Information Theory

Curtis Liu & Sunny Liang

2023-12-03

LaTeX source: https://www.overleaf.com/read/bzfyctdgdmkq#d1682c

Dear instructors, this packet is not meant to be easily done without help. Also, subsequent questions require a good understanding of previous questions. Be sure to go around and guide students, making sure that they have correct answers to key questions.

This worksheet is heavily based on this video. Some problems are taken directly from it. Watching the video will make the worksheet pretty much trivial.

1. Once.
2. Twice.
3. $3^5 = 243$ so five times.
4. https://en.wikipedia.org/wiki/Balance_puzzle#Variations
5. https://en.wikipedia.org/wiki/Balance_puzzle#Twelve-coin_problem
6. Harold is more likely to have 10 toes, having 11 toes is more "informative"
7. $\log_b(\frac{1}{P(x)}) = -\log_b(P(x))$ where b can be any number. For future questions, we will want to choose $b = 2$ to get our answer in bits.
8. For the two latter questions, I used $P(\text{polydactylism in toes at birth}) = 1/500$. Of course, your answer will very depending on your research. Also, if you decide to take into account that most people with polydactylism get surgery for it, things get more complicated. The answer doesn't matter, the point is just that the less likely outcome gives way more information.
   - $I(x = \text{a fair coin lands heads}) = 1.0$ bits
   - $I(x = \text{you roll a 1 on a fair dice}) = 2.6$ bits
   - $I(x = \text{Harold has 10 toes}) \approx 0.003$ bits
   - $I(x = \text{Harold has more than 10 toes}) \approx 9$ bits
9. $H(X) = \sum_i P(x_i) \cdot I(x_i)$, this is just the expected value of the self-information. X is the corresponding discrete random variable, but I'm not pressed about notation.
10. $H(\text{Vancouver}) \approx 1.52$ and $H(\text{Los Angeles}) \approx 0.33$. The weather in Vancouver is "less predictable", therefore it has higher entropy.
11. $S = k_B \ln W$
12. $H = -4 \cdot 0.25 \cdot \log_2(0.25) = 2.0$
13. See Figure 2 on problem document.
14. $H = -(0.5 \cdot \log_2(0.5) + 0.25 \cdot \log_2(0.25) + 0.125 \cdot \log_2(0.125) + 0.125 \cdot \log_2(0.125)) = 1.75$
15. See Figure 4 in this document.

16. Ditto.

17. $H \approx 2.23$. I would expect students who get this to use Shannon-Fano coding (alternative link, go to section Shannon-Fano tree to see procedure). You can achieve 2.30 using Huffman coding but I don't expect this from anybody.

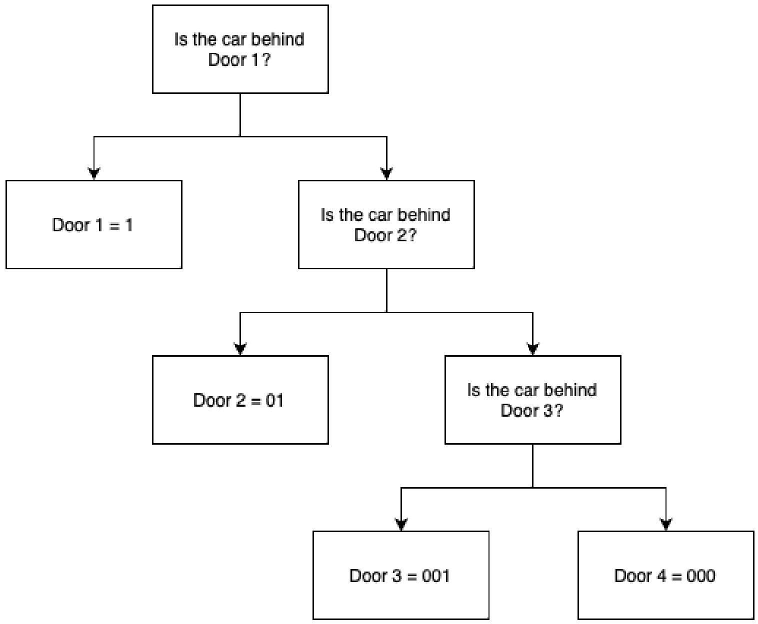

Figure 4: Notice how the symbol-length varies from 1 to 3 inclusive instead of just using two digits for all four outcomes.