stream codes (Mackay chapter 6)

problems with Huffman codes

changing ensembles

- Huffman is optimal for iid sources

the extra bit: we know Huffman gives $H(X) \leq \mathbb{E}[L_C(X)] \leq H(X) + 1$

a	0.001	00000
b	0.001	00001
c	0.990	1
d	0.001	00010
e	0.001	00011
f	0.001	0100
g	0.001	0101
h	0.001	0110
i	0.001	0111
j	0.001	0010
k	0.001	0011

$\mathbb{E}[\text{length}] = 1.034$

$H(X) = 0.114$

$\mathbb{E}[L]/H(X) = 9$

the guessing game

MAJORITY OF AMERICANS -

17 1 12 2 0 0 0 0 1 1 3 1 2 1 0 0 0 0 0

FOUR IN TEN - HATE MATH

20 5 2 1 1 1 1 1 3 1 3 1 13 1 1 1

Code - Huffman Code of (17.1.12.2.0. . .)

how to model data sources

1) iid sources, known distr

2) iid sources, unknown distr
(universal codes)

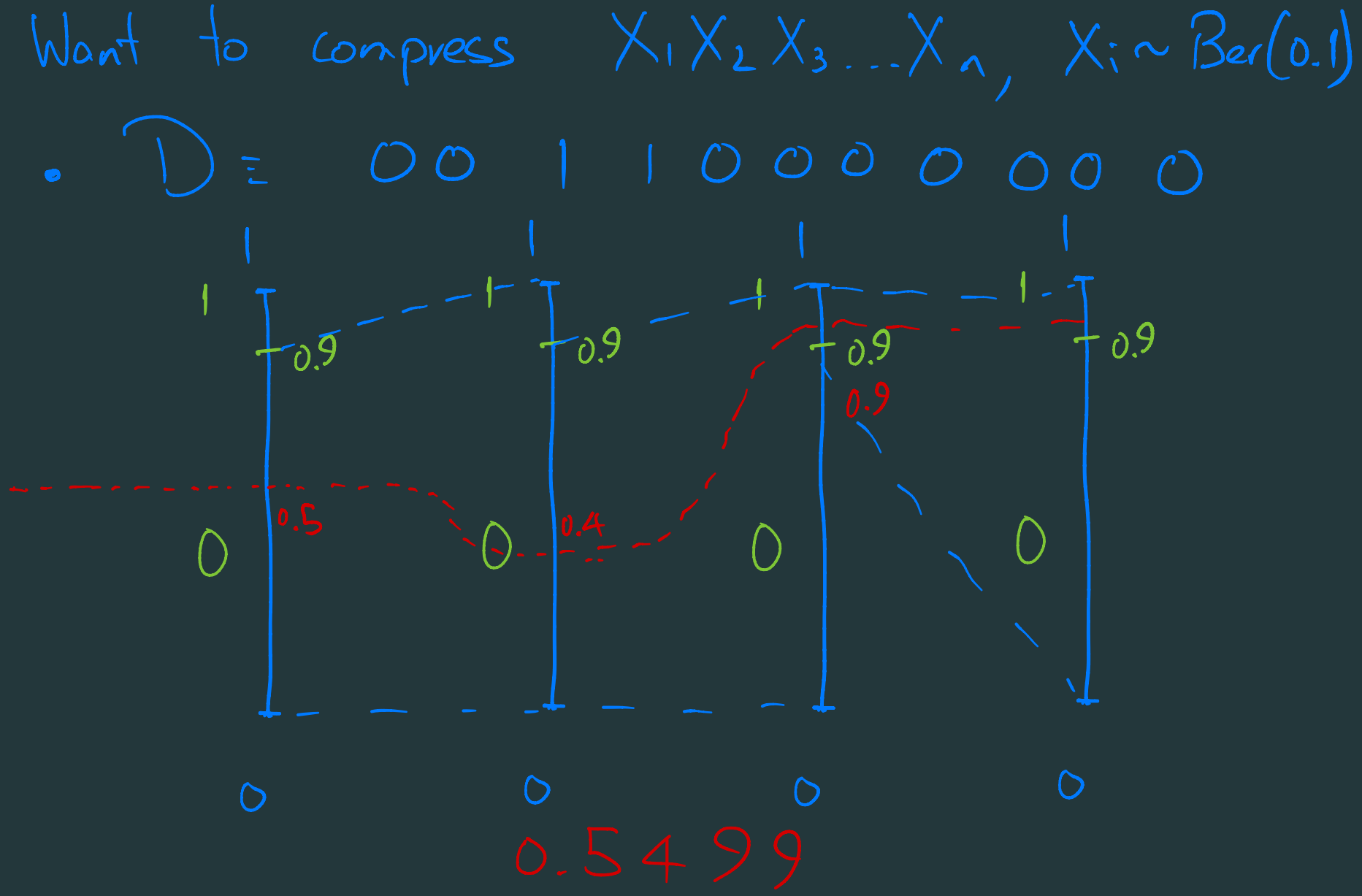
3) oracle-based models (probabilistic model)
(arithmetic coding) $X_1 X_2 \ldots X_n$, $P(x_1)$, $P(x_2|x_1)$, $P(x_3|x_1,x_2)\ldots$

