COMS W4705: Natural Language Processing
Written Homework 4

Sample Solutions

December 17, 2018

Problem 1

Euclidean distance:
$dist_{eucl}(animal, dog) = 3.3166$
$dist_{eucl}(animal, cat) = 8.4261$
$dist_{eucl}(animal, computer) = 4.5826$
$dist_{eucl}(animal, run) = 6.9282$
$dist_{eucl}(animal, mouse) = 9.6437$
Therefore *dog* is the most similar word to *animal*.
Cosine similarity:
$sim_{cos}(animal, dog) = 0.8519$
$sim_{cos}(animal, cat) = 0.7292$
$sim_{cos}(animal, computer) = 0.7620$
$sim_{cos}(animal, run) = 0.6014$
$sim_{cos}(animal, mouse) = 0.6658$
dog is still the most similar word to *animal*.

Problem 2

There are multiple possible solutions to this problem. Here is one: We assume the two senses a and b are given as WordNet synsets. The lexical relations in WordNet (hypernyms, meronyms, etc.) form a graph. We can campute the shortest path between a and b in the graph, i.e. the shortest number of lexical-relations one needs to follow to go from a to b. Polysemes will have a relatively short distance to each other, while homonyms (that are mostly unrelated) will have a large distance.

Problem 3

a)

- ⟦every dog⟧(⟦eats a bone⟧)
  - ⟦every⟧(⟦dog⟧)
    - ⟦every⟧
    - ⟦dog ⟧
  - ⟦eats ⟧(⟦a bone ⟧)
    - ⟦eats ⟧
    - ⟦a ⟧(⟦bone ⟧))
      - ⟦a ⟧
      - ⟦bone ⟧

$[\![$a bone$]\!] = [\![$a$]\!]([\![$bone$]\!]) =$

$$(\lambda R.\lambda S.\exists y R(y) \land S(y))(\lambda y.bone(y)) =$$

$$\lambda S.\exists y(\lambda y.bone(y))(y) \land S(y) = \lambda S.\exists y\ bone(y) \land S(y)$$

$[\![$every dog$]\!] = [\![$every$]\!]([\![$dog$]\!]) =$

$$(\lambda P.\lambda Q.\forall x P(X) \rightarrow Q(x))(\lambda x.dog(x)) =$$

$$\lambda Q.\forall x(\lambda x.dog(x))(x) \rightarrow Q(x) =$$

$$\lambda Q.\forall x\ dog(x) \rightarrow Q(x)$$

$[\![$eats a bone$]\!] = [\![$eats$]\!]([\![$a bone$]\!]) =$

$$(\lambda T.\lambda x.T(\lambda z.eats(x, z)))(\lambda S.\exists y\ bone(y) \land S(y))$$

$$\lambda x.(\lambda S.\exists y\ bone(y) \land S(y))(\lambda z.eats(x, z))) =$$

$$\lambda x.\exists y\ bone(y) \land (\lambda z.eats(x, z))(y) =$$

$$\lambda x.\exists y\ bone(y) \land eats(x, y)$$

$[\![$every dog eats a bone$]\!] = [\![$every dog$]\!]([\![$eats a bone$]\!]) =$

$$(\lambda Q.\forall x\ dog(x) \rightarrow Q(x))(\lambda x.\exists y\ bone(y) \land eats(x, y))$$

$$\forall x\ dog(x) \rightarrow (\lambda x.\exists y\ bone(y) \land eats(x, y))(x)) =$$

$$\forall x\ dog(x) \rightarrow \exists y\ bone(y) \land eats(y, x))$$

b)

$$\forall x\ dog(x) \rightarrow \exists y\ bone(y) \land eats(y, x))$$

$$\exists y\ bone(y) \land \forall x\ dog(x) \rightarrow eats(y, x))$$

c) While one could represent quantifiers in AMR, for example using a ":mod every" edge, AMR cannot represent quantifier scope. The reason for this is that the edges only express local semantic roles for individual concepts, not for entire phrases. As a result, there is no way to distinguish between the two readings from part b). Both interpretations would have the same graph representation.