CPSC 422 2021W1
Homework 4 (Due April 14, 3:30pm) (tot: 100 points)

Question 1 – Similarity Metrics in Ontologies (30pts)

To answer this question you will compare your intuition about the similarity between different word senses (i.e., WordNet synsets) with the output of several similarity measures defined on WordNet.

First, select 15 pairs of synsets from WordNet. You can look for synsets at http://wordnetweb.princeton.edu/perl/webwn. Type in a word to find all its synsets. To get the corresponding sense numbers select *"Show Sense Numbers"* in the *display options* menu and click on the *Change* button. With this interface you can also explore all the relationships between synsets (e.g., hypernym).

Select your synset pairs so that:

- they include pairs of senses that range from very similar ones like *dog#n#5* and *sausage#n#1*, to less similar ones like *dog#n#1* and *tooth#n#1*, to very different ones like *snake#n#2* and *plane#n#2*
- they include 5 pairs of senses from different POS (e.g., *fly#v#1* and *bird#n#1*)

Now rank your pairs according to how similar you think the two senses in each pair are, from the most similar pair to the least similar one. After this, apply the three WordNet similarity measures of *lesk*, *jcn* and *path* provided at http://ws4jdemo.appspot.com/ to your pairs. (In this interface, once it has loaded WordNet, if you type a word followed by # you will be prompted with a list of all the senses for that word, from which you can select the ones you want to compare).

Note: *the ws4jdemo app has a limited number of uses per day, so you should* ***not*** *leave this question until the last minute.*

Finally, for each similarity measure, rank your pairs according to how similar they are according to that measure (again from the most similar pair to least similar one).

Do the different measures produce different rankings? What measure is the most consistent with your own ranking? Also, discuss possible errors/inconsistencies/weaknesses in the rankings, if any.

Question 2 – Probabilistic Context-Free Grammar (PCFG) (50pts):

Consider the *mystery-code.pl* file. It is described as a program that, given a PCFG and a sentence, returns the most likely parse for the sentence according to the grammar. The grammar and the target sentence must be specified in a single file as exemplified in the file *pcfg-grammar.pl* (where the sentence is "*John plays soccer at soccer*").

You can run the code on a grammar/sentence file in a Unix/Linux/MacOS terminal:

```
perl mystery-code.pl pcfg-grammar.pl
```

(a) (20 points) Suppose that you need to use this code in your AI system; but before using it you want to make sure that it is correct (always do this!). As a preliminary step you want to make sure that it gives the right answer on at least the sample input *pcfg-grammar.pl*. Is this the case? Show your work.
(If you know Perl you may also try to figure out what algorithm *mystery-code.pl* is implementing – no points for this, but hey - you can learn something ☺)

(b) (30 points) As previously stated, the file *pcfg-grammar.pl* contains a simple PCFG. As you can see, all the probabilities were assigned assuming a uniform distribution on the rules for each non-terminal. Well, it is your lucky day and you find on the Web a silly corpus that will allow you to get some "better" estimates for your probabilistic CF grammar (see file *silly-corpus*).

- Compute estimates for the grammar probabilities based on *silly-corpus.* The corpus is so small that you can do it by hand, but if you are looking for **(10)** bonus points you can implement a program that would do it for you on a less-silly-corpus. (If you implement such a program, include your code in your submission.)
- Given this new grammar, compute the probability of the most likely parse for each of the following two sentences: *John plays soccer at school* and *John plays soccer* (you can use the mystery code for finding the most likely parse)
- Finally, compute the same two probabilities with the old grammar. Are the results from the old and the new grammar different? Explain why this is (or is not) the case for each of the two sentences.

Question 3 – IBM Watson and what we covered in 322/422 (20 points)

Watson is an AI system that was developed at IBM Research to "compete at the human champion level in real time on the American TV quiz show, Jeopardy" (http://www.jeopardy.com/). Search the Web for reputable sources about how Watson works. You can start from http://www.aaai.org/Magazine/Watson/watson.php.

Identify what representation and reasoning methods we studied in 422 are applied in Watson. Describe in as much detail as you can how they were applied. Relate your descriptions to concepts and methods we studied in 322/422. Include a list of the sources you used. **5 bonus points** will be awarded to students who provide above average details.