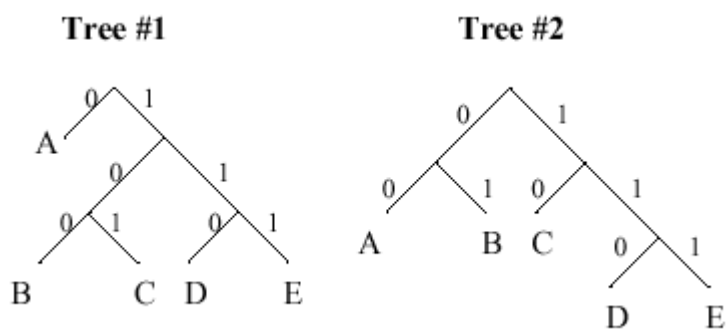




LE1.5.1: Variable-length Encoding

pontos 2 / 2 (sem classificação)

Consider the following two Huffman decoding trees for a variable-length code involving 5 symbols: A, B, C, D and E.



Using Tree #1, decode the following encoded message: "01000111101".

Decoded message (as a string):

ABAEC

✓ Answer: ABAEC

Explanation

To decode the message, start at the root of the tree and consume digits as you traverse down the tree, stopping when you reach a leaf node. Repeat until all the digits have been processed. Processing the encoded message from left-to-right:

"0" => A
"100" => B
"0" => A
"111" => E
"101" => C

Suppose we were encoding messages that had the following probabilities for each of the 5 symbols:

$p(A) = 0.5$
 $p(B) = p(C) = p(D) = p(E) = 0.125$

Which of the two encodings above (Tree #1 or Tree #2) would yield the shortest encoded messages averaged over many messages?

☒ Tree #1

☐ Tree #2



Explanation

Using Tree #1, the expected length of the encoding for one symbol is:

$1 \cdot p(A) + 3 \cdot p(B) + 3 \cdot p(C) + 3 \cdot p(D) + 3 \cdot p(E) = 2.0$

Using Tree #2, the expected length of the encoding for one symbol is:

$2 \cdot p(A) + 2 \cdot p(B) + 2 \cdot p(C) + 3 \cdot p(D) + 3 \cdot p(E) = 2.25$

So using the encoding represented by Tree #1 would yield shorter messages on the average.

Enviar

ⓘ Answers are displayed within the problem

Calculator

Discussion

Topic: 1. Basics of Information / LE1.5

Ocultar discussão

Adicionar publicação

Show all posts ▼by recent activity ▼

I think we can not generalize and say Tree#1 yield shorter messages, it depends on the specific message. right?

< Previous

Next >



edX

- About
- Affiliates
- edX for Business
- Open edX
- Careers
- News

Legal

- Terms of Service & Honor Code
- Privacy Policy
- Accessibility Policy
- Trademark Policy
- Sitemap
- Cookie Policy
- Your Privacy Choices

Connect

- Idea Hub
- Contact Us
- Help Center
- Security
- Media Kit

