GÖTEBORGS UNIVERSITET
FILOSOFI, LINGVISTIK OCH VETENSKAPSTEORI

(partial)
ANSWER SHEET

Exam: Natural Language Processing (LT2103)

December 20, 2010, 09.00-12.000
Room: V30

Write your name and personal ID number (”personnummer”) clearly below. At every page, write only the exam number. This to make the exam anonymous (the name will be removed before the exam correction).

Limits	Pass	20 points
	Pass with distiction	30 points
	Maximum	40 points

Accessories Paper and pencil
The Jurafsky & Martin book
But no hand-written notes!

Notes Write readable! (unreadable = wrong)
Keep it simple! (too complicated = wrong)
Number the pages, and start every question on a new page

Words (12p)

1. Write regular expressions for the following languages. You may use either Perl/Python notation or the minimal "algebraic" notation, but make sure to say which one you are using. By "word", we here mean an alphabetic string separated from other words by whitespace, any relevant punctuation, line breaks, and so forth (the shortcut symbol **\s** is useful here).

(a) the set of all alphabetic strings (2p)

[a-zA-Z]+
or [a-zA-Z]*

(b) the set of all lower case alphabetic strings starting with a b (2p)

b[a-z]*
if "+" instead of "*", 1 point (does not accept "b")
ab[a-z]* also OK

(c) the set of all strings with three consecutive repeated words (e.g., "Humbert Humbert Humbert" and "the the the" but not "the bug bug" or "the big bug") (2p)

([A-Za-z]+)\s\1\s\1
or [A-Za-z]+{3}

(d) all strings that start at the beginning of the line with a word and that end at the end of the line with an integer (2p)

^[A-Za-z]+.*\d$

2. Assume you know the following probabilities:

P(i | <s>)
P(want | i)
P(to | want)
P(fly | to)
P(</s> | fly)

How would you use these to compute (an approximation of)

P(<s> i want to fly </s>)

Write the formula you would use to compute the probability. (4p)

P(</s> | fly) x (fly | to) x P(to | want) x P(want | i) x P (i | <s>)

Syntax (16p)

3.

(a) Complete the context-free grammar below so that it generates (at least) the sentences listed. (4p)

GRAMMAR	LEXICON
S -> NP VP	ProperNoun -> Vincent
NP ->ProperNoun	ProperNoun -> Mia
NP -> ProperNoun Rel	Det -> a
Nominal -> Noun	Det -> her
Nominal -> Noun Rel	Noun -> gun
Rel -> Wh VP	Noun -> robber
VP -> IV	Wh -> who
VP -> TV NP	Wh -> that
VP -> DV NP PP	Preposition -> to
PP -> Preposition NP	IV -> fell
	TV -> loved
	DV -> gave

SENTENCES
Vincent loved Mia.
Mia gave the gun to the man.
Marsellus knew the robber that fell.

Det -> the
ProperNoun -> Marsellus
Noun -> man
NP -> Det Nominal (or NP -> Det Noun)
TV -> knew

(b) Assign parse trees to all three sentences. (3p)

(c) Convert the grammar into Chomsky Normal Form. (2p)

VP -> DV X1
X1 -> NP PP

(or VP -> X1 PP
X1 -> DV NP)

VP -> fell

Nominal -> gun | robber | man

NP -> vincent | mia | marsellus

(d) Given the CNF version of your grammar, draw the parse tree of ”Mia gave the gun to the man”. (2p)

(e) following the CKY algorithm formally described in fig 13.10 (and informally described in Section 13.4.1), complete the CKY parse table for ”Mia gave the gun to the man”. (5p)

Note on algorithm notation: in the algorithm, the notation {A | C(A)} is used, where C(A) is some condition on A. This means "the set of all A's such that C(A) holds". Note also that in a loop beginning with "for n <- from x downto y", if x is lower than y, nothing happens.

Two alternatives (depending on how VP rule was converted to CNF):

mia	gave	the	gun	to	the	man
ProperNoun, NP						S
	DV					VP
		Det	NP			X1
			Noun, Nominal			
				Prep		PP
					Det	NP
						Noun, Nominal

mia	gave	the	gun	to	the	man
ProperNoun, NP						S
	DV		X1			VP
		Det	NP			
			Noun, Nominal			
				Prep		PP
					Det	NP
						Noun, Nominal

Semantics (7p)

4. Here is the WordNet entry for ”sentence”:

Noun

- S: (n) **sentence** (a string of words satisfying the grammatical rules of a language) *"he always spoke in grammatical sentences"*
- S: (n) conviction, judgment of conviction, condemnation, **sentence** ((criminal law) a final judgment of guilty in a criminal case and the punishment that is imposed) *"the conviction came as no surprise"*
- S: (n) prison term, **sentence**, time (the period of time a prisoner is imprisoned) *"he served a prison term of 15 months"; "his sentence was 5 to 10 years"; "he is doing time in the county jail"*

Verb

- S: (v) **sentence**, condemn, doom (pronounce a sentence on (somebody) in a court of law) *"He was condemned to ten years in prison"*

(a) Which of these senses are related by homony, and which are related by polysemy? For any senses which are polysemous, give an argument as to how the senses are related. (4p)

Polysemous:
- Alternative A: 2, 3 and 4 are related by polysemy since they are all related to legal proceedings

- Alternative B (more detailed analyses of polysemy relations):
  - 2-3 Judgement <-> Period of imprisonment as result of judgement in criminal case
  - 2-4 Judgement <-> Act of judging
  - 4-3 Act of judging <-> Period of imprisonment as result of judgement in criminal case

Related by homonymy: 1 to all others (1 point)

(b) Assume the following bag-of-words vector for "sentence":

[grammatical, spoke, condemned, prison, time, jail, serve, term, offense]

What is the correct binary vector representation for the following sentences (using a window size of 20)?

(i) This sentence are not grammatical because it has not followed the grammatical rules of syntax.

[1,0,0,0,0,0,0,0,0]

(ii) Meta-analysis of previous studies shows that a prison sentence does not reduce future offenses.

[0,0,0,1,0,0,0,0,1] (last 1 could be 0)

(iii) We are condemned if we do not know how to use a term in a sentence.

[0,0,1,0,0,0,0,1,0]

(3p)

Pragmatics (5p)

5. Write a finite-state automaton for a dialogue manager for checking your bank balance and withdrawing money at an automated teller machine. (5p)

(This is merely a suggestion. Note that I have used text only; it is better to actually draw an automaton. CAPITALISED words are states.)

START: Welcome. > PIN

PIN: Please enter your PIN code. ->
PIN correct > ALT
PIN incorrect > FAULT1

FAULT1: Sorry, that was the wrong PIN code. Please try again.
PIN correct > ALT
PIN incorrect > FAULT2

FAULT2: Sorry, wrong again. This is your final try.
PIN correct > ALT
PIN incorrect > END

ALT: Do you want to check your bank balance or withdraw money?

”Bank balance” -> You have X kronor on your account. Do you want to continue?
”No”-> END
”Yes”-> ALT

”Withdraw” -> How much do you want to withdraw? > TAKE

TAKE: Please take your cash. Do you want to continue?
”No”-> END
”Yes”-> ALT

END: Thank you for using this service. Goodbye.