Information Theory

Written by Nikita, updated by Nakul

Problem 1.
(a) A magician guessed an integer number between 1 and 100. You can ask him any Yes/No questions about the number. Can you guess the number with full certainty in 7 questions?

(b) Can you do this in 6?

Watch this video by 3Blue1Brown on solving Wordle using information theory.

Problem 2.
Given a balance scale and the weights of 1 lb, 2 lbs, 4 lbs, and 8 lbs (one of each), prove that you can weigh any (integral) load from 1 to 15 lbs.

Problem 3.
Now, you have a balance scale of 1 lb, 3 lbs, 9 lbs, and 27 lbs (one of each). Can you weigh any (integral) load from 1 to 40 lbs? What if, instead, you have two of each?

Problem 4.
(a) You are given a set of scales and 3 coins. The scales are of the old balance variety. That is, a small dish hangs from each end of a rod that is balanced in the middle. The device enables you to conclude either that the contents of the dishes weigh the same or that the dish that falls lower has heavier contents than the other. The 3 coins appear to be identical. In fact, 2 of them are identical, and one is counterfeit and is lighter than the other 2. How do you determine which coin is counterfeit in 1 weighing?

(b) Same problem with 27 coins and 3 weighings.

(c) Can you determine a counterfeit coin among 28 coins in 3 weighings?

Problem 5.
Suppose you got the following first guess in Wordle: ("f" is grey and "ight" are green).

(a) List all the words matching this pattern (you should get eight words, check with the instructor).

(b) Can you guess the word in 5 remaining attempts?

(c) [*] How many extra attempts do you need to guess it for sure?

Problem 6.
You are given 9 labeled marbles, of which 2 are radioactive. For any set of marbles, in one check you can find out if it contains at least one radioactive marble (but you cannot find out how many there are).
(a) Count the number of possibilities for the identity of the two radioactive marbles. (e.g. radioactive marbles are marble 1 and marble 5, or marble 2 and 7, etc.)

(b) How many checks do you need to do to find the two radioactive marbles?

(c) [*] Find an algorithm guaranteeing this number of checks.

Problem 7.
You are given a set of scales and 12 marbles.
The 12 marbles appear to be identical. In fact, 11 of them are identical, and one is of a different weight. Your task is to identify the unusual marble and discard it. You are allowed to use the scales three times if you wish, but not more.
Note that the unusual marble may be heavier or lighter than the others. You are asked to both identify it and determine whether it is heavier or lighter.
Hint: You should distinguish between 24 possibilities. Compose your first weighing so that each of the three outcomes is equally likely.

Problem 8.
You are a detective trying to identify a suspect out of a list of 5 individuals. Based on your investigation, each suspect has a different probability of being the culprit.

1. Alice has a 50% chance of being the culprit.
2. Bob has a 20% chance of being the culprit.
3. Charlie has a 12.5% chance of being the culprit.
4. Daniel has a 12.5% chance of being the culprit.
5. Eve has a 5% chance of being the culprit.

You are allowed to ask a series of Yes/No questions to the entire group (who agree to answer honestly) to determine who the culprit is. Give a strategy that determines the culprit using 2 questions on average.

Problem 9 (*)**.**
You and your friend are trying to create a secret language using only your eyes to communicate – you can either wink your left eye or your right eye to send a message. For the sake of simplicity we denote these actions using 0 and 1. Also, assume that you only want to communicate words that use the letters A, B, C, D and E. Your goal is to come up with a method of encoding these letters in your secret language. For example consider the following encoding: $A \to 000, B \to 001, C \to 010, D \to 011, E \to 100$. Each letter requires 3 winks to communicate which you and your friend find too tiresome. Your friend suggests an alternate encoding: $A \to 0, B \to 1, C \to 10, D \to 11, E \to 100$.

(a) What is the issue with this encoding? **Hint:** What are the encodings for CAB and EB?

(b) Come up with an alternate encoding where some codewords are shorter than 3 symbols but doesn't face the above issue i.e. there is a unique way to decode every codeword.

(c) After analyzing your communication patterns you note that you use the letters with the following frequencies – A: 0.17, B: 0.35, C: 0.17, D: 0.15, E: 0.16. Give an alternate encoding scheme where, on average, the length of a codeword is 2.31 bits.

(d) Can you do any better?

Problem 10 (*).
There are 5 weights of different mass. In one operation, one can choose an ordered triple of weights (A, B, C) and find out if the statement "$m(A) < m(B) < m(C)$" is true (where $m(X)$ denotes the mass of the weight X). Is it possible to use 9 such operations to find out the ordering of the weights?