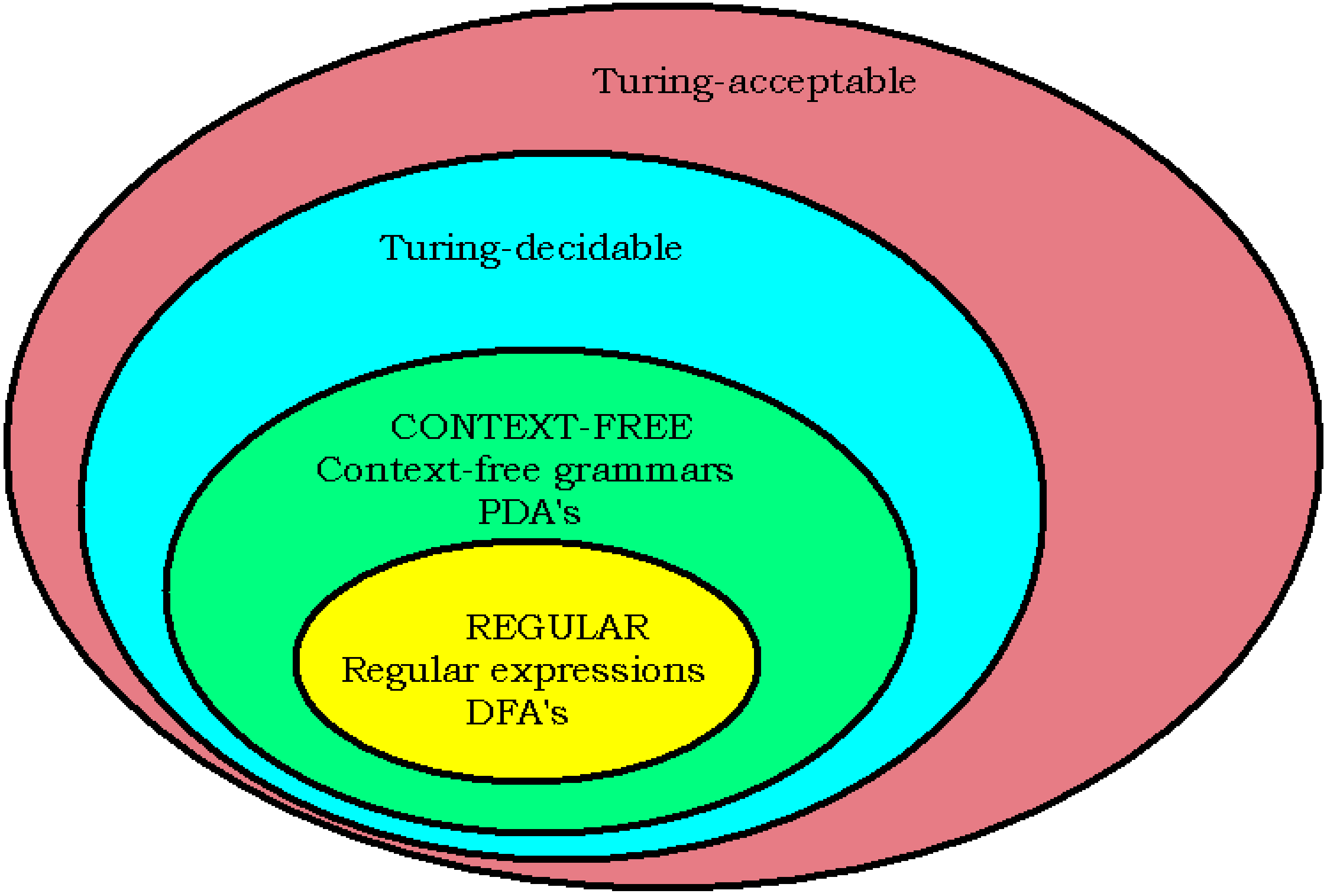
Context-free Grammars (CFG's)- used to specify valid syntax for programming languages, critical for compiling.

