CS 7337 – Natural Language Processing
Midterm Exam

Instructions: Clarity of answers is more important than length of answers. Although not required (unless indicated otherwise), feel free to use graphs, charts, visuals, et al in your answers if you feel these artifacts can help support your answers. There are no bonus points for using these artifacts.

Due date: See instructors note; submission should be in PDF or Word DOCX file format.

Q1. **a. [5 pts] Define homonymy and polysemy and give an example of each.**

b. [5 pts] Define NLU and NLG and give an example of each.

Q2. **You are given the following grammar for expressions:**

E → I	I → a
E → E + E	I → b
E → E * E	I → 0
E → (E)	I → 2

a. [10 pts] Show parse tree(s) for the expression 2 + 2 * 2

b. [10 pts] Describe any interesting observations in your answer to a.
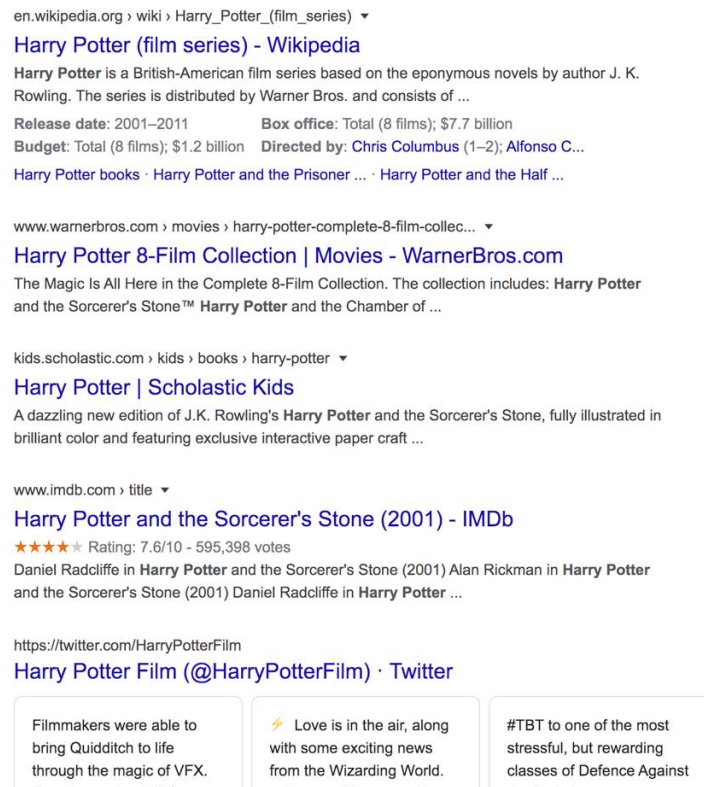
Q3. Consider the following grammar and sentence:

$$S \rightarrow NP \; VP$$
$$NP \rightarrow Pronoun$$
$$NP \rightarrow Proper\text{-}Noun$$
$$NP \rightarrow Det \; Nominal$$
$$Nominal \rightarrow Nominal \; Noun$$
$$Nominal \rightarrow Noun$$
$$VP \rightarrow Verb$$
$$VP \rightarrow Verb \; NP$$
$$VP \rightarrow Verb \; NP \; PP$$
$$VP \rightarrow Verb \; PP$$
$$PP \rightarrow Preposition \; NP$$
Nominal --> Nominal PP

Sentence: *I booked a flight from LA*

   a. [10 pts] In what way is this sentence ambiguous? Describe different interpretations of this sentence.

   b. [10 pts] Show the parse trees for this sentence and where the ambiguity manifests in the parse trees.

Q4. The image below shows Google search results for the query "harry potter"

As the results show, the query could represent any of the seven books in the harry potter franchise, any of the film adaptations of the books, a theme park, or a ride, an audiobook, cartoons, et al.

a. [10 pts] Discuss why google shows a mix of such results and what factors can influence the search results for this query that will be presented to you.
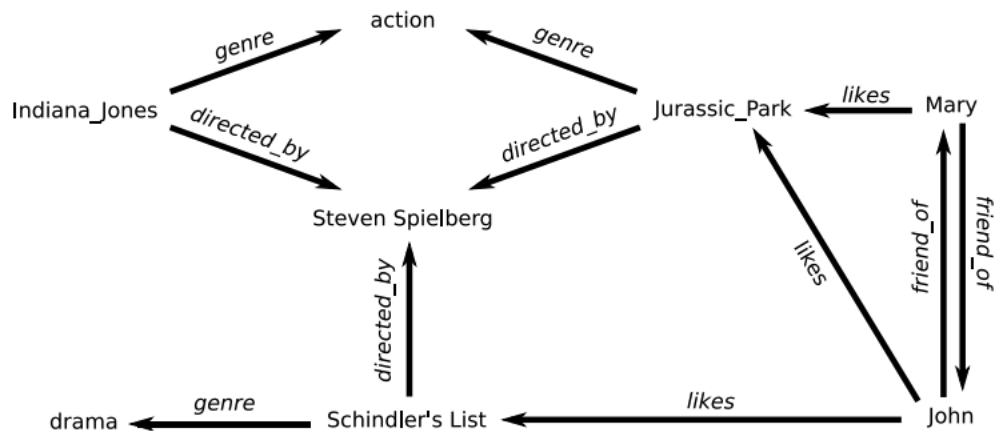
b. [15 pts] Consider the following sentence:

*The **bank** can guarantee deposits will eventually cover future tuition costs because it invests in adjustable-rate mortgage securities.*

The word **bank** has multiple senses. Use Wordnet to show the top two sense, glossaries and examples for **bank** and describe (at a high level) how you can use this information to find the proper sense for this word in a sentence.

Wordnet link: http://wordnetweb.princeton.edu/perl/webwn

Q5. You are building an online moving streaming service which enables looking up information on movies, genres, directors, actors and customer movie preferences.

**(Indiana Jones refers to Raiders of the Lost Ark (1981))

a. [10 pts] What is the customers intent (i.e. what are they looking for) with the following queries? (these are individual queries, not queries entered in succession)

"Drama", "Jurassic Park", "Indiana Jones: Raiders of the lost ark", "Steven Spielberg"

b. [5 pts] A customer searches for "Indiana Jones" but clicks on and watches "Jurassic Park" – what insights can you get from this customer action?

c. [10 pts] The customer searches for "Indiana Jones: Raiders of the lost Ark" but it's not available in their region (US, EU, Asia). What search results would you show the customer? Discuss how you would build that experience from a technical design perspective.