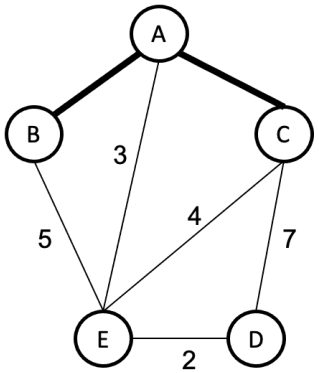
Algorithm Engineering
12 June 2020 – time 30 minutes

Question #1. Let us be given the following symbols and probabilities: p(a) = 1/4, p(b)=1/4, p(c) = 1/2. Compress the text "AC" via Arithmetic coding, using ratios to make computations, and emitting the final sequence of bits.

Question #2. You are given the graph below, in which the bold edges {(A,B), (A,C)} are the ones already inserted in the MST under construction.

- Which is the next edge inserted in the MST by Kruskal's algorithm
- Which is the next edge inserted in the MST by Prim's algorithm

Please motivate the two answers.



Question #3. Given the suffix tree of a string T[1,n], how do you compute the longest substring of T which occurs at least L times? What is the time complexity of the proposed algorithm in function of n?

Question #4. Show the first 3 codewords of the (s,c)-code with s=1 and c=3, hence $s+c = 4 = 2^2$ (comment your calculations).