4) Unknown probabilistic model
(Lempel-Ziv, Dictionary codes)

two approaches to stream coding

- Arithmetic Coding
  - 'Encodes entire document as a single number'
  - Requires probabilistic model
    (Separation of modeling & coding)
- Lempel-Ziv coding (Dictionary coding)
  - Generate and store a 'dictionary'
  - Codeword consists of dictionary + message

arithmetic coding



arithmetic coding

Idea - After encoding x_1, x_2 as $0.c_1c_2c_3$, partition $[0,1]$ according to $P[x_3 | x_1 x_2]$

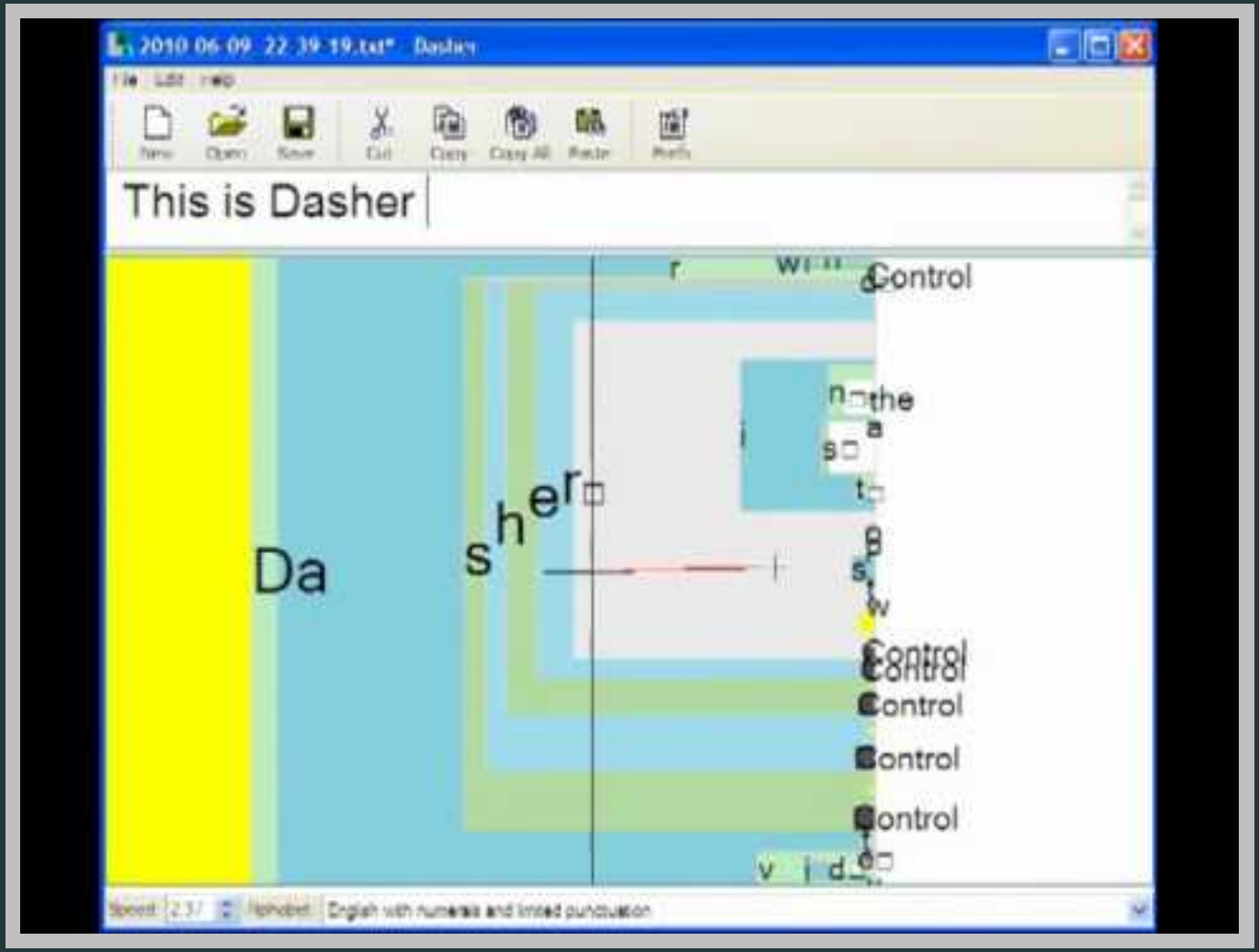
Fact - Symbol x_i heads

$\approx h(x_i | x_1 x_2 \ldots x_{i-1})$

$= \log_2 \frac{1}{P(x_i | x_1 x_2 \ldots x_{i-1})}$ bits

