Information Theory Exercise Sheet #3 (Arithmetic Codes and Entropy Rates)

University of Amsterdam, Master of Logic, Fall 2015
Lecturer: Christian Schaffner
TA: Mathias Madsen

Out: Wednesday, 10 November 2015
(due: Wednesday, 17 November 2015, 9:00)

1. **Straws (McKay, Ex. 15.4)** How can you use a fair coin to draw lots among three people? Come up with at least two methods and compare them in terms of (1) fairness, and (2) expected number of coin flips.

2. **Arithmetic Coding for a Bent Coin** We flip a bent coin with bias $1/4$.
   (a) Construct the Shannon-Fano-Elias code for a single sample from this distribution.
   (b) Construct the Shannon-Fano-Elias code for two independent samples from this distribution.
   (c) For each of your two codes, evaluate how many bits they use per symbol, and compare this performance to the theoretical minimum.

3. **Interval Representations*** (involves programming) We flip a coin with bias $1/3$.
   (a) Write a function that maps a sequences of outcomes (0s and 1s) to an interval in the manner of arithmetic coding.
   (b) Test your function on a sequence of five 1s, and a sequence of five 0s.
   (c) Test it on a sequence of fifty 1s, and a sequence of fifty 0s.

4. **Binary Approximation**
   (a) Prove that any interval contains a binary interval which is at most four times narrower than itself.
   (b) Derive an upper bound on the difference between (i) the surprisal at observing a sequence $(X_1, X_2, \ldots, X_n) = (x_1, x_2, \ldots, x_n)$ and (ii) the length of its arithmetic codeword.
   (c) Derive an upper bound on the difference between (i) the entropy of the distribution over such sequences, and (ii) the expected length of their arithmetic codes.

5. **Binary Interval Names*** (involves programming)

(a) Implement a function that maps an interval $[a, b] \subseteq [0, 1]$ to the corresponding Shannon-Fano-Elias code of that interval.

(b) Test your function on intervals of the form $[1/2 - \varepsilon, 1/2 + \varepsilon]$ for tiny ε.

6. **Forwards and Backwards Prediction** Consider the following two tasks:

   (a) Guessing the next letter of a text given the preceding ones:

   ... `RE PARTICULARLY IMPR_`

   (b) Guessing the previous letter of a text given the following ones:

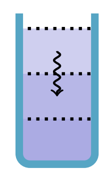
   `_ONTH FOLLOWING THE C` ...

   In general, which task is the more difficult — from a statistical perspective, and from a cognitive? Why?

7. **The Entropy of English** Assume for the purposes of this exercise that English uses a character set of 27 characters, and that it has an entropy rate of $H = 1.5$ bits per character.

   Under these assumptions, what approximately is the probability that a randomly drawn string of $n = 20$ characters is a snippet of English text? (In the context of finite sets, a "random" selection is an element drawn from the uniform distribution over the set.)

8. **Random walk with gravity** A molecule moves around in a glass of water which we consider as divided up into three compartments. Whenever possible, the molecule moves one compartment down with probability $1/5$, and one compartment up with probability $1/20$.

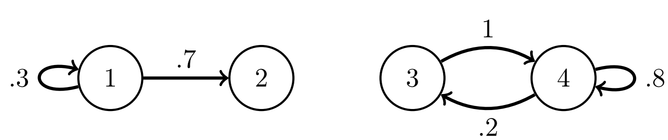


   (a) Write down the transition probabilities for this system.

   (b) Find the stationary distribution of the system.

   (c) What would you guess the equilibrium distribution would look if we had started with k compartments instead of three?

9. **Stationary Distributions** Find all stationary distributions of the family of four-state Markov chains defined by the transition diagram

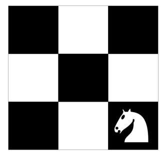


10. **Morse Code (Cover and Thomas, Ex. 4.8)** An alphabet contains a dot which takes one unit of time to transmit and a dash which takes two.

    (a) When the two symbols have probability p and $1 - p$, what's the entropy rate of this process?

(b) For which choice of p is this entropy rate the largest?

11. **Tiny chess** What's the entropy rate of a knight walking on a 3×3 chess board? What about a bishop? Be explicit in your assumptions about the transition probabilities.



Homework is exercises 2, 7, and 10.