

# Homework 1 (25pts)

due before class Feb 8, 2012

1. Describe the syntax of each of the following languages using BNF.
  - (a) Strings of length five or more over the set of terminals  $\{p, q, r\}$ .
  - (b) Strings over  $\{a, b, c, d\}$  that start with  $ab$  and end with  $cd$ .
  - (c) Strings over  $\{0, 1, 2\}$  that contain an even number of tokens, e.g., 2011 and 12 are in the language, but 121 is not.
2. The set of well-formed propositional logic formulas can be considered as a language over the set of terminals  $\{p, \neg, \vee, \wedge, \rightarrow, (, )\}$ , where  $p$  is a propositional variable and can be mapped to an arbitrary identifier (i.e., variable name),  $\neg$  is the negation unary operator, and  $\vee, \wedge, \rightarrow$  are binary operators representing *or*, *and*, *implication* relations respectively. For example, expression  $p1 \wedge p2 \vee p3$  means  $p1$  and  $p2$  or  $p3$ , and expression  $p1 \vee (x \rightarrow y)$  means  $p1$  or (if  $x$  then  $y$ ).
  - (a) Use Backus-Naur Form (BNF) to describe the syntax of the given language.
  - (b) Based on your BNF, give a parse tree and an AST (Abstract Syntax Tree) for the input  $p1 \vee (x \rightarrow y)$ .
  - (c) Is your grammar ambiguous? If yes, rewrite your grammar to be non-ambiguous by enforcing that  $\wedge$  has higher precedence than  $\vee$ ,  $\vee$  has higher precedence than  $\rightarrow$ , and all binary operators are left associative.