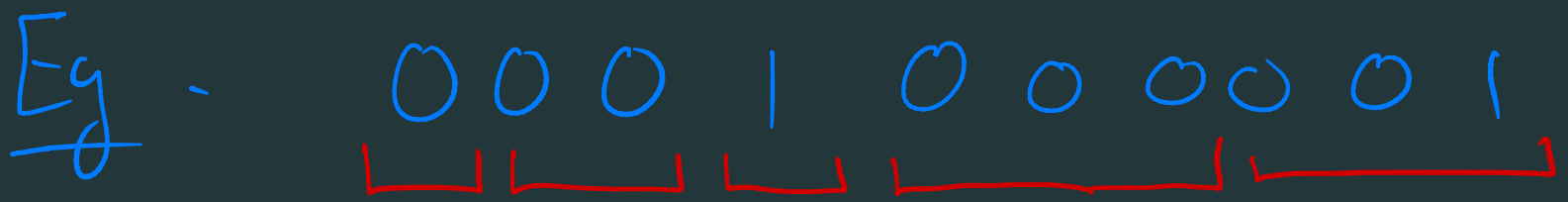
Average codeword length $\approx \log_2 \frac{1}{P(x_1 x_2 \ldots x_n)}$

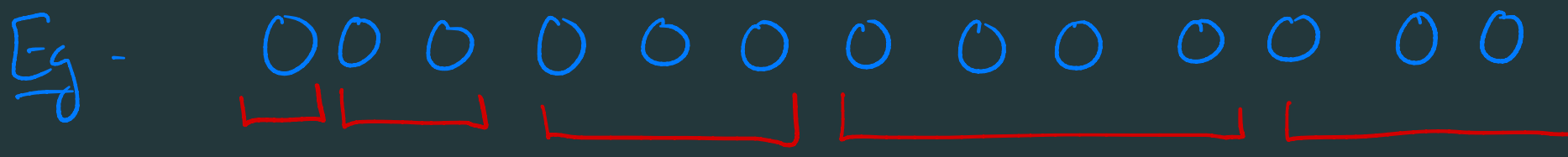
application of arithmetic coding beyond compression



https://www.youtube.com/watch?v=nr3s4613DX8

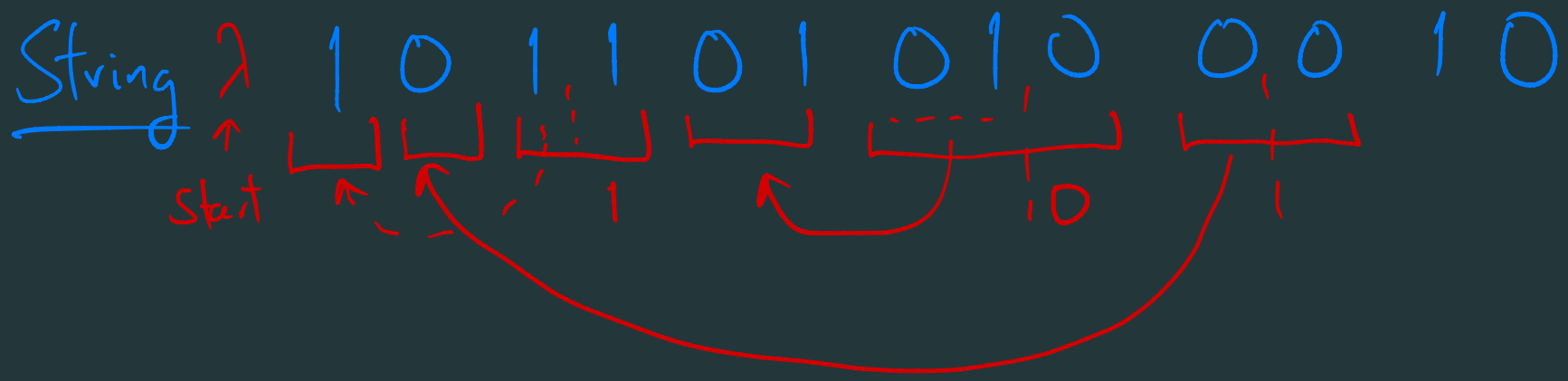
Lempel-Ziv codes (dictionary codes)





Lempel-Ziv-Welch coding

source substrings	λ	1	0	11	01	010	00	10
$s(n)$	0	1	2	3	4	5	6	7
$s(n)_{\text{binary}}$	000	001	010	011	100	101	110	111
(pointer, bit)		(, 1)	(0, 0)	(01, 1)	(10, 1)	(100, 0)	(010, 0)	(001, 0)



- As $n \to \infty$, $|L(X_1 \cdots X_n, C)| \approx H(X_1, \ldots X_n)$