11-411/11-611 Midterm Exam II

David R. Mortensen and Lori Levin

Spring 2023

Name: _____

Andrew ID: _____

INSTRUCTIONS: Complete this exam with a **blue or black pen**. Pencil will not scan properly and you may lose points if you use it. Mark multiple choice and multiple answer questions by filling in the parentheses around the letter (as if it were a bubble). **Checks or marks next to the answer will be missed by the autograder.** Answer all questions in the space given. **The backs of pages will not be scanned** and answers written there will not be evaluated. Multiple answer questions will be scored as $F1 \times$ point_value.

1 (20 points) Attention, Selves!

You have constructed a neural model based on self-attention and inspecting it to see if it works. As a test case, you take the first tweet you can find:

Elon often praises himself

You convert each token to a two-dimensional embedding x_i:

	Elon	often	praises	himself
$\mathbf{x} =$	$\begin{bmatrix}3 & 4\end{bmatrix}$	$\begin{bmatrix}0 & 1\end{bmatrix}$	$\begin{bmatrix}1 & 1\end{bmatrix}$	$\begin{bmatrix}3 & 3\end{bmatrix}$

You have learned three weight matrices, W^Q, W^K, and W^V:

$$W^Q = \begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} \qquad W^K = \begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix} \qquad W^V = \begin{bmatrix}3 & 4\\ 9 & 8\end{bmatrix}$$

1. (1 points) Compute the missing query vector for the embeddings in **x**

$\mathbf{Q} =$	$\begin{bmatrix}7 & 7\end{bmatrix}$	$\begin{bmatrix}1 & 1\end{bmatrix}$	$\begin{bmatrix}2 & 2\end{bmatrix}$	$\begin{bmatrix}6 & 6\end{bmatrix}$

2. (1 points) Compute the missing key vectors for the embeddings in **x**

$\mathbf{K} =$	$\begin{bmatrix}3 & 7\end{bmatrix}$	$\begin{bmatrix}0 & 1\end{bmatrix}$	$\begin{bmatrix}1 & 2\end{bmatrix}$	$\begin{bmatrix}3 & 6\end{bmatrix}$

3. (1 points) Compute the missing value vectors for the embeddings in **x**

$\mathbf{V} =$	$\begin{bmatrix}45 & 44\end{bmatrix}$	$\begin{bmatrix}9 & 8\end{bmatrix}$	$\begin{bmatrix}12 & 12\end{bmatrix}$	$\begin{bmatrix}36 & 36\end{bmatrix}$

You will compute attention for the **last** token in the sequence (*himself*).

4. (2 points) What is the next step in computing self-attention?

   Take the dot product of the query vector (for the token in question) and each of the key vectors (including that for the token in question).

5. (1 point) Perform this step for each token, yielding a vector.

$$\begin{bmatrix}60 & 6 & 18 & 54\end{bmatrix}$$

6. (1 point) Scale each of the resulting values by multiplying them by $\frac{1}{\sqrt{d_k}} \approx 0.701$. Round your answers to the nearest 10th.

$$\begin{bmatrix}42.4 & 4.2 & 12.7 & 38.2\end{bmatrix}$$

7. (2 points) Perform the next step in computing self-attention, converting vector from the preceding step into a vector of probabilities. Present your answers in scientific notation, rounded to one decimal place.

$$\begin{bmatrix} 9.9 \times 10^{-1} & 2.6 \times 10^{-17} & 1.2 \times 10^{-13} & 1.4 \times 10^{-2} \end{bmatrix}$$

8. (2 points) Which token is *himself* attending to most strongly?
Himself attends most strongly to *Elon*.

9. (3 points) What is the next step in computing self-attention?
Multiplying (element-wise) the probabilities by the value vector for each token.

10. (2 points) Carry out this step (hint: the result is a sequence of vectors). Write each vector as a row in the matrix. Express each number in scientific notation and to one decimal place.

$$\begin{bmatrix} 4.4 \times 10 & 4.3 \times 10 \\ 2.3 \times 10^{-16} & 2.1 \times 10^{-16} \\ 1.5 \times 10^{-12} & 1.5 \times 10^{-12} \\ 5.1 \times 10^{-1} & 5.1 \times 10^{-1} \end{bmatrix}$$

11. (3 points) What is the final step?