Pushdown Automata (PDA's)- machine model corresponding to context-free grammars, DFA with one stack.

All regular languages are context-free but not all context-free languages are regular.

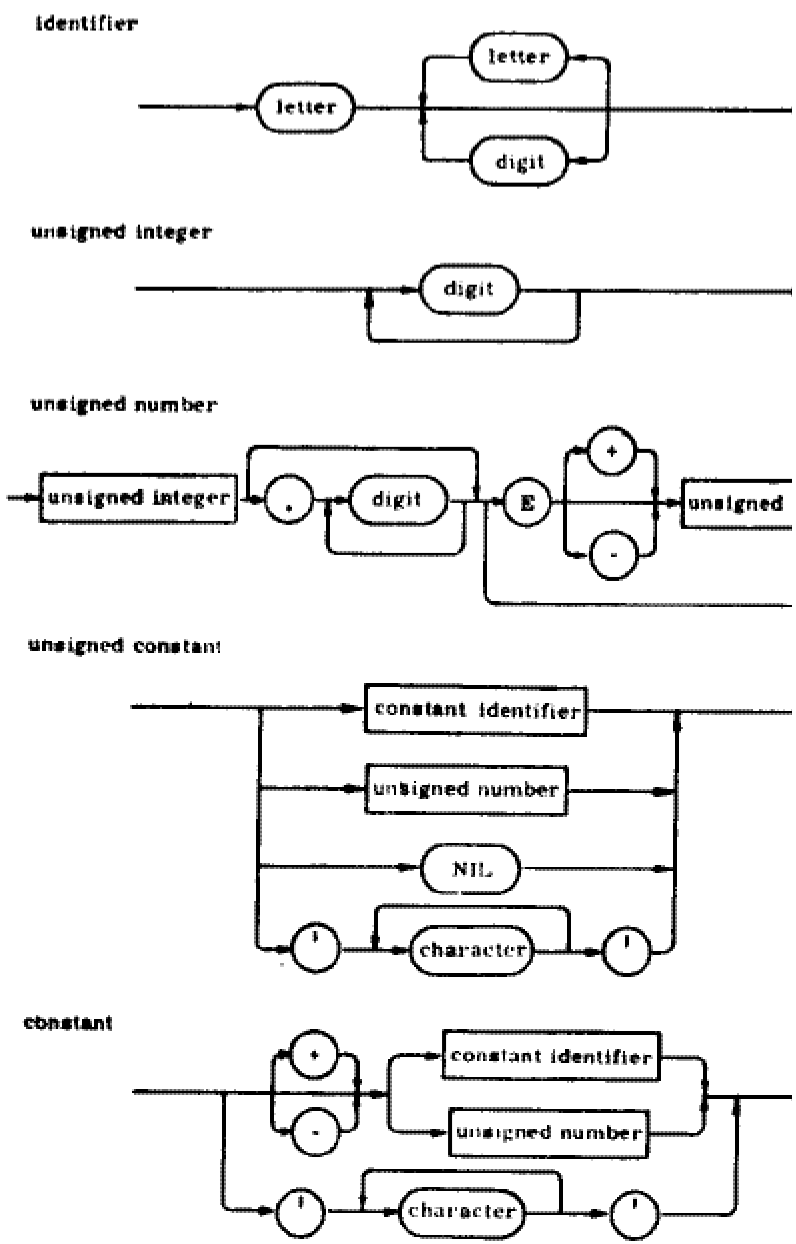
We will study closure properties, equivalence between languages specified by CFG's and PDA's, a pumping theorem and algorithmic questions.

