Name: _____

Dr. Dan Moldovan
Spring 2020
May 4th 2020

Final Exam
CS6320

Instructions:

1. This is an open-book exam.
2. **Do not communicate with anyone during the exam**.
3. If something is not clear to you, write down an assumption you think is reasonable and solve the problem under that assumption.
4. Write your name on the first page and initial all pages.
5. Number all pages.
6. You have 90 minutes for the exam and 15 minutes to scan and upload the exam. Submit only one copy no later than 3:45 pm today.

Problem 1 Regular Expression for NER (10 points)

Write a regular expression for identifying names of doctors in text. Your regular expression should match at least the examples below, but should not recognize either non-names (words not capitalized) or names that do not include the identifying title information (*Dr., Doctor, M.D.)* Do your best to include information about spaces, hyphens, commas and periods, as shown in examples.

William R. Breakey M.D.
Pamela J. Fischer M.D.
Leighton Cluff M.D.
C.M. Franklin, M.D.
Atul Gawande, M.D.
Dr. Talcott
Dr. J. Gordon Melton
Dr. Karl Thombe
Dr. Etienne-Emile Baulieu
Dr. Karl Thomae
Dr. Alan D. Lourie
Doctor Dre
Doctor Dolittle
Doctor William Archibald Spencer
Doctor No

Problem 2 Information Extraction with Bootstrapping (15 points)

Suppose you are building an information extraction system to identify the date, and the city and state in which Governors were born. You want to use bootstrapping to do this.

You know when and where Greg Abbott was born (November 13, 1957, in Wichita Falls, TX) but you don't know about any other governors.

Describe how you could use this information, along with a combination of Google and Wikipedia, to find patterns that could be used in general to determine date and place of birth. Be specific.

Problem 3. Textual Entailment (20 points)

T: Knowledge Representation is the field of AI that focuses on the design of formalisms that are both epistemologically and computationally adequate for expressing knowledge about a particular domain.

H: AI expresses domain specific knowledge.

1. Provide Davidsonian Logic Forms for T and H
2. List the semantic relation triples in the form R(x,y) you identify in T and H
3. Provide a Knowledge Graph Representation of T and H
4. Sketch a proof that T entails H using the Knowledge Graph Representation. Add axioms as necessary to facilitate the proof.

Problem 4 Question Answering (15 points)

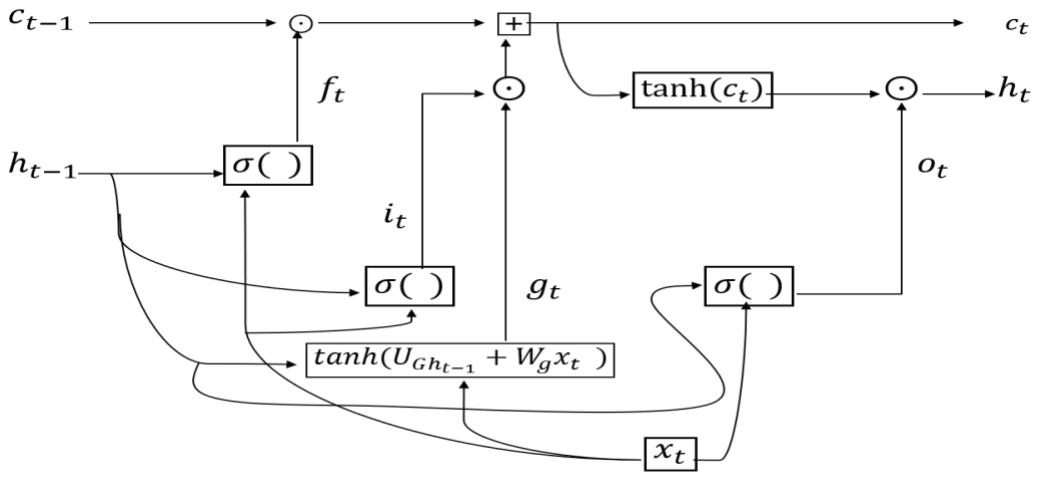
Q: Who is the sponsor of the International Criminal Court?

A: South Africa welcomed the adoption last week of a U.N. – sponsored agreement on the creation of an international criminal court.

Explain how this question may be answered using an enhanced logic representation. Mark Name Entities in text to help with finding the answer.

Problem 5. Neural Network Architecture (15 points)

The figure shows an LSTM cell architecture.



$$g_t = \tanh\left(U_g h_{t-1} + W_g x_t\right)$$
$$i_t = \sigma(U_i h_{t-1} + W_i x_t)$$
$$f_t = \sigma\left(U_f h_{t-1} + W_f x_t\right)$$
$$o_t = \sigma(U_o h_{t-1} + W_o x_t)$$
$$c_t = f_t \odot c_{t-1} + i_t \odot g_t$$
$$h_t = o_t \odot \tanh(c_t)$$

Explain the following:

1. Is it true that when x_t is the 0 vector $h_{t-1} = h_t$?

2. Is it true that if f_t is very small or zero, then error will not be back-propagated to earlier time steps?

3. Can entries f_t , i_t, o_t be viewed as probability distributions (non-negative and they sum to 1) ?

Problem 6 Probabilistic CKY Parsing (15 points)

You are given the following PCFG.

S → NP VP	1.0	Pronoun → I	1.0
NP → Det N	0.5	Verb → ran	1.0
NP → Pronoun	0.2	Noun → race	0.5
NP → NP PP	0.3	Noun → charity	0.5
VP → Verb	0.1	Prep → for	1.0
VP → Verb NP	0.4	Det → a	1.0
VP → VP PP	0.5		
PP → Prep NP	1.0		

Use the probabilistic CKY algorithm to find the most probable parse tree for the sentence

S: *I ran a race for a charity*.

a. Shaw the triangular CKY table with each cell filled with its constituents together with their probabilities. Show how you got those numbers.
b. Show the final parse tree for this sentence that corresponds to your result in a. Show the respective probabilities.

Problem 7 Logistic Regression Classification (10 points)

A logistic regression classifier is used for sentiment classification.

You are given two observation vectors X= $[x_1\ x_2\ x_3]$

$X^1 = [1, 0.5, 0.2]$
$X^2 = [1, 0, -1]$

Assume we learned two weights $w_1\ and\ w_3$ and bias b, but we do not know w_2 yet.

$[w_1\ w_2\ w_3] = [-1, w_{2,}\ 3]$

bias b = − 0.3

What is the range of w_2 such that observation X^1 is less likely to reflect a positive sentiment that X^2.

Hint:

The formula for logistic regression classifier is $P(y = 1) = \frac{1}{1+e^{-(w \cdot x+b)}}$