Compute an element-wise sum of all of the vectors

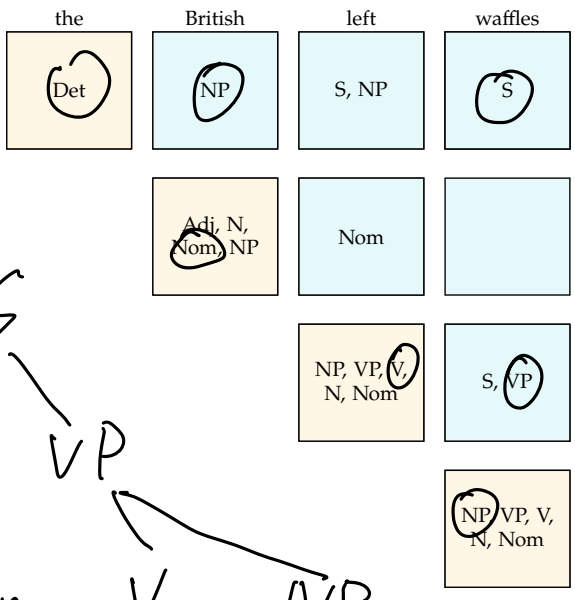
12. (1 point) What is the result? (Hint: the result is a vector.)

$$\begin{bmatrix} 44.9 & 43.9 \end{bmatrix}$$

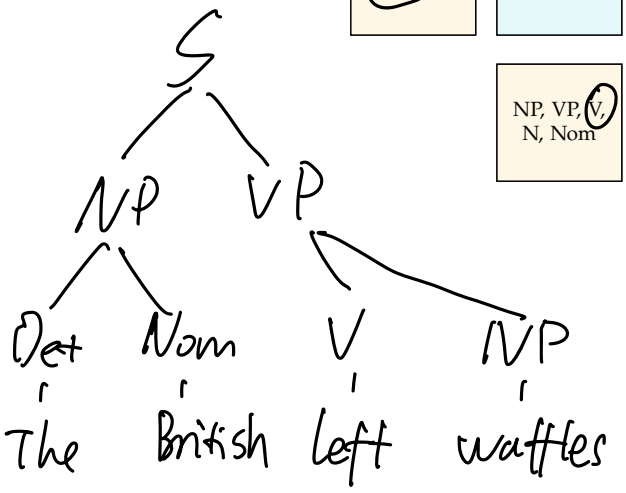
2 Parsing

Consider the ambiguous sentence *The British left waffles*.

1. (6 points) Using the grammar and chart below, draw two phrase structure trees for the sentence. The trees must be derivable given the grammar.




```
N -> British
Nom -> British
NP -> British
Adj -> British
N -> left
Nom -> left
NP -> left
V -> left
VP -> left
Det -> the
N -> waffles
Nom -> waffles
NP -> waffles
V -> waffles
VP -> waffles
NP -> Det Nom
Nom -> Adj Nom
S -> NP VP
VP -> V NP
```

