

CSE6390 3.0 Special Topics in AI & Interactive Systems II  
Introduction to Computational Linguistics  
Instructor: Nick Cercone – 3050 CSEB – [nick@cse.yorku.ca](mailto:nick@cse.yorku.ca)  
Tuesdays,Thursdays 10:00-11:30 – South Ross 104  
Fall Semester, 2010

---

**THE SMALL ASSIGNMENT**

---

1. Write a regular expression for the language L, over alphabet {0, 1, 2}, such that every 0 that is not the last (rightmost) symbol is immediately followed by a 1, and every 1 that is not the last symbol is immediately followed by a 0. By "immediately followed" we mean "with no intervening symbols. Thus, 010 is in L, but 001 is not, because the first 0 is not *immediately* followed by 1. 021 is not in L for the same reason.
2. Let L be the language consisting of all strings of zero or more 0's followed by one or more 1's followed by two or more 2's. For example 001122, 122, and 0111122 are in L: 012 (too few 2's), and 0112122 (a 2 precedes a 1) are not.

Write a regular expression whose language is L

3. Give a context free grammar generating the same language, It is sufficient just to give the productions, assuming that S is the start symbol

4. Arithmetic expressions with operator + and parentheses can be generated by the grammar

$$E \rightarrow E+E \mid (E) \mid a$$

where *a* stands for any number. This grammar is ambiguous.

- (a) Give an example of a string that has two or more leftmost derivations or parse trees.
- (b) Design an unambiguous grammar for the same language.

5. Here is a context free grammar  $G = (\{S, A, B\}, \{0, 1\}, P, S)$ , where P is the set of productions\_

$$S \rightarrow 0A \mid 1B \mid \epsilon$$

$A \rightarrow 1S \mid 0AA$

$B \rightarrow 0S \mid 1BB$

Intuitively, A generates strings with one more 1 than 0, B generates strings with one more 0 than 1, and S generates the strings with equal numbers of 0's and 1's

Give a leftmost derivation of the string 0011

Give a rightmost derivation of the string 001011

Draw a parse tree for the string 1001

6. Show how one would represent the following natural language statements in predicate calculus form.

(a) Sidney Crosby is a centre.

(b) If Fred is a dog and Bruce is a cat, then Fred chases Bruce.

(c) No one likes Bill.

(d) Bill gave the dirty old green book to his mother.

(e) Computing Science is part of the Faculty of Applied Science.

(f) All dogs chase some cat (or other).

(g) All women are mortal.

(h) Someone is loved by everybody.