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?

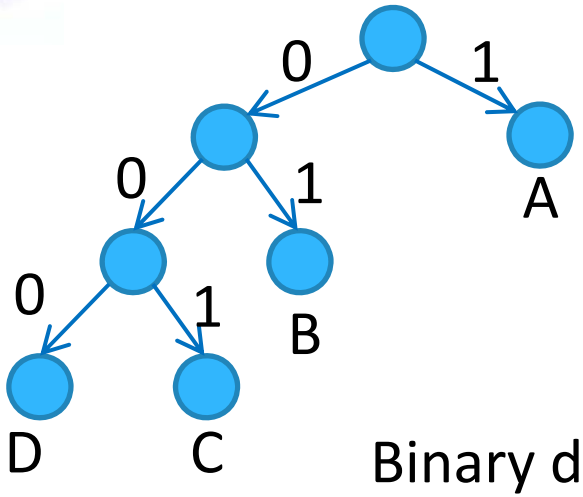
Is it B?

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

codewords

Yes or No questions

Is it A? 0 / 1
Is it B? 0 / 1
Is it C? 0 / 1



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

Can we do better?

Average: 2.25 bits

Yes or No questions

ACABDCBDBDACCDAB …

10011010000010100001000100100100101 …

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

Average: 2.25 bits

Average: 1.5 bits

Yes or No questions

ACABDCBDBDACCDAB …

1001101000001010000100010010010001 01 …

1001010000100101000001 01…

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

Average: 1.5 bits

You receive 0001. What message was sent?

