

**SGN-4106 Speech Recognition**  
exam 25.9.2006

You may answer either in Finnish or English.

**Problem 1.**

- a) What 3 features are used to characterise how consonants are produced? (3 points)
- b) Explain how the formula  $Mel(f) = 2595 \log_{10}(1 + f/700)$  is used to design a Mel-scale filter bank. (2 points)
- c) Why is scaling used in the forward-backward algorithm? (1 point)

**Problem 2.** Give the statistical formulation of the speech recognition problem and explain what each term in it means. (6 points)

**Problem 3.** For the HMM in Figure 1, consider the output sequence

$$O = 01011.$$

The HMM is in state *A* at time 1. Calculate the following (the state at time *t* is denoted by  $q_t$  and the output at time *t* by  $o_t$ ):

- a)  $P(O)$ . (2 points)
- b)  $P(q_2 = A|O)$ . (2 points)
- c)  $P(o_5|o_1, o_2, q_1 = A, q_2 = C, q_5 = A)$ . (2 points)
- d) What is the most likely state sequence to have produced *O*? (2 points)

**Problem 4.**

Draw a diagram of the standard MFCC front-end of a speech recognizer. Explain how it operates and what properties of speech it relies on. (6 points)

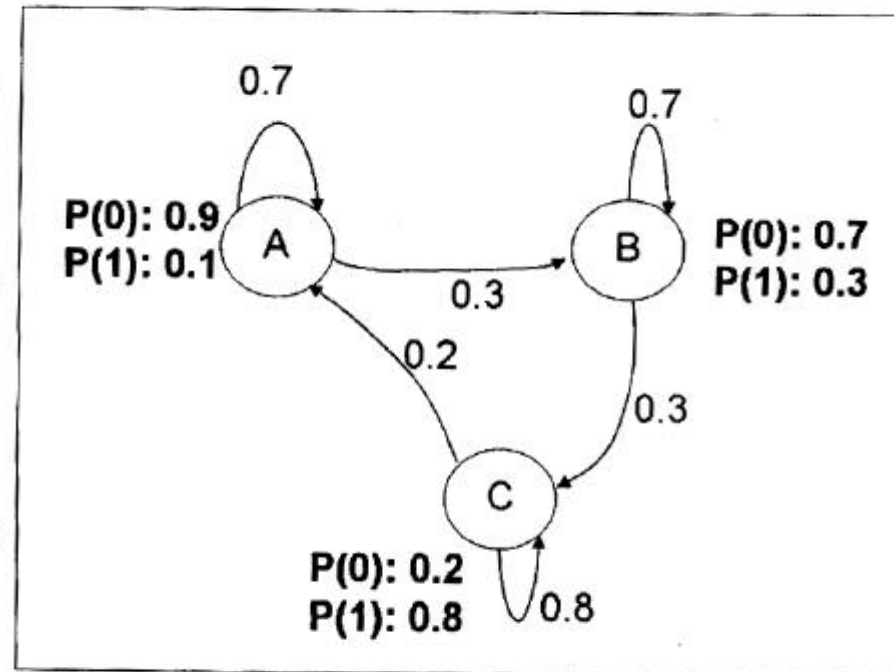


Figure 1: Hidden Markov model for Problem 3.