UNIVERSITY OF OSLO

Faculty of Mathematics and Natural Sciences

Exam in INF5820 Language technological applications
Day of exam: 18 December
Exam hours: 0900-1300
This examination paper consists of 7 page(s).
Appendices: 0
Permitted materials: None

Make sure that your copy of this examination paper is complete before answering.

- **You may answer in English, Norwegian, Danish or Swedish.**
- **You should answer all questions. The weight of the various questions are indicated.**
- **You should read through the whole set to see whether anything is unclear so that you can ask your questions to the teachers when they arrive.**
- **If you think some assumptions are missing, make your own and explain them!**

1 Analysis of spoken dialogue (20%)

On the next page, you will find a transcript from a real human-human conversation recorded in the spoken part of the BNC corpus. You are asked to answer the following questions based on this transcript:

1. How are the dialogue turns structured, based on the observed linguistic cues?
2. What kind of speech acts are used through the dialogue, according to Searle's taxonomy?
3. What are the grounding signals & strategies used through the dialogue?
4. Can you find some examples of conversational implicatures?
5. List a few (2-3) deictic markers occurring in the dialogue.
6. Finally, find one example of disfluency in the transcript, and structure them based on Shriberg's disfluency model.

Please be brief in your answers: a few lines of explanation should be sufficient for each question. You can use the line numbers on the right to refer to specific parts of the transcript.

Transcript 1.1 Dialogue transcript from the British National Corpus

CONVERSATIONAL CONTEXT: [*John and Mary are sorting things for a car boot sale.*]

JOHN:	I really wouldn't mind selling all my camera gear.	1
	... Su and er it's su such a waste.	2
	I never use it.	3
MARY:	It is isn't it?	4
	You, you've had no pleasure really out of it.	5
JOHN:	No.	6
	I mean if I sold it cheap.	7
MARY:	Eh?	8
JOHN:	Even if I only sold it, you know for a couple of hundred pounds. ...	9
MARY:	Well you could buy yourself some ... golf stuff for a hundred couldn't you?	10
	Or	11
JOHN:	Well it's worth more than a hundred.	12
MARY:	Well	13
JOHN:	I couldn't sell it	14
MARY:	you c you could maybe I dunno	15
JOHN:	that cheap.	16
	What do you want Rick?	17
	If you're not gonna help, clear off.	18
	Rick being the dog. ...	19
MARY:	You wanna keep that don't you?	20
JOHN:	Oh yeah that, that will be handy.	21
MARY:	Yeah.	22
	... And the there's your writing set as well.	23
JOHN:	Yeah.	24
	I'll I'll keep the writing set.	25
MARY:	I think yeah.	26
	I mean that's lovely.	27
	Take that downstairs.	28
	You can write to Mark.	29
	Then you've got yours and I've got mine then.	30

(http://www.phon.ox.ac.uk/SpokenBNCdata/KCL.html)

2 Probabilistic modelling (20%)

Imagine you are building a multi-modal, multi-party dialogue system, where users can talk about other persons present in the room. This reference can be done either linguistically (using the person's name), but also via pointing gestures. Your task is to implement a simple reference resolution engine that uses both linguistic and visual cues, using a Bayesian Network to represent the problem.

More specifically, assume the following:

- you have 3 persons in the room: *Jan Tore*, *Stephan*, and *Pierre*;
- A speaker never refers to himself, and can refer to the two other persons with equal probability;
- We have a linguistic observation $SpeakerMentions$ with 4 distinct values: $JanTore$, $Stephan$, $Pierre$, and $None$;
- We also have a gestural observation $SpeakerPointsTo$ also with 4 distinct values: $JanTore$, $Stephan$, $Pierre$, and $None$;
- If the speaker intends to refer to a particular person X, there is a probability 0.7 that $SpeakerMentions$ will be equal to X, a probability 0.2 that it will be equal to $None$, and a probability 0.1 that it will be equal to another person (i.e. 0.05 for each of the two other persons).
- if the speaker intends to refer to a particular person X, there is a probability 0.4 that $SpeakerPointsTo$ will be equal to X, a probability 0.3 that it will be equal to $None$, and a probability 0.3 that is will be equal to another person (i.e. 0.15 for each of the two others).

Based on this information, you are asked to:

1. Draw the Bayesian network and the associated probability distributions that correspond to the outlined specifications;
2. Compute the probability

$$P(Reference=? \mid Speaker=Pierre,\\ SpeakerMentions=JanTore,\\ SpeakerPointsTo=Stephan)$$

(please show your calculation steps)

3 Natural language generation (10%)

You want to build a small module for generating referring expressions about a set of objects on a table, using Dale and Reiter's *Incremental algorithm*. For each object, you have access to two of its properties: (1) its type (e.g. a mug or a plate) and (2) its colour (e.g. blue or red).

Show how Dale and Reiter's *Incremental algorithm* can be applied to find the right referring expression for the object pointed by the arrow in Figure 1. Detail how the algorithm proceeds step-by-step.

Figure 1: Tabletop with three objects, specified as $1 = \{col = blue, type = plate\}, 2 = \{col = blue, type = mug\}, 3 = \{col = red, type = mug\}$.

Bonus question: Dale and Reiter's algorithm often performs well in practice, but does not always produce optimal descriptions. Can you explain why this is the case, and give an example where the algorithm actually produces a sub-optimal referring expression for an entity?

4 Phrase-based statistical machine translation (30%)

The basic model for phrase-based statistical machine translation may be expressed by the following formula.

$$e_{\text{best}} = \text{argmax}_e \prod_{i=1}^{I} \phi(\bar{f}_i|\bar{e}_i)\ d(start_i - end_{i-1} - 1) \prod_{i=1}^{|\mathbf{e}|} p_{LM}(e_i|e_1 \ldots e_{i-1})$$

(a) Answer each of the following three questions using one to four sentences for each.

1. What does the part $\prod_{i=1}^{|\mathbf{e}|} p_{LM}(e_i|e_1 ... e_{i-1})$ model?
2. What is the sequence: $(\bar{f}_1, \bar{e}_1), (\bar{f}_2, \bar{e}_2) \ldots (\bar{f}_I, \bar{e}_I)$?
3. What does $\phi(\bar{f}_i|\bar{e}_i)$ express?

(b) There are three stages to learning a phrase table:

1. Word align the sentence aligned parallel corpus.
2. Extract phrase pairs.
3. Score phrase pairs.

We are translating between German and English, and we assume we have accomplished the first step and established a word aligned parallel corpus. The table shows the result for one sentence pair.

	Mary	had	been	watching	Bill
Mary	x				
hat		x			
Bill					x
beobachtet				x	

List all the phrase pairs which may be extracted from this word alignment.

(c) Explain how the third step, the scoring of the phrase pairs, is carried out.

5 Translation (20%)

There is a slogan which says that

> All you need to translate is a dictionary.

Professional translators argue that this slogan is false.

Also the first idea that comes to mind when people start thinking about machine translation is to base it on a bilingual dictionary. Why isn't a dictionary enough for (machine) translation? Choose three different types of phenomena that may face a machine translator and which cannot be solved by dictionary look-up. Present the phenomena and explain why they make a problem for the dictionary method.

END