Yes or No questions

A. DDDA
B. BAB
C. CB
D. CDAC
E. DDBA

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

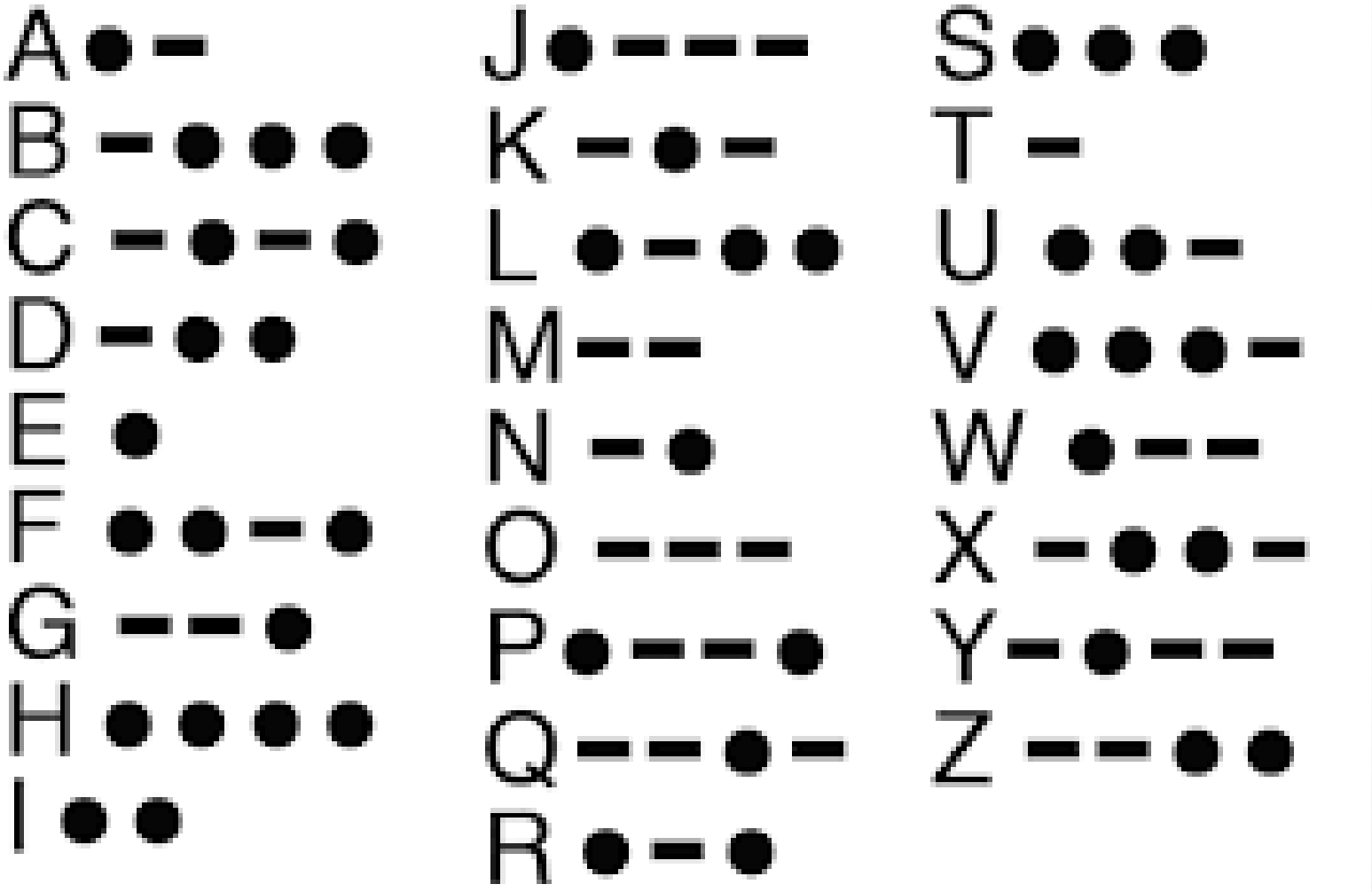
Average: 1.5 bits

You receive 00/0/1. What message was sent?

Yes or No questions

C D A

Morse code



Morse code

A ●−	J ●−−−	S ●●●
B −●●●	K −●−	T −
C −●−●	L ●−●●	U ●●−
D −●●	M −−	V ●●●−
E ●	N −●	W ●−−
F ●●−●	O −−−	X −●●−
G −−●	P ●−−●	Y −●−−
H ●●●●	Q −−●−	Z −−●●
I ●●	R ●−●	

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	Quamoclit
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	Quatemion
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	Quastor
Will be here until	Quebracho
Will be in	Quebrith
Immediate specifications, three months delivery	
... ship... cent	
5 per cent	Quitclaim
7½ per cent	Quitture
10 per cent	Quittance

At what price and how soon can you furnish?

Quadrants

Prefix code

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

Non-prefix code

A	1
B	0 1
C	0 0
D	0

00101

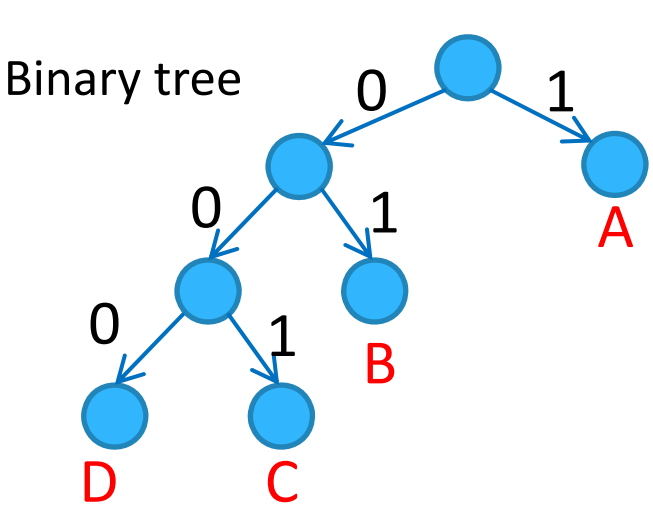
CB

No whole code word is a prefix of any other code word.

