

## SGN-4106 Speech Recognition

Exam 20.8.2013

Annamaria Mesaros, annamaria.mesaros@tut.fi

Jani Nurminen, jani.nurminen@tut.fi

Students may use their own calculators (a standard scientific calculator). The answers should be in English. The questions do not need to be returned.

### Problem 1

Explain the meaning of the following terms (1 point each):

(Write maximum two sentences for each.)

- a) formants
- b) left-to-right HMM
- c) Gaussian mixture model
- d) critical band
- e) backward algorithm
- f) linear prediction

### Problem 2

Explain the steps required for building a word recognition system that will recognize the ten digits ("one", "two", "three" and so on) spoken in isolation. Specifically, what are the components of the system and how it is used for recognition?

(Write maximum one page.) (6 points)

### Problem 3

Describe the processing steps needed for calculating mel-frequency cepstral coefficients (MFCCs) in a standard front end of a speech recognition system. Explain shortly the purpose of each operation. No formulas needed!

(Write maximum one page.) (6 points)

### Problem 4

a) Explain the Viterbi algorithm and what problem it solves in speech recognition. No formulas needed! (Write maximum half a page.) (3 points)

b) Explain shortly the construction of a bigram language model and how it is used in a speech recognition system. No formulas needed! (Write maximum half a page.) (3 points)

### Problem 5

A two-state discrete-output HMM has the following parameters  $\lambda$ :

$$P(q_1|q_1) = 0.7, P(q_2|q_1) = 0.3,$$

$$P(q_1|q_2) = 0.2, P(q_2|q_2) = 0.8,$$

$$b_1(A) = 0.1, b_1(B) = 0.4, b_1(C) = 0.5,$$

$$b_2(A) = 0.3, b_2(B) = 0.5, b_2(C) = 0.2.$$

At time  $t = 1$  the HMM is in state 1.

a) Calculate the total probability  $P(O|\lambda)$  of observation sequence  $O = o_1o_2o_3o_4 = CBAB$ . (3 points)

b) What is the most likely state sequence given the above observation sequence  $O$ ? (3 points)