# Breadth: Compressing Information

Information

Is the enemy coming?

Yes / No

10 torches

Message = number of torches lit

Enemy is coming

We need more food

Olaf fell down the well again. LOL.

…

10 torches in 2 groups

Polybius square

Torches in group 1

Torches in group 2

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I/J	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

A   0.5   0
B   0.5   1
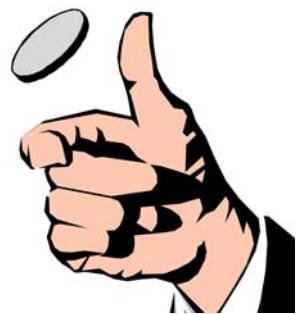
1 bit of information

Yes or No questions

Is it B?

Is it B?

A	0.25	1		
B	0.25	0	1	
C	0.25	0	0	1
D	0.25	0	0	0

codewords

Is it A?   0 / 1
Is it B?   0 / 1
Is it C?   0 / 1
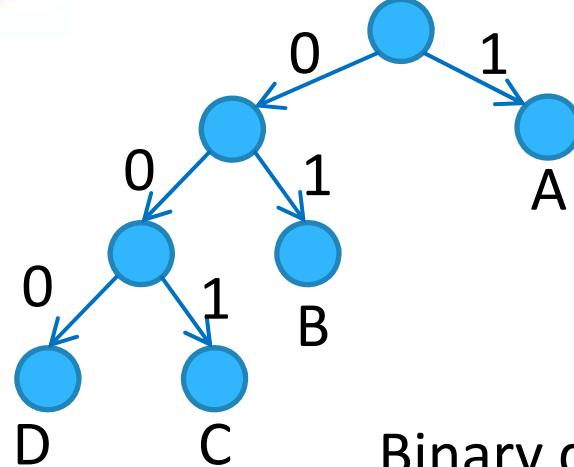
Yes or No questions

Binary decision tree

A	0.25	1			1 bit
B	0.25	0	1		2 bits
C	0.25	0	0	1	3 bits
D	0.25	0	0	0	3 bits

Average: $\sum p \cdot (\#bits) =$ 2.25 bits

Yes or No questions

A   0.25
B   0.25
C   0.25
D   0.25

Can you come up with a better strategy?

How many bits on average will it use?