CADA
DDADA
DBDA
DBB
CAB

Prefix code

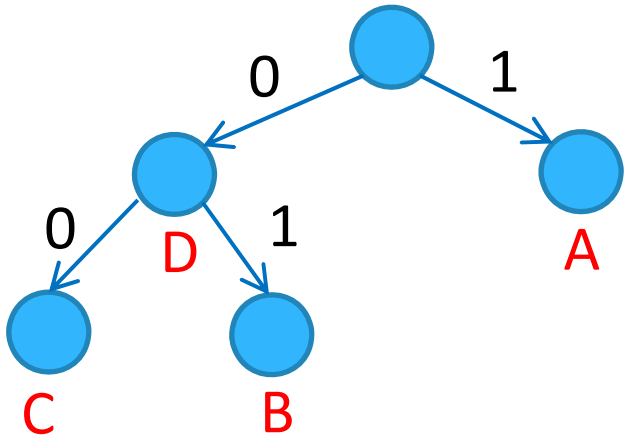
A	1
B	0 1
C	0 0 1
D	0 0 0



E.g. Huffman code

Non-prefix code

A	1
B	0 1
C	0 0
D	0

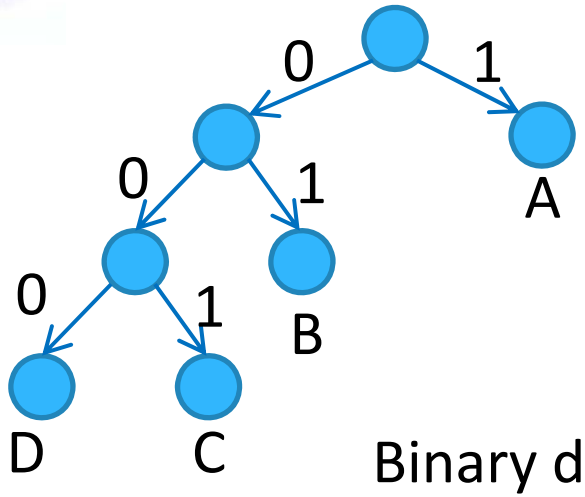


E.g. Morse code

Prefix code

- Heavily used in practical compression algorithms
- For example, part of JPEG, GZIP, etc.

Yes or No questions

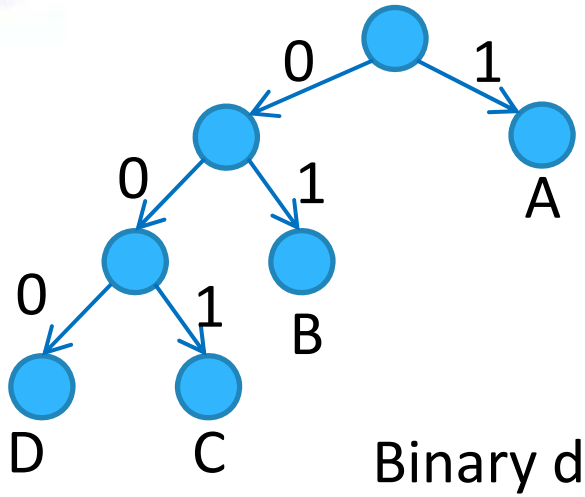


Binary decision tree

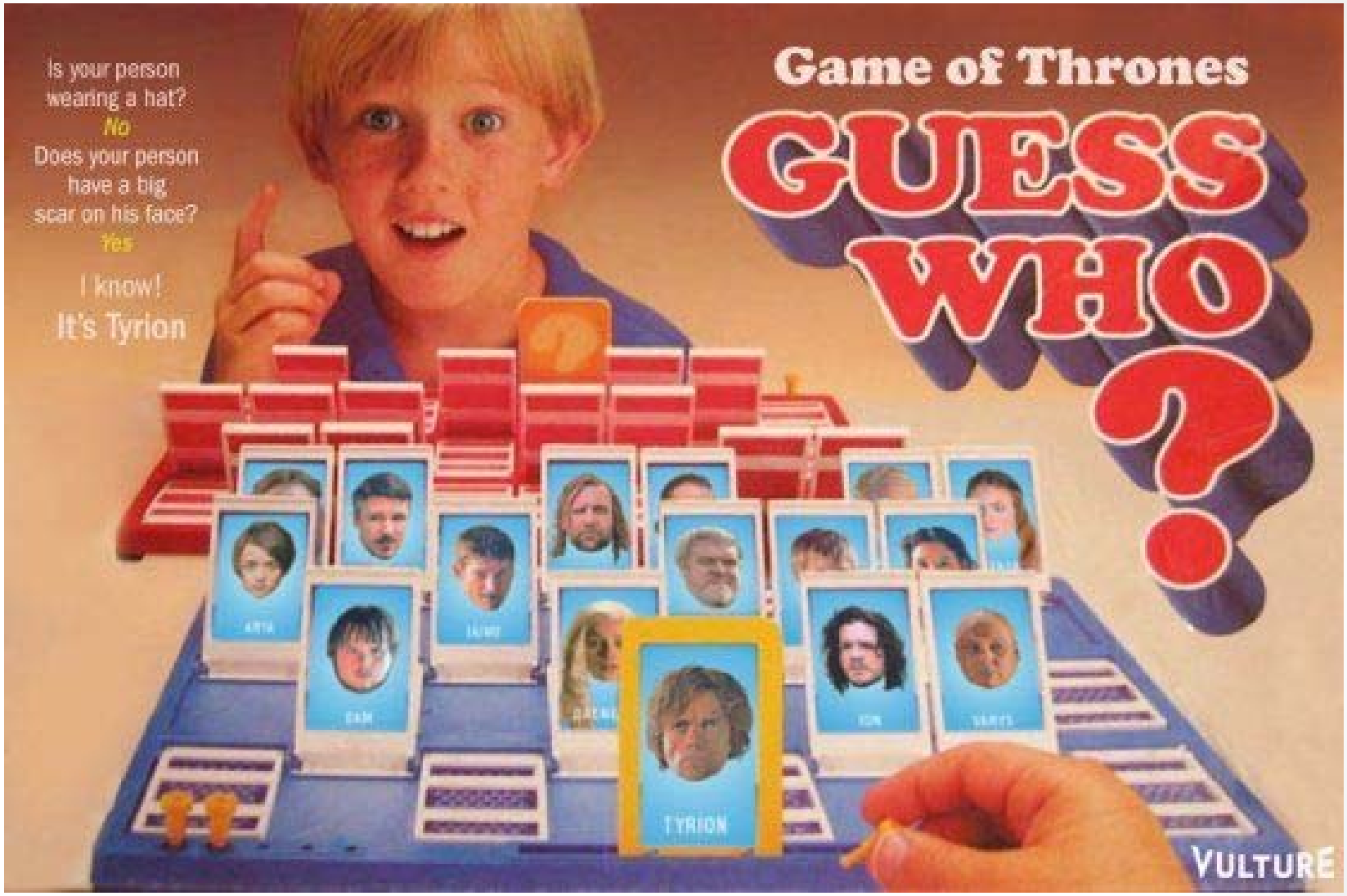
A	0.25	1	1 bit
B	0.25	0 1	2 bits
C	0.25	0 0 1	3 bits
D	0.25	0 0 0	3 bits

