Homework 3

Haukur Páll Jónsson
NLP 2017

December 10, 2017

Question 1. CCG parsing

Answer a): Each rule application is given with
|word part 1 + word part 2| resulting category | rule applied
The ordering of rule application is:
1. |two + units| N | Forward appliaction
2. |the + two units| NP | Forward appliaction
3. |to + the two units| PP | Forward appliaction
4. |four + Boeign-747s| NP | Forward appliaction
5. |The + company| NP | Forward appliaction
6. |added + four Boeign-747s| (S \\ NP)/PP | Forward appliaction
7. |added four Boeign-747s + to the two units| S \\ NP | Forward appliaction
8. |in + 1994| (S NP)\\(S \\ NP) | Forward appliaction
9. |added four Boeign-747s to the two units + in 1994| (S \\ NP) | Backward appliaction
10. |The company + added four Boeign-747s to the two units in 1994| S | Backward appliaction

Question 2. Max-ent models

Answer a): A maximum entropy model predicts the most likely y given $\boldsymbol{x}$. For each y we compute:

$$P(y|\boldsymbol{x}) = \frac{e^{w \cdot f(\boldsymbol{x},y)}}{Z}$$

Where $Z = \sum_{y'} w \cdot f(\boldsymbol{x}, y')$ is a normalizing constant. We can take the log of this expression and still get the same prediction as the logarithm is monotonically increasing:

$$\log P(y|\boldsymbol{x}) = \frac{w \cdot f(\boldsymbol{x}, y)}{\log Z}$$

$$\log P(y|\boldsymbol{x}) = \sum_{i=1}^{k} w_i * f_i(\boldsymbol{x}, y) - \log Z$$

The maximum entropy model is called a log linear model since if you take the logarithm of the expression, the model is simply a linear combination of its parameters.

Answer b): Now we compute $\sum_{i=1}^{k} w_i * f_i(\boldsymbol{x}, y)$ for each y. Note that *animal* does not appear in the text ($\boldsymbol{x}$) so f_4 to f_6 are always zero.

$$\sum_{i=1}^{9} w_i * f_i(\boldsymbol{x}, 1) = 2 - 0.1 = 1.9$$

$$\sum_{i=1}^{9} w_i * f_i(\boldsymbol{x}, 2) = 1.8 + 1.1 = 2.9$$

$$\sum_{i=1}^{9} w_i * f_i(\boldsymbol{x}, 1) = 0.3 + 2.7 = 3.0$$

$$Z = \log\left(e^{1.9} + e^{2.9} + e^{3}\right) = 3.8054$$

Thus,

$$P(y = 1|\boldsymbol{x}) = 1.9 - 3.8054 = -1.9054$$

$$P(y = 2|\boldsymbol{x}) = 2.9 - 3.8054 = -0.9054$$

$$P(y = 3|\boldsymbol{x}) = 3.0 - 3.8054 = -0.8054$$

We can verify that this is indeed correct: $e^{-1.9054} + e^{-0.9054} + e^{-0.8054} = 1.0$. The most probable sense is $y = 3$, a noun.

Question 3. First order logic to natural language

Answer a): All bears are furry

Answer b): Sergii is eating a pizza with a fork

Answer c): (All) the students lift Marie

Answer d): The students lifted Marie

Question 4. Natural language to first order logic

Answer a): Juan hates pasta

$$\exists\, e, x\; hating(e) \wedge hater(e, Juan) \wedge hated(e, x) \wedge pasta(x)$$

Answer b): Some student likes every class

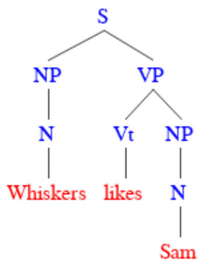
This sentence is ambigious, it can mean "A single student likes each and every class" or "Each student likes some class but possibly different classes". The first sentence:

$$\exists\, s.\, student(s) \wedge \forall\, c.\, class(c) \implies likes(s, c)$$

And the other sentence:

$$\forall\, s.\, student(s) \implies \exists\, c.\, class(c) \wedge likes(s, c)$$

Question 5. Semantic attachments



Answer a): The parse for the sentence is given by the figure. Then we compute the semantics of the sentence:

$$
\begin{aligned}
S.sem &= VP.sem(NP.sem) \\
&= VP.sem(N.sem) \\
&= VP.sem(Whiskers) \\
&= Vt.sem(NP.sem)(Whiskers) \\
&= Vt.sem(N.sem)(Whiskers) \\
&= Vt.sem(Sam)(Whiskers) \\
&= \lambda x.\lambda y.\exists e.liking(e) \wedge liker(e,y) \wedge likee(e,x)(Sam)(Whiskers) \\
&= \lambda y.\exists e.liking(e) \wedge liker(e,y) \wedge likee(e,Sam)(Whiskers) \\
&= \exists e.liking(e) \wedge liker(e,Whiskers) \wedge likee(e,Sam)
\end{aligned}
$$