Yes or No questions

A. 1 bit
B. 1.5 bits
C. 1.75 bits
D. 2 bits
E. 2.25 bits

A	0.25	0	0	2 bit
B	0.25	0	1	2 bits
C	0.25	1	0	2 bits
D	0.25	1	1	2 bits

Average: 2 bits

Binary decision tree

Yes or No questions

A	0.5	0			1 bit
B	0.25	1	0		2 bits
C	0.125	1	1	0	3 bits
D	0.125	1	1	1	3 bits

Average: 1.75 bits
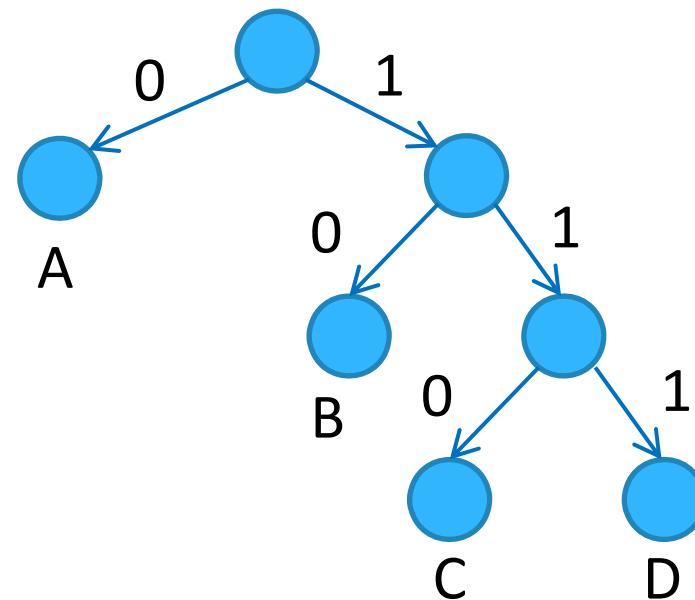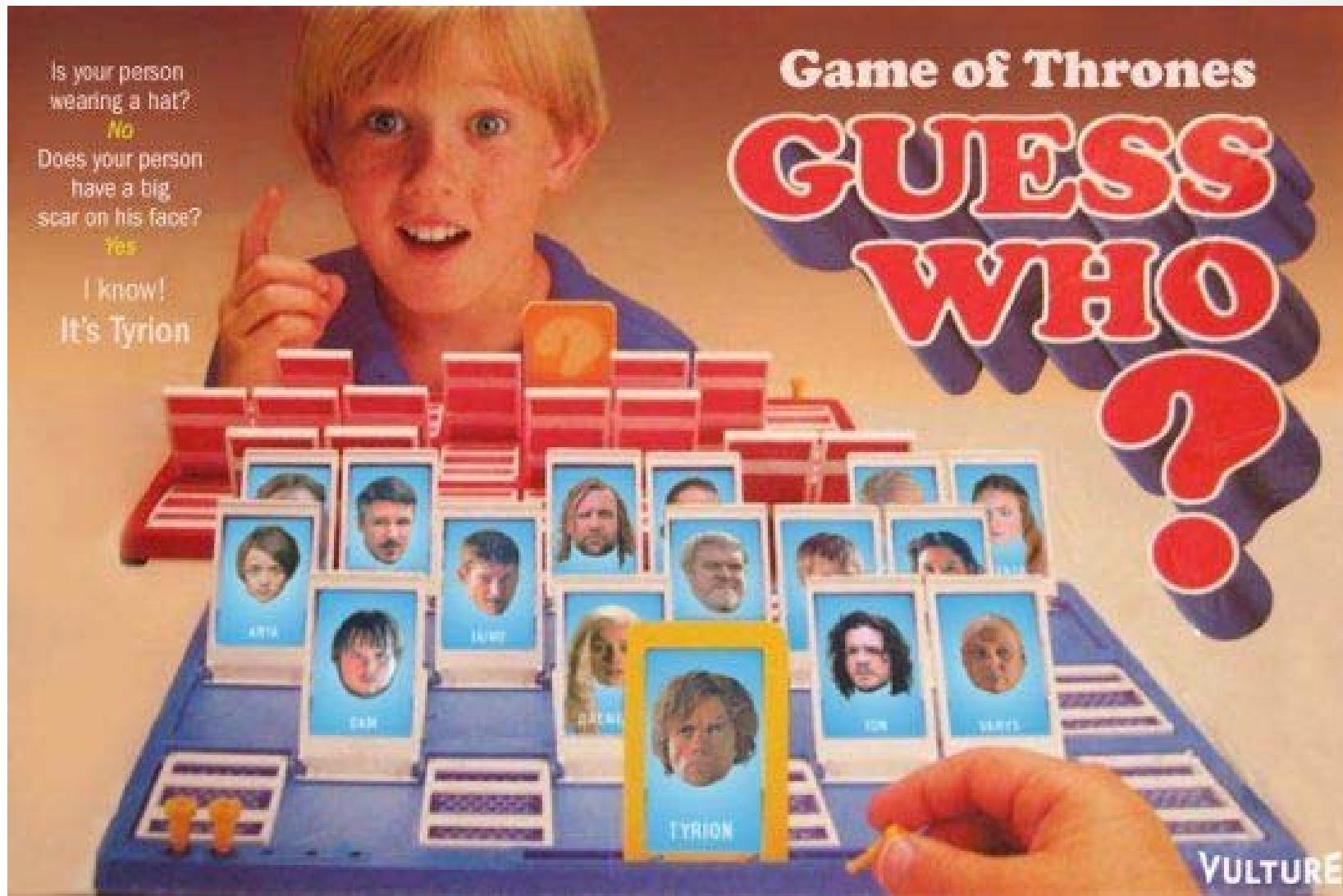
A   0.4

B   0.1

C   0.2

D   0.3

Can you come up with a good strategy?

How many bits on average will it use?

