Please answer any THREE Questions from the FIVE questions provided

This paper will be taken on-line and this is the paper format which will be available as a back-up

The use of electronic calculators is NOT permitted
1. Here are two structurally ambiguous sentences:

   i) Chris bought some apples with red skin.
   ii) Chris gave the apples to his brother.

Here is a toy probabilistic grammar:

\[
\begin{align*}
S &\rightarrow NP \ VP & 1.0 \\
NP &\rightarrow Det \ N & 0.5 \\
NP &\rightarrow Adj \ N & 0.3 \\
NP &\rightarrow PN & 0.2 \\
N &\rightarrow N \ PP & 0.3 \\
VP &\rightarrow V \ NP & 0.8 \\
VP &\rightarrow V \ NP \ PP & 0.2 \\
PP &\rightarrow P \ NP & 1.0
\end{align*}
\]

a) How does one calculate the probability of a complete sentence structure, given the probabilities of the individual rules? (2 marks)

b) Which of the two possible structures will be assigned to these sentences by this grammar? What is the one critical choice of rule involved? (3 marks)

c) What additional information could you build into the rules to get a more accurate result? (5 marks)

d) Discuss the advantages and disadvantages of “shallow parsing” as opposed to “deep parsing”. Give examples. (10 marks)

2. Natural Language Engineering systems typically are built from a number of separate modules. Name two modules in which domain-specific systems differ significantly from each other and from open-domain systems. Discuss the differences, and the problems and opportunities they present, with specific examples. (NB while two, very different examples have been emphasised in the course, there is a case to be made for others if you wish.)

   a) i) Name of module (2 marks)
      ii) Discussion (8 marks)

   b) i) Name of module (2 marks)
      ii) Discussion (8 marks)
3. In the assessed coursework for this course you tested a part of speech tagger, a shallow parser, and WordNet. Based on your experience, how well would the tagger and parser together, and Word Net, deal with each of these two texts?

a) Then this ebony bird beguiling my sad fancy into smiling,
   By the grave and stern decorum of the countenance it wore,
   'Though thy crest be shorn and shaven, thou,' I said, 'art sure no craven.
   Ghastly grim and ancient raven wandering from the nightly shore -
   Tell me what thy lordly name is on the Night's Plutonian shore!'
   Quoth the raven, 'Nevermore.'
   (10 marks, 5 + 5)

b) Paul Tierney grabbed the only goal of the game as Stockport edged out local rivals Macclesfield in the Johnstone's Paint Trophy second round tie.

   Tierney surged forward, played a neat one-two with Leslie Thompson and shot home after keeper Jonathan Brain saved his initial effort.

   Stockport dominated the early stages with Thompson spurning three chances.

   But Macclesfield also had their chances and Francis Green saw a spectacular volley bounce back off the upright.
   (10 marks, 5 + 5)

4. Of the applications we have studied in the course, choose one which you believe is already sufficiently well developed to be genuinely useful, or soon will be.

a) Which application have you chosen? (1 mark)

b) What are the most significant achievements so far in this area? (5 marks)

c) What are the most significant remaining problems or obstacles? (5 marks)

d) Why did you choose this particular application? Contrast it, in detail, with one or two others which you believe are less promising, to justify your choice. (9 marks)
5. Here is a short English text, and the result of using BabelFish to translate it into German and then back to English:

Dormice are small rodents with soft, orange-brown fur and long tails of a similar colour which are furred all over. They are such attractive animals that many people would like to be able to watch them but because they are almost entirely nocturnal this presents a challenge to observers.

Dormice are completely more rüber small rodents with soft, orange-brown fur and long end pieces of a similar color, which are, furred. They are such attractive animals, which would like many people that are, to watch out them but, because they are nearly completely nocturnal this a challenge observers represent.

The main chance of seeing one would be at twilight when many changes are taking place. For almost all birds it is time to go to roost and to finish the day's activities but for many animals it is the reverse: they are beginning their main spell of activity.

The main probability of seeing would be at the dawn, if many changes take place. For nearly all birds it is time to go to roost and to terminate the activities of the daily, but for many animals it is the back: they begin their main spell of the activity.

In Britain this small mammal is known as the 'common dormouse' but a more appropriate name would be the hazel dormouse. The animal's scientific name, as above, shows the deep connection with the hazel tree whose scientific name is (Corylus avellana). Dormouse in German is Haselmaus another connection with hazel, and in French it is Muscardin. The fact that the dormouse hibernates is reflected in some of the local English names. In the counties of Hampshire and Cornwall it is known as 'dory mouse' and 'dozing mouse' respectively: in many counties it is known as the 'sleeper', the 'seven sleeper', or 'sleep mouse'. It's attractive appearance has a fairy tale charm and other delightful local names include 'chestle crumb' and 'derry mouse'.

In Great Britain this small mammal admits as the 'common Dormouse ', but a more suitable name would be the Haselnussdormouse. The scientific name of the animal, like above, shows the deep connection with the Haselnussbaum its scientific Namensis (Corylus avellana). Dormouse on German is Haselmaus another connection with haselnuss, and on French it is Muscardin. The fact that the Dormouse hibernates is reflected in some the local English names. In the counties of Hampshire and of Cornwall it admits as ' rowing boat mouse ' and a ' doesende mouse ' and/or: in many counties it admits sleep mouse ' like the ' camp threshold ' , the ' camp threshold sieves ' or of '. It is attractive appearance lets a charm fairy of history and other wonderful local name ' chestle krumem ' and ' derry mouse ' also include.

Make the best judgement you can, the marks are for intelligent thinking not for “accuracy”. Where possible, explain exactly what you think has happened. Concentrate on the most interesting mistakes.

a) Which errors are just details of idiom or minor word variations, and do not affect our understanding of the output? (5 marks)

b) Of the serious errors, which do you think are due to mistakes in the translation of individual words? (6 marks)

c) Of the serious errors, which do you think are due to mistakes in parsing? (4 marks)

d) Of the serious errors, which do you think make no sense? - you need to explain why in each case. (If you believe that seeing the German intermediate text would help, then say so.) (5 marks)

END OF EXAMINATION