Average: 2.35 bits

Yes or No questions



Binary decision tree


Is your person wearing a hat?
No
Does your person have a big scar on his face?
Yes
I know!
It's Tyrion
Game of Thrones
GUESS WHO?
TYRION
VULTURE

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

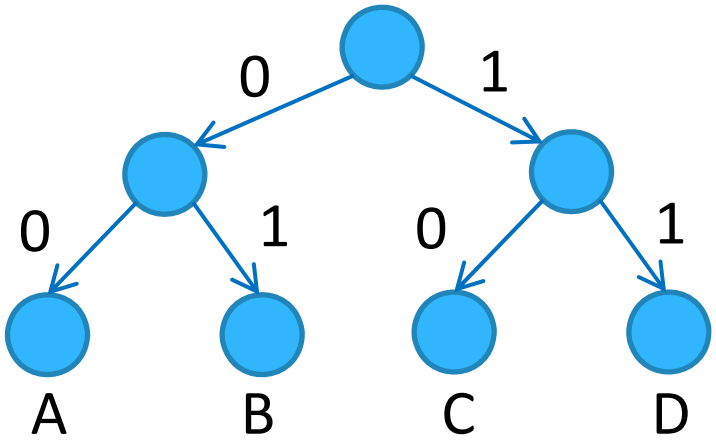
A. 1 bit
B. 1.5 bits
C. 1.75 bits
D. 2 bits
E. 2.25 bits

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

Yes or No questions

Binary decision tree



A	0.5
B	0.25
C	0.125
D	0.125

Can you come up with a good strategy?

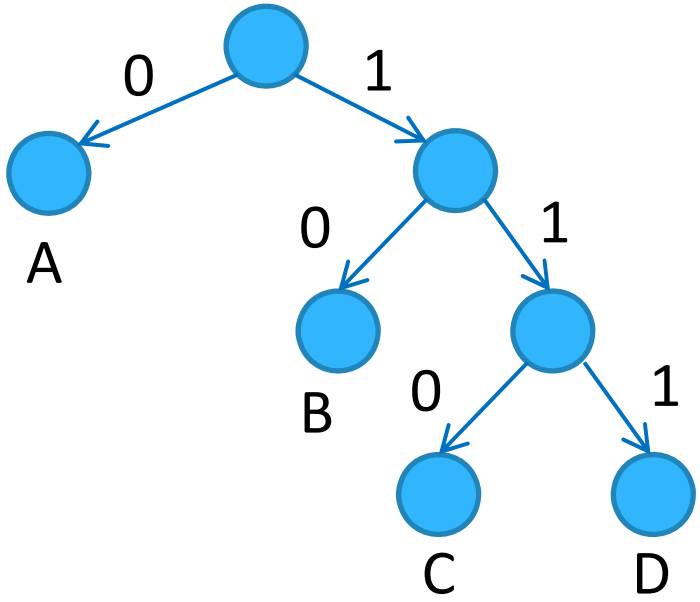
How many bits on average will it use?

