Coreference Resolution Exercises

A) Coreference resolution

1. Given the following text:
   John Smith and his assistant will go to Barcelona with their boss,
   Dan Farrow. She is now preparing the meeting.
   a) Identify all the mentions for the problem of identity noun phrase coreference.
   b) List the most informative positive and negative examples to learn a mention-pair model
       biased to select the closest antecedent.
   c) Assume we have already learned a mention-pair model. Consider m1, m2, m3, m4, m5,
       m6, m7, m8, m9 as the sequence of mentions ordered by occurrence in the text. Given
       the following probabilities for mention pairs, provide the m7 antecedent that results
       from applying closest-first strategy and from applying best-first strategy, assuming a
       coreference threshold of 0.6.

       \[
       P(\text{CO}|<m1,m7>)=0.8; \quad P(\text{CO}|<m2,m7>)=0.6; \quad P(\text{CO}|<m3,m7>)=0.4; \quad P(\text{CO}|<m4,m7>)=0.6;
       \]
       \[
       P(\text{CO}|<m5,m7>)=0.7; \quad P(\text{CO}|<m6,m7>)=0.6; \quad P(\text{CO}|<m7,m8>)=0.9; \quad P(\text{CO}|<m7,m9>)=0.5;
       \]

2. Provide the first iteration trace of the constraint relaxation algorithm explained in class to
   the following graph of mentions. Consider a scale factor of 5. Concretely, specify the
   normalized values achieved for all the label supports and the new probabilities for all the
   labels at the first iteration.

   ![Graph of mentions]

   - h1 = 1
   - h1 = 0.33
   - h2 = 0.67
   - h1 = 0.33
   - h2 = 0.33
   - h3 = 0.34
   - h1 = 0.25
   - h2 = 0.25
   - h3 = 0.25
   - h4 = 0.25

B) Active learning for coreference resolution

Consider three initial mention-pair ME models (M1,M2,M3) to apply active learning. The
following table shows probabilities of mention pairs xi to co-refer, P(CO|xs,Mk) assigned by ead
model Mk to a set of mention pairs xi=<mi,mj>:

<table>
<thead>
<tr>
<th></th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
</tr>
</thead>
<tbody>
<tr>
<td>x1</td>
<td>0.7</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>x2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>x3</td>
<td>0.8</td>
<td>0.75</td>
<td>0.85</td>
</tr>
<tr>
<td>x4</td>
<td>0.7</td>
<td>0.75</td>
<td>0.6</td>
</tr>
<tr>
<td>x5</td>
<td>0.7</td>
<td>0.5</td>
<td>0.1</td>
</tr>
</tbody>
</table>
Questions:

1. Apply uncertainty least confident sampling, uncertainty margin sampling and query-by-committee vote entropy sampling for each possible combination of models and examples. What examples are selected by each of the 9 frameworks?

2. Are all these query strategies appropriated for this problem? Why?