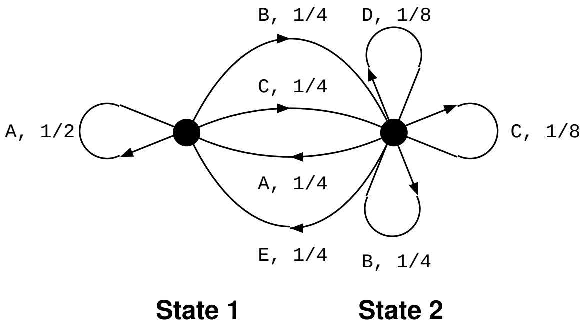
6 Information Theory and Coding (JGD)

(*a*) A two state Markov process emits the letters {A, B, C, D, E} with the probabilities shown for each state. Changes of state can occur when some of the symbols are generated, as indicated by the arrows.



(*i*) What are the state occupancy probabilities? [1 mark]

(*ii*) What is the probability of the letter string AD being emitted? [1 mark]

(*iii*) What is the entropy of State 1, what is the entropy of State 2, and what is the overall entropy of this symbol generating process? [5 marks]

(*b*) A fair coin is secretly flipped until the first head occurs. Let X denote the number of flips required. The flipper will truthfully answer any "yes-no" questions about his experiment, and we wish to discover thereby the value of X as efficiently as possible.

(*i*) What is the most efficient possible sequence of such questions? Justify your answer. [2 marks]

(*ii*) On average, how many questions should we need to ask? Justify your answer. [2 marks]

(*iii*) Relate the sequence of questions to the bits in a uniquely decodable prefix code for X. [1 mark]

(*c*) Define complex Gabor wavelets, restricting yourself to one-dimensional functions if you wish, and list four key properties that make such wavelets useful for encoding and compressing information, as well as for pattern recognition. Explain how their self-Fourier property and their closure under multiplication (i.e. the product of any two of them is yet again a Gabor wavelet) gives them also closure under convolution. Mention one disadvantage of such wavelets for reconstructing data from their projection coefficients. [8 marks]