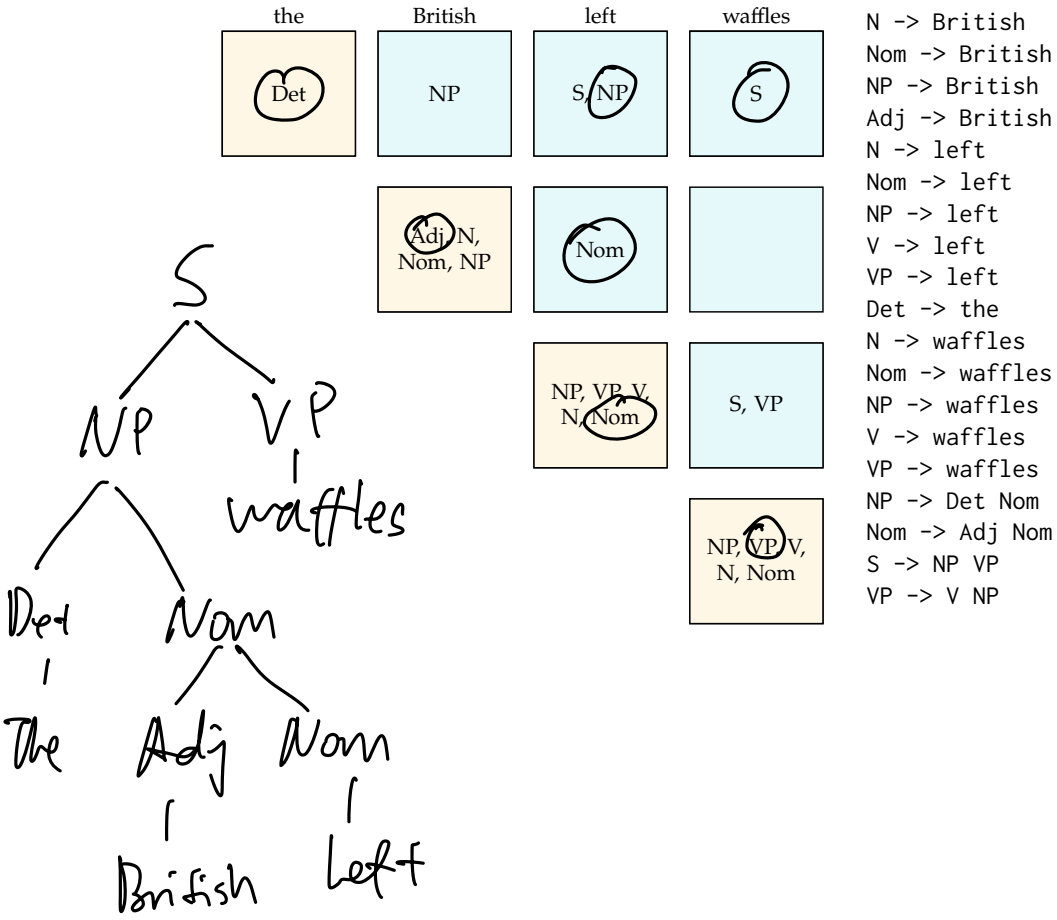


2. (4 points) In your own words describe the two meanings of *The British left waffles*. Note that *waffle* can either mean "a particular type of batter cake" or "vacillate; go back and forth between two opinions or positions."

   The political Left Wing of Britain equivocates. The British left breakfast food behind.

3. (5 points) In the chart, draw a circle around each non-terminal symbol (the letters, not the square) that is used in your first tree.

4. (5 points) In the chart, draw a square around each non-terminal symbol (the letters, not the square) that is used in your second tree. (Hint: a single symbol may be used for more than one tree).




```
N -> British
Nom -> British
NP -> British
Adj -> British
N -> left
Nom -> left
NP -> left
V -> left
VP -> left
Det -> the
N -> waffles
Nom -> waffles
NP -> waffles
V -> waffles
VP -> waffles
NP -> Det Nom
Nom -> Adj Nom
S -> NP VP
VP -> V NP
```

3 Semantics short questions

1. (5 points) What does this Neo-Davidsonian notation say? Write an English sentence.

   $\exists e.\text{OPEN}(e) \wedge \text{AGENT}(e, \text{Alex}) \wedge \text{PATIENT}(e, \text{door}) \wedge$
   $\text{INSTRUMENT}(e, \text{key})$

   Alex opens the door with a key.

2. (10 points) Model-Theoretic Semantics

   Consider the following model:

   ENTITIES: Pat, Sam, Kim, hip_hop, classical, pop

   PREDICATES:

   person: Pat, Sam, Kim
   musical_genre: hip_hop, classical, pop

   RELATIONS:

   like: (Pat, classical), (Pat, hip_hop), (Sam, pop), (Sam, classical), (Sam, hip_hop), (Kim, classical), (Kim, hip_hop)

   favorite_genre: (Pat, classical), (Sam, pop), (Kim, hip_hop)

   Which of the following are true? (fill in the parentheses around (a), (b), and/or (c)).

