CS6320, Spring 2019
Dr. Mithun Balakrishna
Homework 6 Solution

1. Coreference Resolution

Apply the Lappin and Leass pronoun resolution algorithm to resolve the pronouns in the following text passage:

Neil Ferguson is prolific, well-paid and a snappy dresser. Stephen Moss hated him, at least until he spent an hour being charmed in the historian's Oxford study.

The following salience weights can be used for the computation:

Sentence recency	100
Subject emphasis	80
Existential emphasis	70
Accusative (direct object) emphasis	50
Indirect object and oblique complement emphasis	40
Non-adverbial emphasis	50
Head noun emphasis	80

Lappin and Leass pronoun resolution algorithm:

1. Collect the potential referents.
2. Remove referents that do not agree in number, gender
3. Remove referents that do not pass intrasentence syntactic constraints.
4. Compute the total salience value.
5. Select referent with highest salience value.

Example

Niall Ferguson is prolific, well-paid and a snappy dresser. Stephen Moss hated **him**, *at least until he spent an hour being charmed in the historian's Oxford study.*

- Discourse Referents:
  - $N_0 = \{$*Niall Ferguson*$\} = 105$
    (subject 80+ head-noun emphasis 80 + non-adverbial 50)/2
  - $S_0 = \{$*Stephen Moss*$\}$ ***does not pass syntax filter***
- New Discourse referents
  - Add *him* to N_0; $N_1 = \{$*Niall Ferguson, him*$\}$

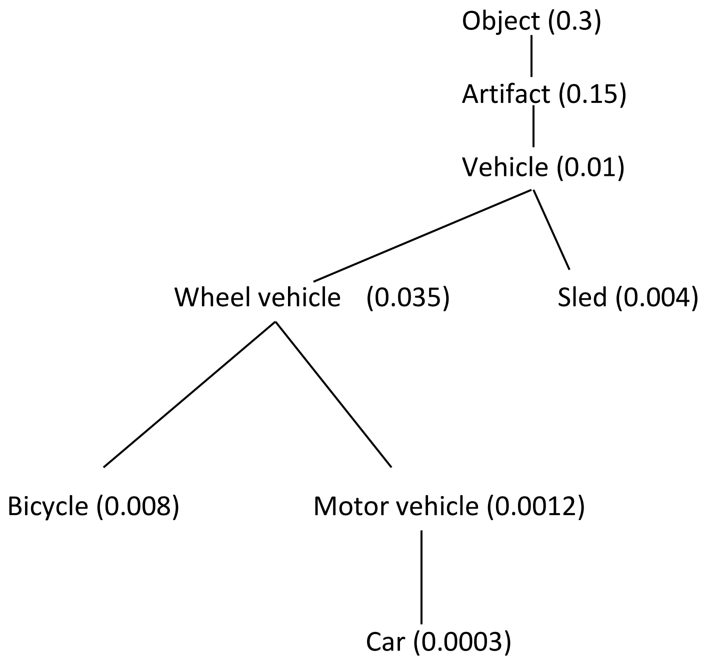
Example

Niall Ferguson is prolific, well-paid and a snappy dresser. Stephen Moss hated him, at least until **he** *spent an hour being charmed in the historian's Oxford study.*

- Discourse Referents:
  - $N_1 = \{$*Niall Ferguson, him*$\} = 405$
    (subject 80+ head-noun emphasis 80 + non-adverbial 50)/2 + direct object 70+ head-noun emphasis 80 + non-adverbial 50 + recency 100
  - $S_1 = \{$*Stephen Moss*$\} = 310$
    subject 80 + head-noun emphasis 80 + non-adverbial 50 + recency 100
- New Discourse referents
  - Add *he* to N_1; $N_2 = \{$*Niall Ferguson, him, he*$\}$

Please note that *he* was mapped incorrectly to Niall Ferguson.

2. Word Similarity



The number near each concept *c* denotes *P(c)*: probability that a random word in a corpus is an instance of that concept *c*.

a.) Using $sim_{path}\,(c_1, c_2) = \frac{1}{pathlength\,(c_1,c_2)}$, which pair of words is more similar? Explain your answer.

(*bicycle, sled*) or (*car, artifact*)

Solution:

$\text{sim}_{\text{path}}(bicycle, sled) = 1/3$

$\text{sim}_{\text{path}}(car, artifact) = 1/4$

$\text{sim}_{\text{path}}(bicycle, sled) > \text{sim}_{\text{path}}(car, artifact)$, hence (*bicycle, sled*) is more similar than (*car, artifact*)

b.) Using $\text{sim}_{\text{Lin}}(c_1, c_2) = \frac{2 \text{ x} \log P(\text{LCS}(c_1,c_2))}{\log P(c_1) + \log P(c_2)}$, which pair of words is more similar? Explain your answer.

