CSC	495	and	791

H2

Natural Language Processing

(Signature) _

Fall 2023

(Write your name above)

Student's affirmation: I certify that I have neither taken help in completing this assignment nor helped anyone else with this assignment. I have never discussed this assignment with anyone other than the instructor and TA.

Mandatory: Affirmation and signature on the first page; name on every page; submission as PDF. If you make assumptions about any problem, state them, but be prepared to justify why they were necessary.

Problem	1	2	3	4	5	Total
Points:	10	25	20	23	22	100
Score:						

This assignment has 5 problems, for a total of 100 points.

Throughout, I prefer that you think afresh but if you come across a source on which you base your answer, please be sure to cite it. Please note that if you are citing a source to bolster your claim then the source should be authoritative (such as our textbook or a book by another credentialed author). Blog posts by strangers are not credible.

- 1. (10 points) Mark the following statements true or false. Provide a short explanation of about 10–20 words. You can and should provide a source where appropriate, including the specific page and line numbers.
 - A. In dependency parsing, a criterion for determining the modifier of a head-modifier construction is that the modifier determines the semantic category of the construction
 - B. Although appositives are not referents, some datasets identify them as a way to provide parenthetical information on a referent
 - C. Since named entities rarely arise consecutively in English, an important feature in judging whether something is a named entity is whether it is adjacent to a named entity
 - D. People whose language doesn't include markers for singular or plural can't talk about the numbers of entities in that language
 - E. The selectional preference of a verb captures whether it supports dative alternation
- 2. Consider the sentence "She sketched a scene in the courthouse that was dull."

Where possible use only the dependency relations specified in the textbook and copied in our lecture notes. If you need to go beyond that list, make a case for doing so but don't assume I will accept the extension.

- (a) (12 points) Show a *projective* dependency parse tree for (some interpretation of) this sentence. State the corresponding interpretation of this sentence.
 - Show a series of shift-reduce moves of the Arc Standard Parser through which this parse could be created, or explain that no shift-reduce sequence can produce this parse.
- (b) (13 points) Show a nonprojective dependency parse tree for (some interpretation of) this sentence.
 - State the corresponding interpretation of this sentence.
 - Show a series of shift-reduce moves of the Arc Standard Parser through which this parse could be created, or explain that no shift-reduce sequence can produce this parse.
- 3. Consider the following passage from an example in the textbook with modifications.

Fall 2023

(Write your name above)

- 1 The Justice Department is conducting an anti-trust trial against Microsoft Corp. with evidence that the company is increasingly attempting to crush competitors
- 2 Microsoft is accused of trying to forcefully buy into markets where its own products are not competitive enough to unseat established brands
- 3 The case revolves around evidence of Microsoft aggressively pressuring Netscape into merging browser software
- 4 Microsoft damaged Netscape through a price war
- 5 Macroeconomic conditions killed Netscape
- 6 The government may file a civil suit ruling that conspiracy to curb competition through collusion is a violation of the Sherman Act
- 7 Microsoft continues to enjoy increased earnings despite the trial
- (a) (12 points) Identify each occurrence of the thematic roles listed in Figure 19.1 of the textbook, which is reproduced in the table below.

Th. Role	Definition	Sample Words
Agent	Volitional causer (includes accidents)	Kick
Experiencer	One who experiences it	Has (a feeling)
Force	Nonvolitional causer	Tsunami
Theme	One (most) directly affected	Shut (the door)
Result	Outcome	Wrote (a book)
Content	Proposition of a propositional event	Asked
Instrument		With a screwdriver
Beneficiary		For his son
Source	Origin of the object in a transfer event	Shipped
Goal	Destination of the object in a transfer event	Delivered

(b) (8 points) Recall FrameNet's *Change Position on a Scale* frame from the textbook, especially in Figure 19.3, which enumerates its frame elements. Apply this frame to the following sentence (simplified from https://en.wikipedia.org/wiki/Netscape):

"Netscape rapidly lost market share falling from more than 90 percent in the mid-1990s to less than 1 percent in 2006."

State which of the specified frame elements are defined in this sentence.

- 4. Consider the following passage taken from https://en.wikipedia.org/wiki/Netscape (with end notes removed).
 - 1 Netscape stock traded from 1995 until 1999 when the company was acquired by AOL in a pooling-of-interests transaction ultimately worth US \$10 billion.
 - 2 In February 1998, approximately one year prior to its acquisition by AOL, Netscape released the source code for its browser and created the Mozilla Organization to coordinate future development of its product.
 - 3 The Mozilla Organization rewrote the entire browser's source code based on the Gecko rendering engine, and all future Netscape releases were based on this rewritten code.
 - 4 When AOL scaled back its involvement with Mozilla Organization in the early 2000s, the Organization proceeded to establish the Mozilla Foundation in July 2003 to ensure its continued independence with financial and other assistance from AOL.
 - 5 The Gecko engine is used to power the Mozilla Foundation's Firefox browser.

Explain each answer briefly, taking about 60–80 words for all parts.

Fall 2023

(Write your name above)

- (a) (5 points) Identify all noun phrases in this passage that may be referents.
- (b) (8 points) Identify all coreference clusters based on this passage.
- (c) (10 points) Identify all phrases that are appositives, predicate nominals, expletives or pleonasts, clefts, extrapositions, numerics, or embedded in larger mentions. (Note that there may be no matches for any of these categories.)
- 5. Explain your criteria for each answer pertaining to the following text extracted from https://en.wikipedia.org/wiki/Obdulio_Varela.
 - 1 Obdulio Jacinto Muiños Varela was a Uruguayan football player.
 - 2 He's most remembered as the captain of the Uruguay team that won the 1950 FIFA World Cup.
 - 3 The decisive match was against the hosts Brazil.
 - 4 Five minutes in the second half, Brazil scored, and Varela took scene, when he intentionally walked slowly to his goal, picked up the ball and then argued with the English referee George Reader about a nonexistent offside, with the intention of delaying the restart of the game so the crowd cooled off.
 - 5 After that he said to his teammates 'Now it's time to win the game' and the Brazilians in the crowd were in a silent mood.
 - (a) (4 points) Show the direct and indirect realizations for each utterance.
 - (b) (10 points) Compute the values for C_b , C_f , and C_p for each utterance. Show these values in a table with one row for each utterance.
 - (c) (4 points) Show the resulting centering transitions. Show the transitions in a table with one row for each utterance.
 - (d) (4 points) Are the conclusions regarding the transitions (and the imputed discourse coherence) reasonable? If so, explain. Otherwise, explain how you would alter the method to produce more acceptable results.