Yes or No questions

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

A	0.5	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?

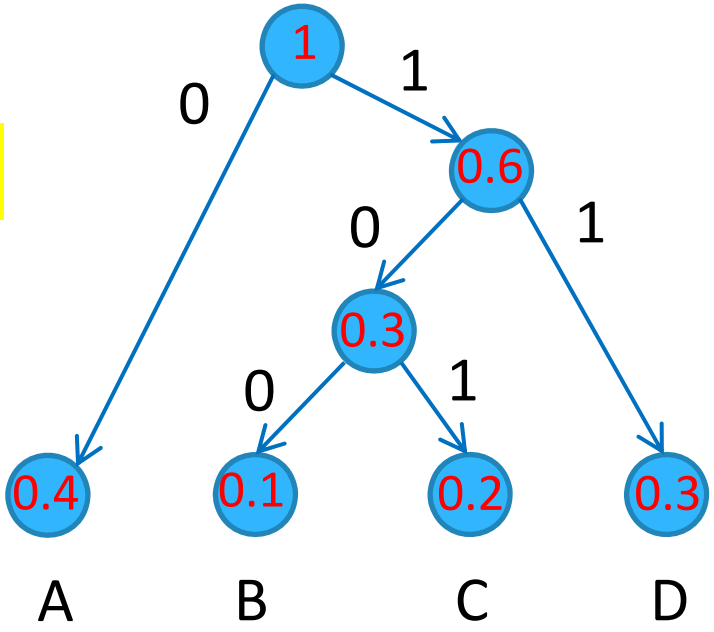
Yes or No questions

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

00101101010111011100011101010O …

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 …

00101101010111011100011101010 0 …

0001101001101000110001 …

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 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