Yes or No questions

A. ~ 1.7 bit   D. ~ 2 bits
B. ~ 1.8 bits   E. ~ 2.1 bits
C. ~ 1.9 bits

A	0.4	0			1 bit
B	0.1	1	0	0	3 bits
C	0.2	1	0	1	3 bits
D	0.3	1	1		2 bits

Average: 1.9 bits

Huffman coding
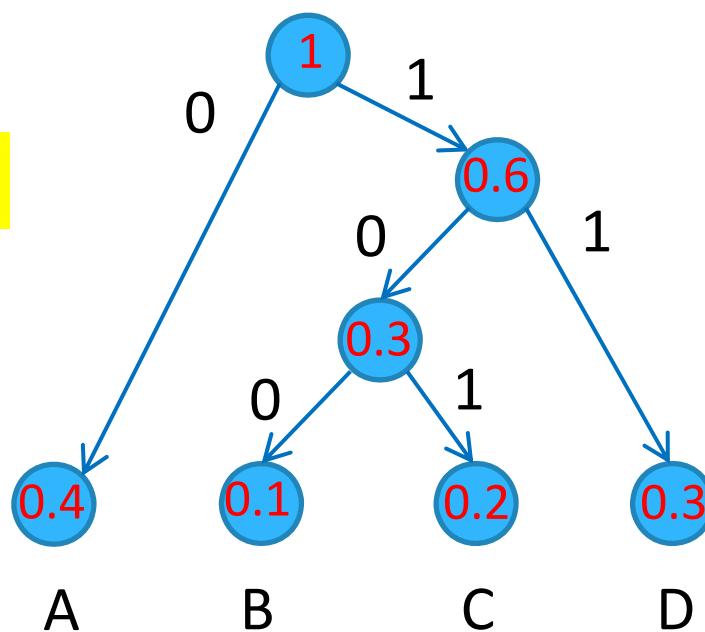
A	0.4	0
B	0.1	1 0 0
C	0.2	1 0 1
D	0.3	1 1

→ Yes or No questions

AABDACADCDAAADCAB …

0010110101011101110001110 10100 …

A   0.05
B   0.05
C   0.15
D   0.2
E   0.2
F   0.35

What is the Huffman code?

Yes or No
questions

A	0.4	0			0		1 bit
B	0.1	1	0	0	0	1	2 bits
C	0.2	1	0	1	1	0	2 bits
D	0.3	1	1		1		1 bit

Average: 1.9 bits   Average: 1.3 bits

→

Yes or No questions

AABDACADCDAAADCAB …

0010110101011101110001110101000 …
00011010011010001100001 …

A	0.4
B	0.1
C	0.2
D	0.3

0		1 bit
0	1	2 bits
1	0	2 bits
1		1 bits

Average: 1.3 bits

You receive 0101. What message was sent?

Yes or No questions

A. ADAD
B. BB
C. AAC
D. ACA
E. BC

A	0.4	0	1 bit
B	0.1	0 1	2 bits
C	0.2	1 0	2 bits
D	0.3	1	1 bits

Average: 1.3 bits

You receive 0/10/1. What message was sent?

Yes or No questions

A C D

Morse code
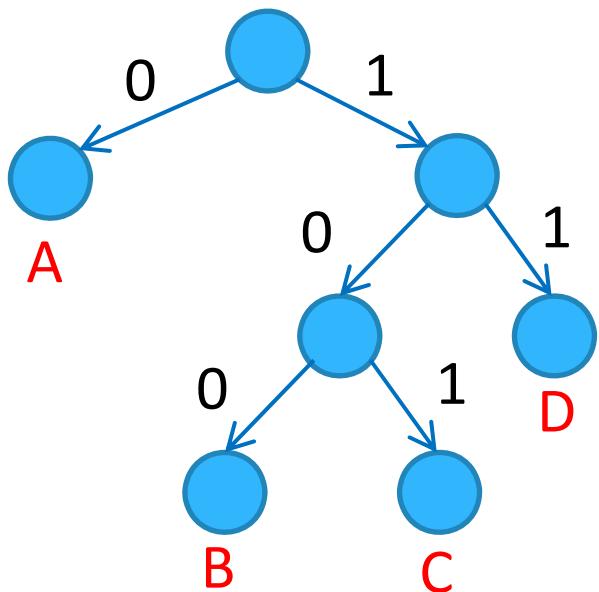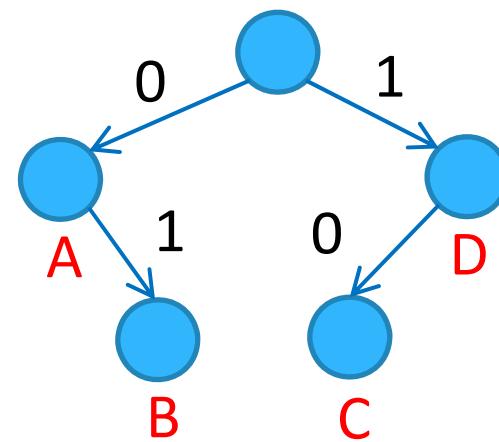
Telegraph Code

In writing a cipher message, please observe the following:

First—Begin every cipher word with a capital letter.