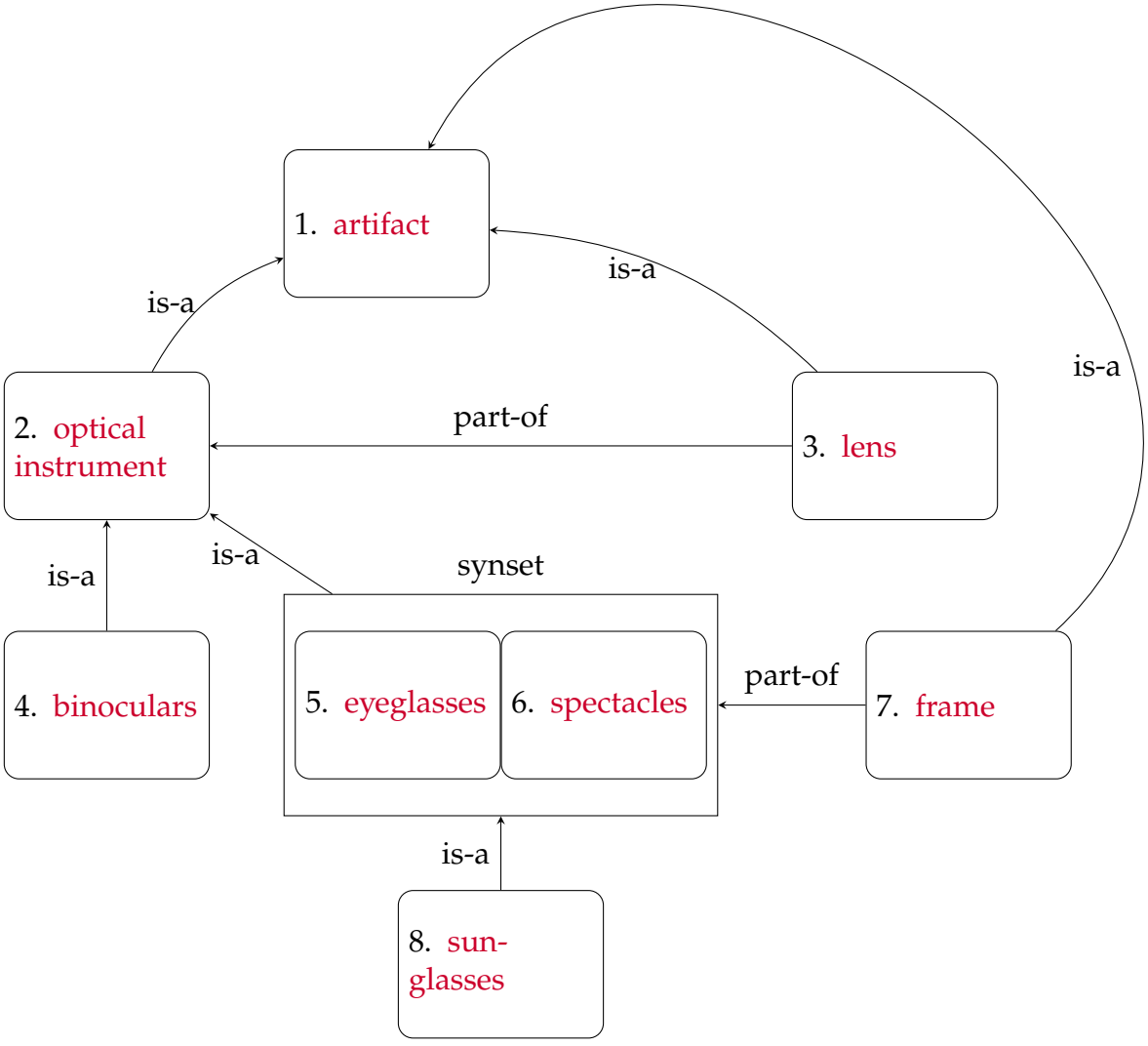
   (a) $\forall x.\text{musical\_genre}(x) \Rightarrow \text{like}(\text{Sam}, x)$ (Correct)

   (b) $\lambda x.\text{like}(\text{Kim}, x)(\text{pop})$

   (c) $\lambda y.\text{favorite\_genre}(y, \text{classical})(\text{Pat})$ (Correct)

3. (10 points) Wordnet

   Your task is to put the following words in a graph with blank nodes (but with IS-A, SYNONYM, and PART-OF links): *eyeglasses*, *spectacles*, *binoculars*, *sunglasses*, *lens*, *frame*, *optical instrument*, *artifact*



   Spectacles are these:

4 Semantics: Long Question

Consider the following passage:

> The computer was propelled by Arvind from the back of the lecture hall to the front row, where it shattered violently into fragments of plastic, glass, and silicon.

