

Data Compression 1.3.2010

1. The probabilities of symbols a , b , c and d are $1/3$, $1/3$, $1/4$ and $1/12$ respectively. Construct the Huffman codes for these symbols. What is the average code length. Compare it with the entropy. (Exact numeric calculations are not necessary.) (4p)
2. The probabilities of a , b and c are $1/2$, $1/3$ and $1/6$. Arithmetic code is 1001. What three symbols were coded? (4p)
3. Encode string emme räökkääkkään kääkkää with LZ77 algorithm of unbounded dictionary and buffer sizes. Decode the result. (4p)
4. Give short answers to the following:
 - (a) Why is prefix property important in coding? (1p)
 - (b) What is predictive compression? (1p)
 - (c) What is the purpose of the Burrows-Wheeler transformation? (1p)
 - (d) For what kind of problems no compression method is successful? (1p)