Natural Language Processing

Assignment- 8

TYPE OF QUESTION: MCQ

Number of Questions: 9 [Question 4 carries 2 marks] **Total Marks: 10**

Question 1:

Consider the following statements. Which of the following is/are True?

1. Car is a hyponym of vehicle.
2. Flower is a hypernym of rose.
3. Vehicle is a hyponym of car.
4. Tulip is a hypernym of flower.

Answer: 1, 2

Question 2:

Which of the following is False?

1. Hypernym: From concepts to superordinates
2. Hyponym: From concepts to subtypes
3. Troponym: From verbs to the verbs they entail
4. Part Meronym: From wholes to parts

Answer: 3

Solution: Refer to Week 8 Lecture 37

In reality, troponyms describe a way of doing the action of the main verb, rather than verbs entailing other verbs. For example, "to sprint" is a troponym of "to run" because sprinting is a specific way of running. Let me know if you'd like further clarification!

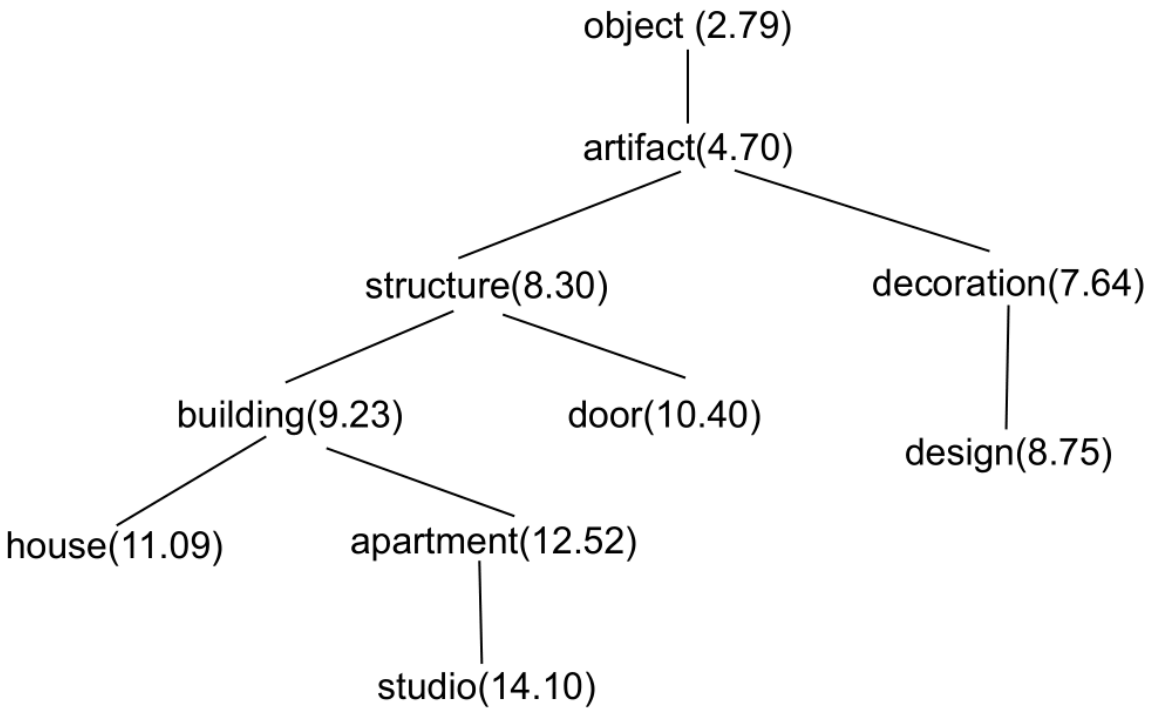
Question 3:
Two concepts along with their glosses are given below. Find the similarity score between concepts “book” and “novel” using the Extended Lesk’s algorithm. (Note: Do not consider the stop words.)

book: a set of written or printed pages bound together
novel: a long written work of fiction bound in pages

1. 2
2. 3
3. 5
4. 8

Answer: 2
Solution:
Common words are: **written, pages, bound**
Score = 1^2 + 1^2 + 1^2 =3

For Question 4 to 6, consider a hypothetical wordnet noun taxonomy with their information content as shown in Figure 1. Question 4 carries 2 marks
Note: Use base 10 in logarithmic calculations