1. (10 points) Following are two tables, slightly modified from the PropBank Lexicon that tell you how to label the semantic roles for the two verbs in the sentence.

propel	
ARG0-propel-agent	causer of motion
ARG1-propel-patient	destination of motion
ARG2-propel-goal	ending point of motion
ARG3-propel-source	starting point of motion

shatter	
ARG0-shatter-agent	breaker
ARG1-shatter-patient	thing broken
ARG2-shatter-instrument	tool used to break ARG1
ARG3-shatter-result	pieces

Your job is to (1) identify argument spans: in the sentence below draw square brackets around strings of words that correspond to arguments of *propel* and *shatter*. Not all words will be part of an argument span. (2) Label each span with one of the ARGn labels from the table. Not all labels will be used.

```
[The computer] was propelled by [Arvind]
ARG1-propel                     ARG0-propel
from [the back of the lecture hall] to [the front row],
ARG3-propel                       ARG2-propel
where [it] shattered violently
ARG1-shatter
into [fragments of plastic, glass, and silicon].
ARG3-shatter
```

2. (2 points) What is the name of the task you just did?
   Semantic Role Labeling

3. (3 points) If you modeled this task as an end-to-end sequence labeling task, what would be labeled and what would the labels be? (mark one by filling in the parentheses).

   (a) Verb phrases would be labeled and the labels would be semantic roles

   (b) Tokens would be labeled and the labels would be BIO labels plus semantic roles

   (c) Noun phrases would be labeled and the labels would be constituency parses

   (d) Arcs would be labeled and the labels would be grammatical relations

5 Discourse

What follows is a made-up conversation between a banking customer (U) and an artificially-intelligent customer service agent (A):

A1 Hello! How may I help you today?
U1 I would like to transfer funds from my money market account to my checking account.
A2 How much would you like to transfer from it?
U2 Two hundred dollars.
A3 When would you like to transfer it?
U3 Today.
A4 I'm sorry. Your money market account does not have funds sufficient to cover this transaction.
U4 Oh.
A5 Can I help you with anything else?
U5 Can you tell me the balance on my credit card?
A6 The balance on that account is twenty-five thousand six hundred and twenty-four dollars and sixty-four cents.
U6 Oh.
A7 Can I help you with anything else?
U7 No. [Hangs up.]

Answer the following questions, based on the dialog:

1. (4 points) Identify four coreference chains in the dialog. **Remember that each word in a coreference chain must refer to the same entity in the discourse model**. Write them as sequences of coreferent words or phrases separated by right arrows (→). If more than one turn separates two coreferent words/phrases, separate them with pipes (|). To take a hypothetical example:

   Manling → her → she | Manling → she

   I | I'm | I
   I | me
   my money market account → it | money market account

funds | it
credit card → that account

2. (3 points) Identify three lexical chains in the dialog. Write them as you did the coreference chains.
   transfer → transfer → transfer → (transaction)
   funds → two thousand dollars → funds
   account → account | account | account
   balance → balance

3. (4 points) How many tasks are present in the dialog?

   Two.

4. (9 points) Imagine that you are an algorithm. Using only the information in (1) and (2), justify the the division of the dialog into topics (that is, the location of the boundary(s) between topics).

   The coreference chaines for *money marked account* and *funds* end in A4 and A3 respectively while the chain for *credit care* starts in U5. This suggest a change of topic between A4 and U5. This is confirmed by the *transfer* chain, which ends at A3 and the *account* chain that ends at A4. Together, this suggests a boundary after A4 and before U5.