Assignment 3
Due Wed. 30^{th} Nov.
Total points: 40

1. Consider the following context free grammar with start symbol S:

$$
\begin{array}{ll}
S \rightarrow NP\ VP & PP \rightarrow Pre\ NP \\
S \rightarrow \textit{I}\ VP\ PP & V \rightarrow \textit{ate} \\
NP \rightarrow Det\ N & Det \rightarrow \textit{the} \ | \ \textit{a} \\
VP \rightarrow \textit{ate}\ NP & N \rightarrow \textit{fork} \ | \ \textit{salad} \\
VP \rightarrow V & Pre \rightarrow \textit{with}
\end{array}
$$

   (a) Convert this grammar to Chomsky Normal Form [2 points]

   (b) Use the CYK algorithm to parse the sentence, representing the CKY chart in matrix form. [3 points]
   I ate the salad with a fork

   (c) How many complete analyses of S do you get? Draw their parse trees. [1 points]

2. (a) Derive a PCFG grammar from the corpus below. Write down the rules of the grammar and calculate their probabilities. Assume for the purpose of the grammar that the corpus is lowercased (e.g. *The* has been replaced by *the*). [4 points]

```
(S
    (NP She)
    (VP
        (VP
            (V saw)
            (NP (Det the) (N man)))
        (PP
            (P from)
            (NP (Det a) (N distance)))))
```

```
(S
    (NP Here)
    (VP
         (V is)
         (NP  (Det a)  ( N telescope))))
```

```
(NP
    (NP (Det the) (N man))
    (PP (P with) (NP (Det the) (N guitar))))

(S
    (NP He)
    (VP
        (V saw)
        (NP (Det the) (N girl))))

(S
    (NP (Det The) (N man))
    (VP
        (V saw)
        (NP
            (NP (Det the) (N girl))
            (PP
                (P with)
                (NP (Det the) (N flowers))))))
```

(b) Draw all possible parse trees for the following sentence, and compute their probabilities using the probabilistic grammar you have created. 2 points
He saw the man with the telescope

(c) You can also use your probabilistic grammar for text prediction much like the suggestion box in a search engine. If a user typed "The girl saw" , what is the most likely suggestion for a possible completion? Justify your answer. 3 points

3. Consider the PCFG below: 6 points

S	→	Subj VP (1.0)
VP	→	V Obj (0.5) | V Obj Obj (0.3) | V Small (0.2)
Small	→	Obj V (1.0)
Subj	→	*I* (0.3) | NP (0.7)
Obj	→	*her* (0.2) | NP (0.8)
NP	→	N (0.5) | Det N (0.5)
V	→	make (0.6) | duck (0.4)
N	→	duck (0.5) | goose (0.5)
Det	→	her (1.0)

Use a probabilistic CYK-style algorithm to find the most probable parse for the sentence
I make her duck

Note: For this example, do not convert the grammar to Chomsky Normal Form. Instead, you should draw up a CYK-style parse chart for the grammar as it stands. For small examples this is perfectly feasible, but you should check that you understand why a CYK- style algorithm has difficulties with non-CNF grammars in general.

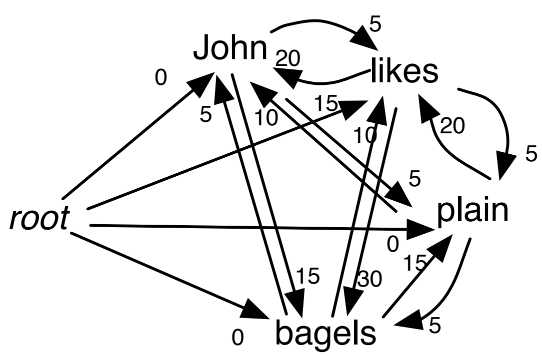
Hint: The sentence has three readings or meanings; 'I make a duck for her' , 'I make a duck that is her's ', and finally, one in which *duck* is a verb and *her duck* forms a "small clause", corresponding to a reading in which 'she' is 'ducking'.

4. Dependency parsing / MST **[6 points]**

   Imagine that you are using an arc-factored non-projective dependency parsing model. You are given a sentence: *John likes plain bagels.* In this exercise, we consider that there are no labels on arcs. So, the goal is to choose a dependency structure which maximises the sum of arc scores:

$$G = \arg\max_{G \in T(G_x)} \sum_{(i,j) \in G} w_{ij} \tag{1}$$

   The scores for each directed arc are given in the following graph (the table is the corresponding matrix):

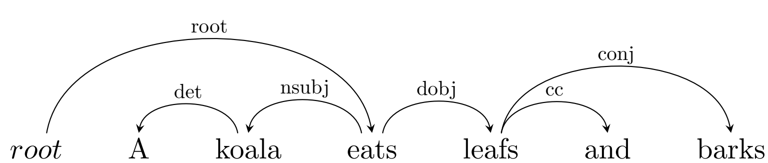


dependends \ *heads*	root	John	likes	plain	bagels
John	0	-	20	10	5
likes	15	5	-	20	10
plain	0	5	5	-	15
bagels	0	15	30	5	-

   Explain, step by step, how the Chu-Liu-Edmonds (CLE) algorithm is applied to the given example (similarly to how it was done in our slides) and show what the resulting MST is.

5. Dependency parsing / Transition-based **[6 points]**

   Consider a sentence "A koala eats leafs and barks". Assume that a correct dependency tree for this sentence is the following:



   Will a transition-based dependency parser (using the arc-standard system described in class) be able to correctly predict this structure. List transitions (i.e. configurations and decisions made). If there is a mistake, explain what was the mistake (something like: in configuration X the parser made the decision W instead of Z). If it is tricky for you to draw that many dependency graphs, it is sufficient to show for each configuration the stack, the buffer, what kind of action was made, and which arcs are added to the arc set A (if any).

For extra credit (4 points over and above the points for this question), also solve using the *arc-eager version* of Nivre's algorithm.

6. (a) Give a CCG derivation for the following sentence: **3 points**

   The company added four Boeing-747s to the two units in 1994

   Indicate clearly which CCG combinator is being used at each step, and assume the following lexical categories:

   added : $((S \setminus NP)/NP)/PP$
   to : PP/NP
   in : $((S \setminus NP) \setminus (S \setminus NP))/NP$
   The, four : NP/N
   two : N/N
   1994: NP
   Boeing-747s, company, units : N

   Note: CCG assigns different lexical categories to prepositional phrases that are 'arguments' of verbs and those that are 'adjuncts'. In the above sentence, the prepositional phrase 'to the two units' is an argument of *add*. The other prepositional phrase 'in 1994' is an adjunct, and will combine using a rule for modifiers (this rule is in addition to the ones discussed in class)

   $X \;\; X \setminus X \Rightarrow X$
   $X/X \;\; X \Rightarrow X$
   where X can be any category.

   (b) Now give a derivation for the following noun phrases, proposing lexical categories for *which* in each case: **3 points**

   the company which added four Boeing-747s to the two units in 1994
   the four Boeing-747s which the company added to the two units in 1994

   (c) Give an example of a sentence in Dutch which illustrates the non-context free nature of some structures in language. Clearly indicate the dependencies. (For this question, you might have to do some reading/internet research on your own, since we only briefly touched upon this in the lectures) **1 points**