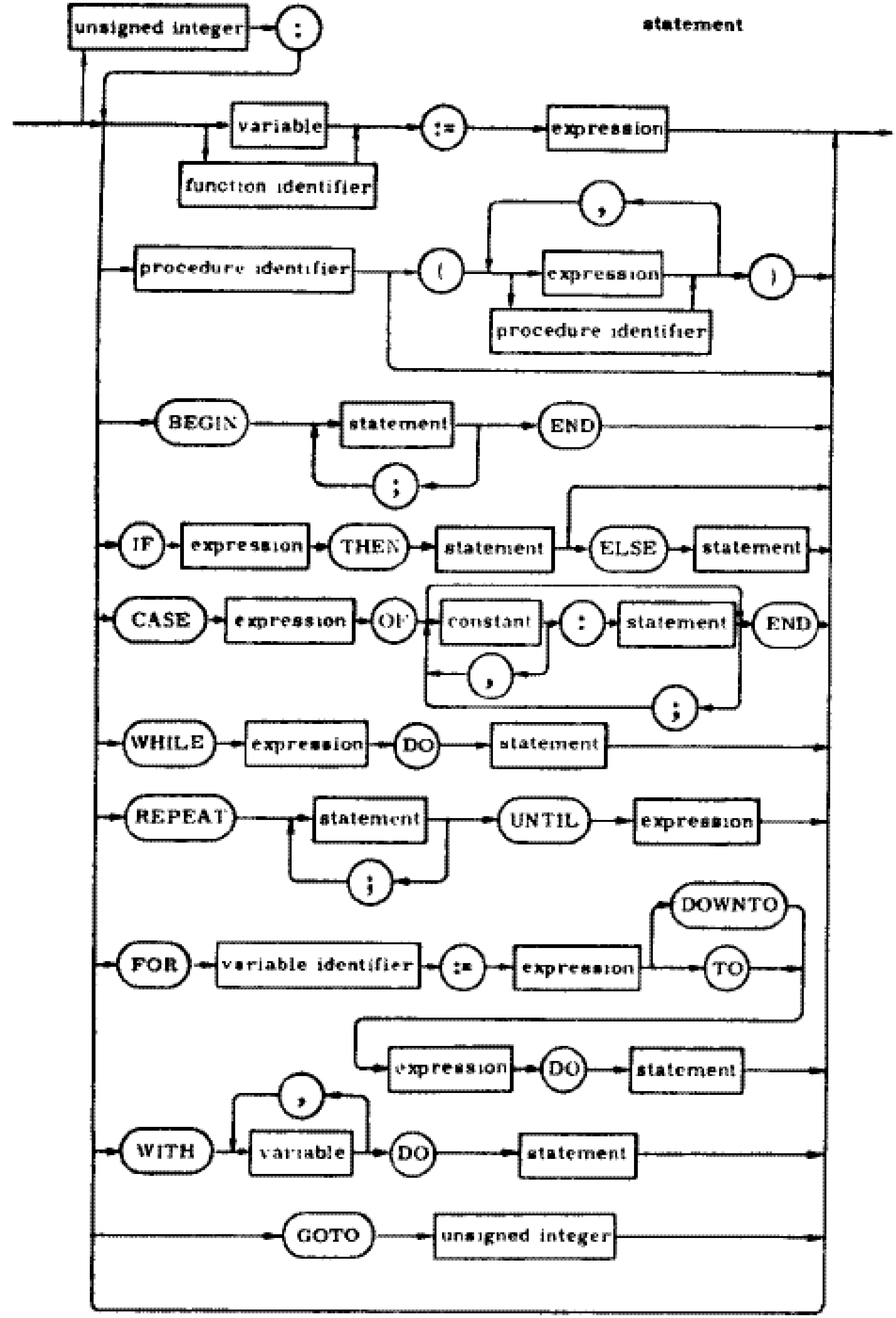
Classes of Languages



PASCAL

Part of the CFG for Pascal





Example 1:

$L = \{ a^n cc b^n : n \geq 0 \}$

Example 2:

L= (aa U aba)* bb (bb U bab)*

Example 3:

L= a a* b b* c c*

Notation

a, b, c	lower case	terminals
A, B, S	upper case	non-terminals
e	often reserved for empty string	

Strings in the generated language consist of terminals only.

$\rightarrow$ Used to describe rules.

$\Rightarrow$ Used for derivations.

$\Rightarrow^*$ Derives in zero or more steps.