Second—Whenever a blank occurs in a sentence, the word or words supplying such blank must immediately follow the cipher word of the sentence.

Quotations and Correspondence

	Cipher Word
At what price and how soon can you furnish	Quadrants
Quote best price on	Quadrate
In market for	Quadrangle
Quote best price on ……square feet of standard (38-inch) height of …. Radiators	Quadrille
Wire reply quickly	Quadroon
Will wire you to-morrow morning	Quaffed
Have written	Quaggy
Must have information immediately	Quahog
Answer by first mail	Quagmire
See our letter of…… giving full particulars	Quaintly
Have received no reply from you to our letter of	Quakingly
Referring to your telegram of	Quakerism
Referring to your letter of	Qualify
Have received no reply to our telegram of	Quamedit
Referring to our telegram of	Quandary
Referring to our letter of	Quarried
Referring to telephone conversation of to-day	Queerness
Do not understand the meaning of	Quartette
Inclosure mentioned in your letter of…… not received, mail same at once	Quantum
We quote you for immediate acceptance	Quash
F. O. B. factory less the actual rate of freight in carloads or less to any railroad point of destination not to exceed 30c per 100 lbs	Quaternion
Wire at once less than carload freight rate on	Quaternary
Answering your wire of date less than carload rate per cwt. on	Quaternate
Wire carload freight rate on	Quartzite
Answering your wire of date the carload rate per cwt. and minimum weight on	Quaternity
Change my route to read as follows	Questor
Will be here until	Quebracho
Will be in	Quebrith
Immediate specifications, three months delivery	
5 per cent	Quitclaim
7½ per cent	Quittore
10 per cent	Quittance

265

At what price and how soon can you furnish?" ………………… Quadrants

A   1/4   0
B   3/4   1

What is the Huffman code?

1 bit / symbol

Yes or No questions

Can we do better?

A   1/4

B   3/4

AA   1/16   0 0 0   3 bits

AB   3/16   0 0 1   3 bits

BA   3/16   0 1   2 bits

BB   9/16   1   1 bits

Average: 1.6875 bits

0.8438 bits / symbol

We are only considering independent events!

0 16/16
        1
   0 7/16
        1
  0 4/16
     0 1
1/16 3/16 3/16 9/16
AA   AB   BA   BB
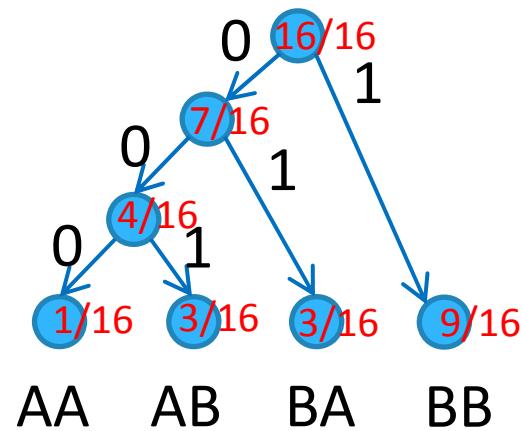
Yes or No questions

A	1/4	AAA	1/64	000000	6 bits
B	3/4	AAB	3/64	000001	6 bits
		ABA	3/64	00001	5 bits
		ABB	9/64	010	3 bits
		BAA	3/64	0001	4 bits
		BAB	9/64	011	3 bits
		BBA	9/64	001	3 bits
		BBB	27/64	1	1 bit

Average: 2.4844 bits

0.8281 bits / symbol

Grouping of 1 symbol   1 bit/symbol

Grouping of 2 symbols   0.8438 bits/symbol

Grouping of 3 symbols   0.8281 bits/symbol

...

Grouping of $\propto$ symbols   ?? bits/symbol

                              < 0.8113 bits/symbol

A   1/4

B   3/4

$$I = \frac{1}{4} \cdot \log_2 \frac{4}{1} + \frac{3}{4} \cdot \log_2 \frac{4}{3}$$

$$= 0.8113 \text{ bits/symbol}$$

Entropy

Claude Shannon

$$I = \sum p \cdot \log_2 \frac{1}{p}$$

$$= -\sum p \cdot \log_2 p$$

Claude Shannon
1916 - 2001

A	0.5
B	0.5

$I = 1$

A	0.25
B	0.75

$I = 0.8113$

A	0
B	1

$I = 0$

$$I = -\sum p \cdot \log_2 p$$

Bits

Information relates to *uncertainty*.

Note: context matters as well

"Tomorrow, it will be over 55 degrees in La Jolla"