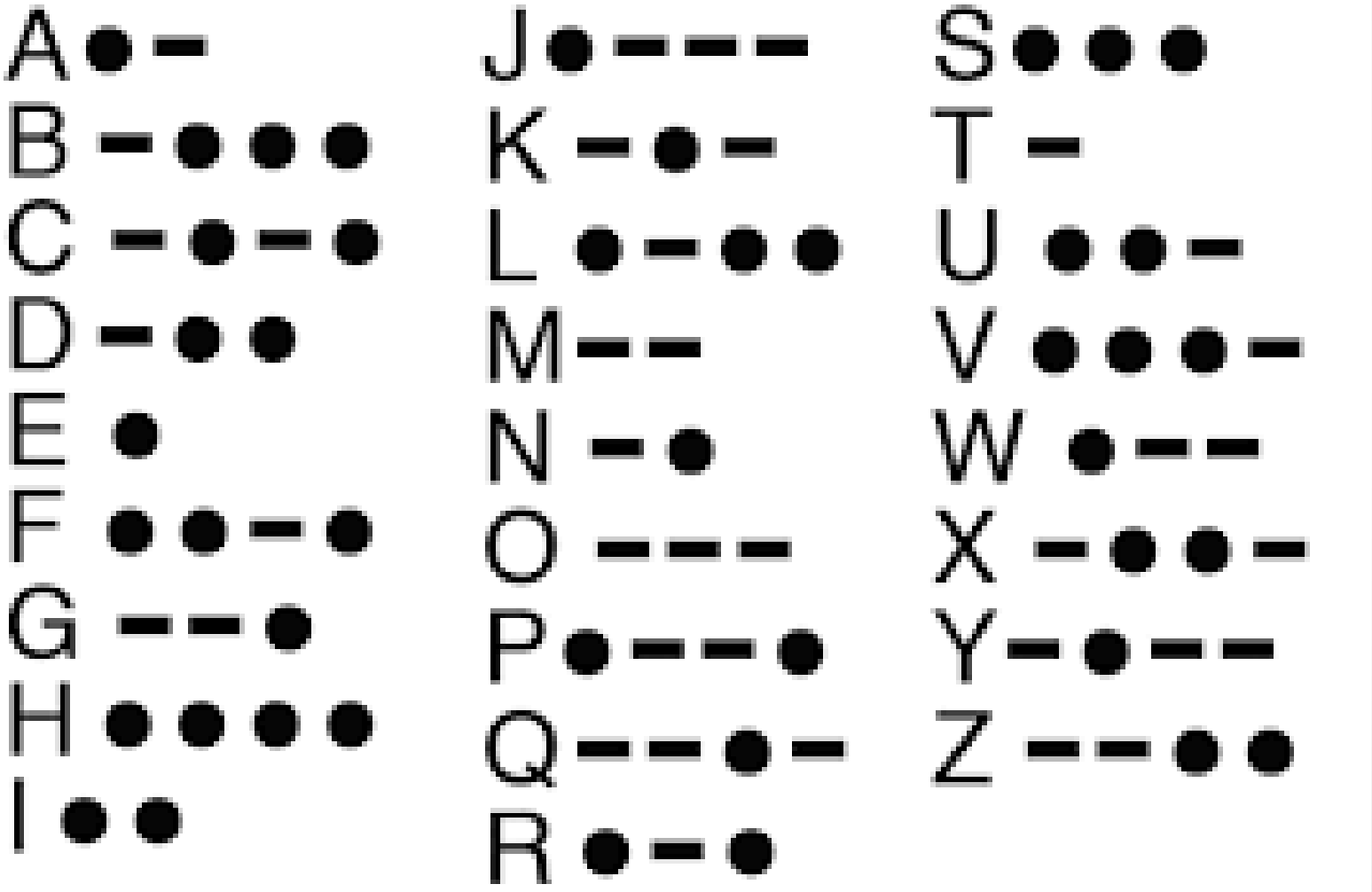
Average: 1.3 bits

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

Yes or No questions

A C D

Morse code



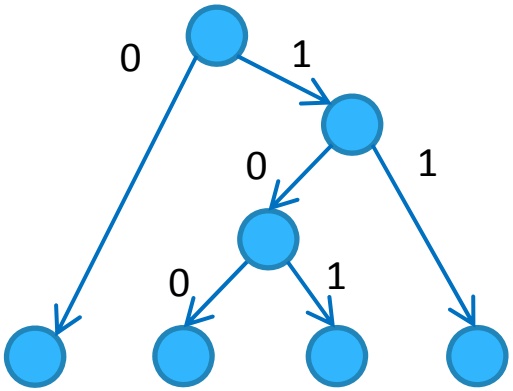
Prefix code

E.g. Huffman code

Non-prefix code

E.g. Morse code

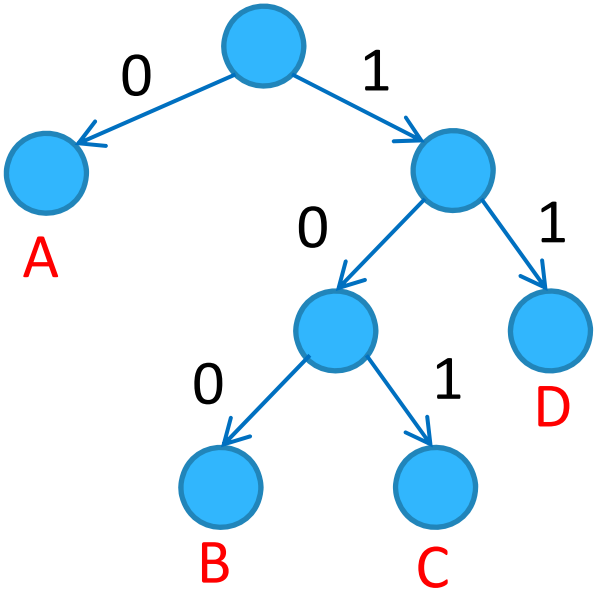
Binary tree



00101101010111011100011101010O ...