Question 4:

What is the Lin similarity between **house** and **design**?

a. 0.564
b. 0.433
c. 0.466
d. 0.473

SimLin(c1,c2) = (2*IC(LCS(c1,c2))) / (IC(c1) + IC(c2))

Answer: d

Solution: (2×4.7)/(11.09+8.75) ≈ 0.473

Question 5:

What is the Resnic similarity between **building** and **door**?

a. 11.09

b. 8.30

c. 9.23

d. 4.70

Answer: b SimResnik(c1,c2) = IC(LCS(c1,c2))

Solution:

Question 6:

What is the Leacock–Chodorow similarity between **building** and **design**?

a. 0.398

b. 0.699 d is max depth of tree

c. 0.097

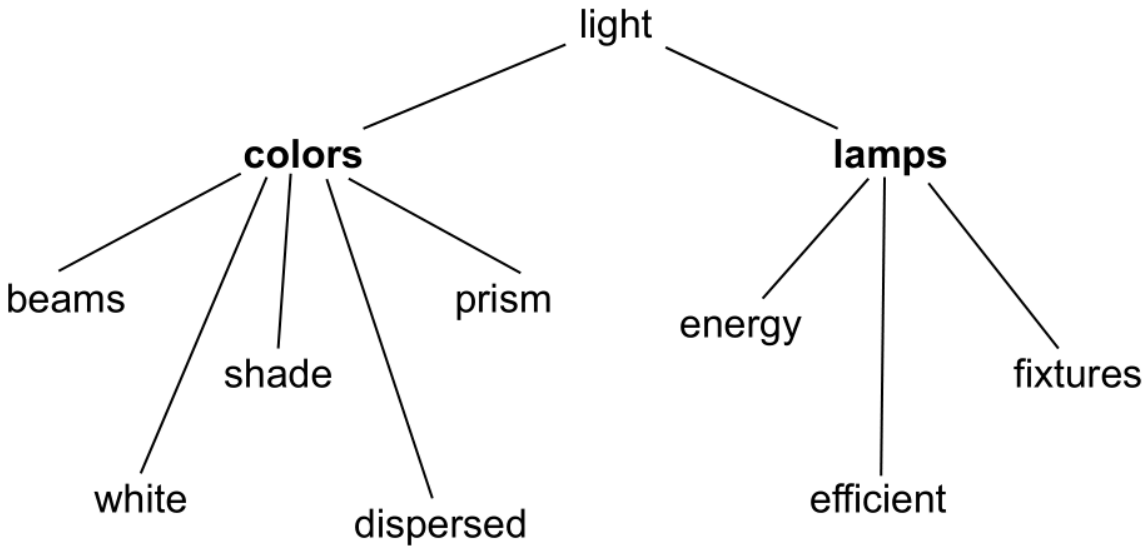
d. None of the above

Answer: a

Solution:

LC similarity = −log pathlen(c1,c2)/2d = −log 4/(2×5) ≈ 0.398

For Question 7 to 9 consider the network of words for disambiguation of the word "light" as shown in Figure 3. The hubs are "colors" and "lamps". Note: Take the distance between two words as the path length between them.



Question 7:
Compute the scores for (i) the hub "colors" and the component "white" and (ii) the hub "colors" and the component "fixtures".

a. 0.2, 0.25
b. 1.0, 0.0
c. 0.5, 0.25
d. None of the above

Score = 1/(1+pathlenFromHubToComponent)

Answer: d

Solution:
(i) 1/(1+1) = 0.5
(ii) 0 as "colors" is not an ancestor of "fixtures"

Question 8:

What are the scores of the hubs “colors” and “lamps” respectively?

a. 0.6, 0.4
b. 0.20, 0.33
c. 2.5, 1.5
d. None of the above

ComponentScore = 1/(1+PathLength)
Calc for Each Component the Score.
Sum of Each ComponentScore for Both Branches

(0.5)5, (0.5)3
2.5, 1.5

Answer: c

Solution: Each component’s score is 0.5

Question 9:

Which is the most appropriate sense for the word “light”?

a. colors 0.5*5 = 2.5
b. lamps 0.5*3 = 1.5
c. both colors and lamps are appropriate
d. Not enough data

Answer: a

Solution: “colors” has the highest score