Example: $L = \{ w w^R : w \in \{a, b\}^* \}$

Rules: Start symbol S

$S \rightarrow aSa$

$S \rightarrow bSb$

$S \rightarrow \varepsilon$

A derivation:

$S \Rightarrow b\,S\,b \Rightarrow b\,b\,S\,b\,b \Rightarrow b\,b\,a\,S\,a\,b\,b \Rightarrow$ bbaabb

Shorthand:

$S \overset{*}{\Rightarrow} b\,b\,a\,a\,b\,b$

A context-free grammar G is a quadruple

(V, Σ, R, S) where

V is an alphabet,

Σ (the set of terminals) is a subset of V,

R (the set of rules) is a finite subset of

$(V - \Sigma) \times V^*$, and

S (the start symbol) is an element of (V-Σ).

Elements of (V- Σ) are called non-terminals.

If (A, u) is in R, we write $A \rightarrow u$

(A can be replaced by u).

If $u, v \in V^*$, then $u \Rightarrow v$ (u derives v) if and only if there are strings x, y and z in V^*

and a non-terminal A

such that $u = x A y$, $v = x z y$, and $A \rightarrow z$ is a rule of the grammar.

$L(G) = \{ w \in \Sigma^* : S \Rightarrow^* w$, S is the start symbol$\}$

Language L is context-free if it is L(G) for some context-free grammar G.

Prove the following language is context-free by designing a context-free grammar which generates it:

L= {w in {a,b}*: the number of a's is even and the number of b's is even}

Another example:

L= { w ∈ {0,1}* : 11 is not a substring of w}

Context-free grammars and regular languages.

More examples of context-free languages.

All regular languages are context-free and a sub-class of context-free languages (those with regular context-free grammars) are regular.

Theorem:

Not all context-free languages are regular.

Proof:

$\{a^n b^n : n \geq 0\}$ is context-free but not regular.

Context-free grammar:

Start symbol S.

$S \rightarrow a S b$

$S \rightarrow \varepsilon$

Definition: A regular context-free grammar is a context-free grammar where each rule has its righthand side equal to an element of

Σ^* $(\{\varepsilon\} \cup (V - \Sigma))$

[0 or more terminals] then [at most one non-terminal]

Which rules below are not in the correct form to correspond to a regular context-free grammar?

1. $S \rightarrow A\ B$
2. $S \rightarrow B\ b$
3. $S \rightarrow a\ A$
4. $S \rightarrow aaaA$
5. $B \rightarrow b$
6. $B \rightarrow \varepsilon$
7. $S \rightarrow A\ S\ B$
8. $A \rightarrow a$
9. $A \rightarrow \varepsilon$
10. $A \rightarrow aaa\ bb$

[0 or more terminals] then [at most one non-terminal]

Theorem: If L is regular, then L is context-free.

Proof: A context-free grammar can be constructed from a DFA for L.

Definition: A regular context-free grammar is a context-free grammar where each rule has its righthand side equal to an element of

$$\Sigma^* \qquad (\{\varepsilon\} \cup (V - \Sigma))$$

[0 or more terminals] then [at most one non-terminal]

Our proof constructs a regular context-free grammar.

Create a NDFA which accepts the language generated by this context-free grammar. Start symbol: S

$S \rightarrow aa\ S$

$S \rightarrow \varepsilon$

$S \rightarrow M$

$M \rightarrow bbb$

$M \rightarrow ab\ M$

$M \rightarrow b\ S$

$M \rightarrow E$

$E \rightarrow aa$

Given the regular context-free grammar $G=(V, \Sigma, R, S)$ construct a NDFA

$M = (K, \Sigma, \Delta, s, F)$ where

$K= (V - \Sigma) \cup \{ f \}$, $s= S$, $F = \{ f \}$

For each rule $T \rightarrow u R$ with $u \in \Sigma^*$, $R \in V-\Sigma$,
add a transition (T, u, R) to Δ.

For each rule $T \rightarrow u$ with u in Σ^*,
add a transition (T, u, f) to Δ.