CSCI 5832 Spring 2011 Exam 2

Name: _____

On my honor, as a University of Colorado at Boulder student, I have neither given nor
received unauthorized assistance on this work. _____ .

1. **(5 points)** What is your favorite color?

2. **(5 points) True/False:** In developing a grammar, the primary motivation for the
   inclusion of a potential **constituent** (i.e., a non-terminal category) in the grammar is
   the meaning of the phrases covered by that category.

3. **(5 points)** Given the grammar on the attached sheet, provide a valid parse tree for
   the following example: *Linda flew United to Chicago*. Use the back of this page.

4. **(5 points) True/False:** The previous example is ambiguous given this grammar.

5. **(5 points) True/False:** Given a sentence *S*, the following equation adequately
    describes the computation performed by the probabilistic CKY algorithm given in
    the text.

$$P(S) = \sum_{Ts.t.S=Yield(T)} P(T)$$

6. Probabilistic parsing with PCFGs can be improved via parent annotation.
   a) **(5 points)** Illustrate this technique on the tree you gave for question 3
      (make a copy of it; don't annotate the original).
   b) **(5 points)** Out of all the non-terminal categories in this grammar, which
      one has the potential to be most improved (helped) by this technique?

7. **(15 points)** Provide the missing semantic attachments for the relevant rules on the
   attached sheet such that the following FOL form could be produced as the
   representation for the given sentence given in question 3.

   ∃*e flying(e) ∧ flier(e,Linda) ∧ Carrier(e,United) ∧ Destination(e,Chicago)*

8. **(5 points)** Consider the problem of named-entity recognition where we're looking
   for **people**, **organizations**, and **locations**. Provide an appropriate IOB-style
   annotation for the "United" example on the following page.

9. **(10 points)** Consider the problem of extracting relations between given named entities in this example. Describe three **types** of features that would be useful in classifying the relation between Rahsaan Johnson and the second United token in this example.

10. **(10 points)** A relation classification system would also have to consider possible relations between Rahsaan Johnson and Chicago. What problem does this pose and what kind of features might be helpful in avoiding it? Be specific.

Grammar 1

1. S → NP VP   vp.sem(np.sem)
2 NP → PN
3 VP → Verb NP NP
4 VP → Verb NP PP
5 VP → VP PP
6 PP → P NP

7 Verb → *flew*

9 PN → *Linda*
10 PN → *United*
12 PN → *Chicago*

13 P → *to*

A United jet lost electrical power forcing an emergency landing at the New Orleans airport, said Rahsaan Johnson, a spokesman for Chicago based United.