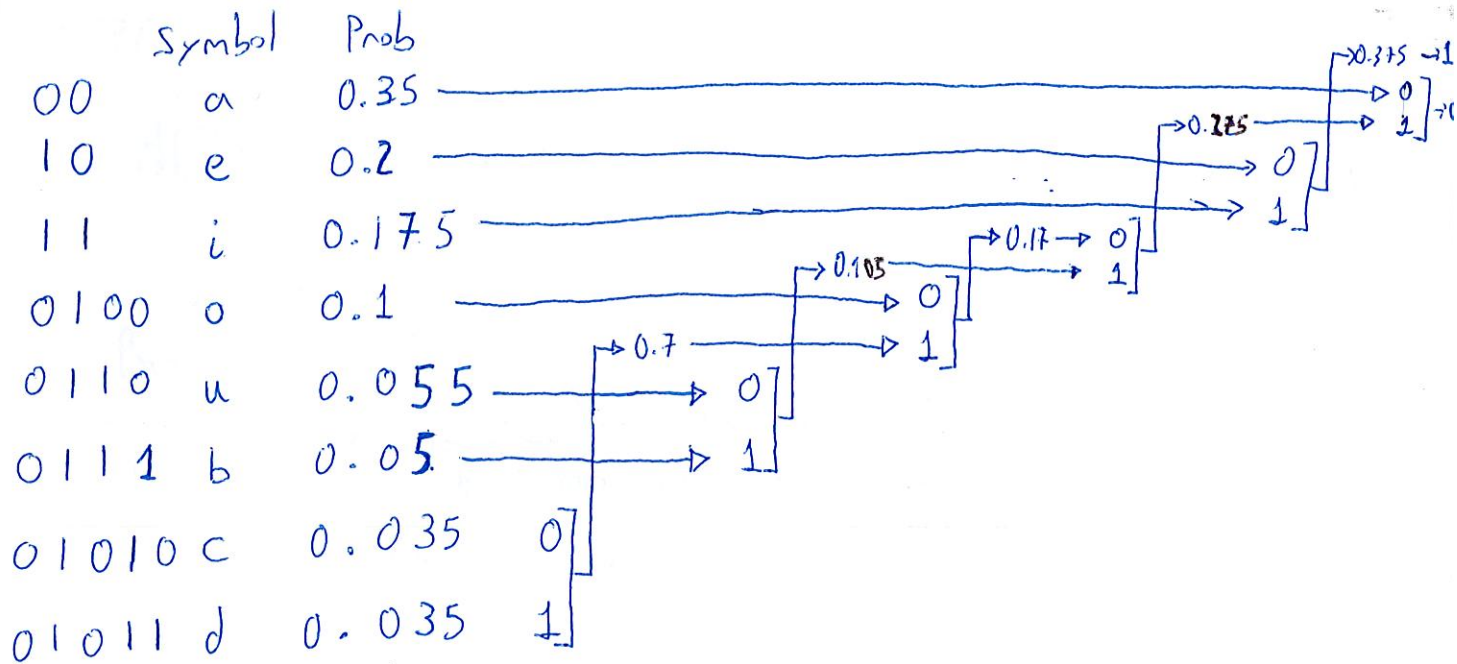


Source coding ~ compression plus some other symbols . , : () ' ...

- E.g. 26 letters in the alphabet! If we assign 5 bits to each letter, lots of waste. Use fewer bits for most common

• Say



↑
Backtrack

- Important to observe that no sequence is a prefix to any other, to avoid confusion when sent as string
- Average length^{bits} per letter:

$$(0.35 + 0.2 + 0.175) \cdot 2 + (0.1 + 0.055 + 0.05) \cdot 4 + 0.07 \cdot 5 = 2.62$$

- Optimal when encoding symbols separately, but does not capture correlations. We can compress more by dealing with words instead of letters, etc.

^
3!!!