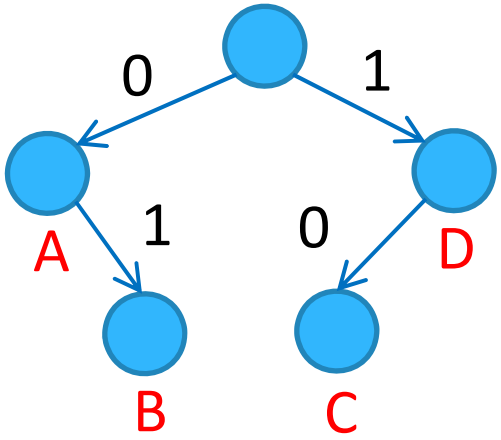
Prefix code

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



Non-prefix code

A	0
B	0 1
C	1 0
D	1



Morse code

A	·−	J	·−−−	S	···
B	−···	K	−·−	T	−
C	−·−·	L	·−··	U	··−
D	−··	M	−−	V	···−
E	·	N	−·	W	·−−
F	··−·	O	−−−	X	−··−
G	−−·	P	·−−·	Y	−·−−
H	····	Q	−−·−	Z	−−··
I	··	R	·−·		

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	Quamoclit
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	Quatemion
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	Quastor
Will be here until	Quebracho
Will be in	Quebrith
Immediate specifications, three months delivery	
5 per cent	Quitclaim
7½ per cent	Quiture
10 per cent	Quittance

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

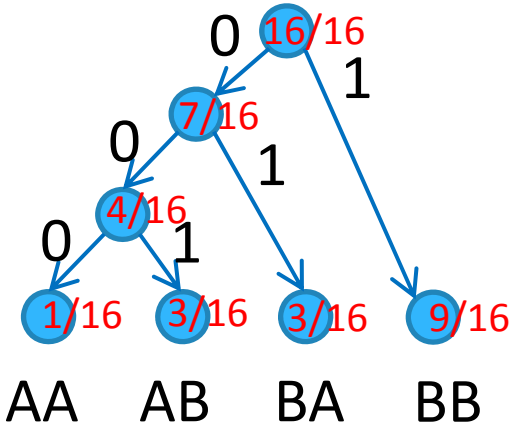
AA	1/16	0 0 0	3 bits
AB	3/16	0 0 1	3 bits
BA	3/16	0 1	2 bits
BB	9/16	1	1 bits

Average: 1.6875 bits

We are only considering independent events!