(*bicycle, sled*) or (*car, artifact*)

Solution:

sim_{Lin}(*bicycle, sled*) = (2*log10(0.01))/(log10(0.008)+log10(0.004)) = 0.8899073341
sim_{Lin}(*car, artifact*) = (2*log10(0.15))/(log10(0.0003)+log10(0.15)) = 0.37908857681

sim_{Lin}(*bicycle, sled*) > sim_{Lin}(*car, artifact*) , hence (*bicycle, sled*) is more similar than (*car, artifact*)

3. Word Association and Similarity

Using the feature values listed in the table below, compute the similarity between (apple, IOS) and (apple, orange) using Jaccard similarity and probability association. Which word pair is more similar?

	f1	f2	f3
apple	5	3	2
IOS	3	6	4
orange	0	4	10

$$\text{assoc}_{\text{prob}}(w, f) = P(f|w)$$

$$\text{sim}_{\text{Jaccard}}(\vec{v}, \vec{w}) = \frac{\sum_{i=1}^{N} \min(v_i, w_i)}{\sum_{i=1}^{N} \max(v_i, w_i)}$$

Solution:

$$P(f|w) = \frac{\text{count}(f, w)}{\text{count}(w)}$$

	P(f1|w)	P(f2|w)	P(f3|w)
apple	0.5	0.3	0.2
IOS	0.23	0.46	0.31
orange	0	0.29	0.71

$$\mathrm{sim}_{\mathrm{Jaccard}}\left(\overrightarrow{apple}, \overrightarrow{IOS}\right) = \frac{0.23+0.3+0.2}{0.5+0.46+0.31} = 0.57$$

$$\mathrm{sim}_{\mathrm{Jaccard}}\left(\overrightarrow{apple}, \overrightarrow{orange}\right) = \frac{0+0.29+0.2}{0.5+0.3+0.71} = 0.32$$

Based on the above scores computed using a domain feature vector, the word pair (apple, IOS) is more similar than (apple, orange)

4. Logic and Semantic Representation (25 points)

Bill Gates, the founder of Microsoft, generously donates money to charities every year.

a. Provide a Davidsonian logic representative of this sentence.

Solution:

bill_NN(x1) & gates_NN(x2) & nn_NNC(x3,x1,x2) & _human_NE(x3) & founder_NN(x4) & of_IN(x4,x5) & Microsoft_NN(x5) & _organization_NE(x5) & generously_RB(e1) & donate_VB(e1,x3,x6,x7) & money_NN(x6) & charity_NN(x7) & every_JJ(x8) & year_NN(x8)

b. Identify the semantic relations in the sentence. Write them as semantic triples R(x,y). Consider only the sematic relations in the list below.

Semantic Relation	Definition
agent(X,Y)	X is the agent for Y
beneficiary(X,Y)	X is a beneficiary of Y
cause(X,Y)	X causes Y
instrument(X,Y)	X is an instrument in Y
justification(X,Y)	X is the reason/motive/justification for Y
location(X,Y)	X is the location of Y or where Y take place
manner(X,Y)	X is the manner in which Y happens
part-whole(X,Y)	X is a part of Y
quantity(X,Y)	X is a quantity of Y; Y can be an entity or event
result(X,Y)	X is the result consumed in/from/of Y
synonymy(X,Y)	X is a synonym/name/equal for/to Y
theme(X,Y)	X is the theme consumed in/from/of Y
time(X,Y)	X is the time of Y (when Y take place)
value(X,Y)	X is a value of Y

Solution:

synonymy(Bill Gates, founder)

part-whole(founder, Microsoft)

manner(generously, donate)

agent(Bill Gates, donate)

theme(money, donate)

beneficiary(charity, money)

time(year, donate)

value(every, year)

c. Provide a new logic representation that includes semantic relations.

Solution:

bill_NN(x1) & gates_NN(x2) & nn_NNC(x3,x1,x2) & _human_NE(x3) & founder_NN(x4) & Microsoft_NN(x5) & _organization_NE(x5) & generously_RB(x6) & donate_VB(e1) & money_NN(x7) & charity_NN(x8) & every_JJ(x9) & year_NN(x10) & synonymy_SR(x3,x4) & part-whole_SR(x4,x5) & manner_SR(x6,e1) & theme_SR(x7,e1) & & beneficiary_SR(x8,x7) & time_SR(x10,e1) & value_SR(x9,x10)

d. Using the basic semantic relations that you identified for Question 4.b, write the semantic calculus rule to create a new semantic relation:

Donates-To(*X*, *Y*): *X* donates to *Y*

Solution:

agent(X,A), theme(Z,A), beneficiary(Y,Z) → Donates-To(X,Y) [A is donate verb/concept]