0.8438 bits / symbol

Yes or No questions



A	1/4
B	3/4

AAA	1/64	000000	6 bits
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

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

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

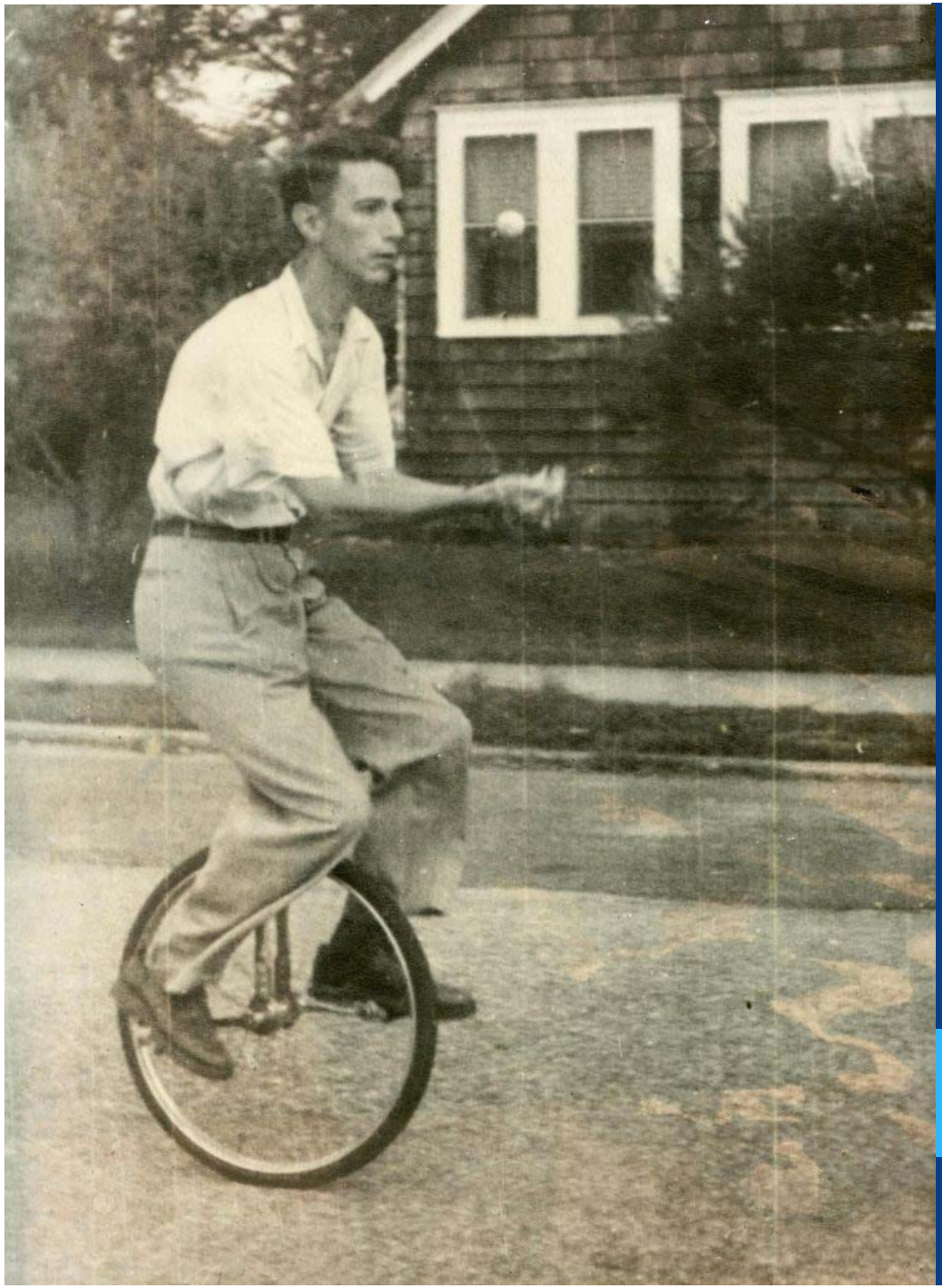
Entropy

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

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

Claude Shannon

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

Unique code for each possible message

The great power of this principle of selection is not hypothetical. It is certain that several of our eminent breeders have, even within a single lifetime, modified to a large extent their breeds of cattle and sheep. In order fully to realise what they have done it is almost necessary to read several of the many treatises devoted to this subject, and to inspect the animals. Breeders habitually speak of an animal's organisation as something plastic, which they can model almost as they please. If I had space I could quote numerous passages to this effect from highly competent authorities. Youatt, who was probably better acquainted with the works of agriculturalists than almost any other individual, and who was himself a very good judge of animals, speaks of the principle of selection as "that which enables the agriculturist, not only to modify the character of his flock, but to change it altogether. It is the magician's wand, by means of which he may summon into life whatever form and mould he pleases." Lord Somerville, speaking of what breeders have done for sheep, says: "It would seem as if they had chalked out upon a wall a form perfect in itself, and then had given it existence." In Saxony the importance of the principle of selection in regard to merino sheep is so fully recognised, that men follow it as a trade: the sheep are placed on a table and are studied, like a picture by a connoisseur; this is done three times at intervals of months, and the sheep are each time marked and classed, so that the very best may ultimately be selected for breeding.

What English breeders have actually effected is proved by the enormous prices given for animals with a good pedigree; and these have been exported to almost every quarter of the world. The improvement is by no means generally due to crossing different breeds; all the best breeders are strongly opposed to this practice, except sometimes among closely allied sub-breeds. And when a cross has been made, the closest selection is far more indispensable even than in ordinary cases. If selection consisted merely in separating some very distinct variety and breeding from it, the principle would be so obvious as hardly to be worth notice; but its importance consists in the great effect produced by the accumulation in one direction, during successive generations, of differences absolutely inappreciable by an uneducated eye—differences which I for one have vainly attempted to appreciate. Not one man in a thousand has accuracy of eye and judgment sufficient to become an eminent breeder. If gifted with these qualities, and he studies his subject for years, and devotes his lifetime to it with indomitable perseverance, he will succeed, and may make great improvements; if he wants any of these qualities, he will assuredly fail. Few would readily believe in the natural capacity and years of practice requisite to become even a skilful pigeon-fancier.

The same principles are followed by horticulturists; but the variations are here often more abrupt. No one supposes that our choicest productions have been produced by a single variation from the aboriginal stock. We have proofs that this is not so in several

Note: context matters as well


“Tomorrow, it will be over 50 degrees in La Jolla”