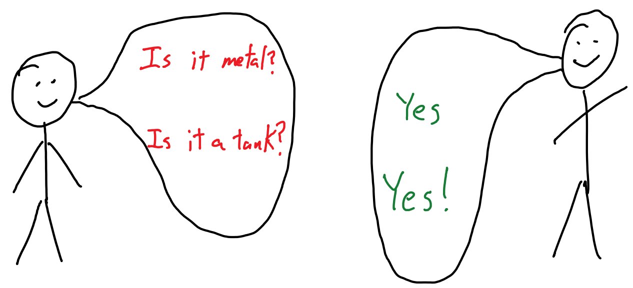
Lecture 13

Tuesday, April 01, 2014
5:43 PM

20 Questions



Is it a transform? No
Is it tangible? Yes
Is it metal? No
Is it edible? No
Is it a filter? No
Is it organic? Yes

Is it a material? No
Person? Yes
Alive? No
Fourier? No
Male? Yes
Did he create something for this class? Yes
Mathematician? Yes
Laplace? No
Euler? Yes

15 guesses:

If objects are picked from a set $\Omega = \{\text{monkey, ball, apple, tank} \dots\}$

How large can Ω be? $|\Omega| \le 2^{20} = 1{,}048{,}576$

This many combinations of yes-no.
Each question cuts the set in half at best.

Suppose $\Omega = \{\text{apple, orange, pear, banana}\}$

How many question? 2

What if $\text{Prob(apple)} = \frac{1}{4}$
$\text{Prob(orange)} = \frac{1}{8}$
$\text{Prob(pear)} = \frac{1}{2}$
$\text{Prob(banana)} = 1/8$

Minimize the average number of questions needed.

$\xrightarrow{\text{yes (1)}}$ pear

Is it a pear? $\xrightarrow{\text{yes (1)}}$ pear

Is it a pear? $\xrightarrow{\text{no (0)}}$ Is it an apple? $\xrightarrow{\text{yes (1)}}$ apple

Is it an apple? $\xrightarrow{\text{No (0)}}$ Is it an orange? $\xrightarrow{\text{Yes (1)}}$ orange

Is it an orange? $\xrightarrow{\text{No (0)}}$ banana

Let X be the answer and $L(X)$ be the number of questions asked to determine X.

$$\text{Average} = E\,L(X) = \sum_{x\in\Omega} p_X(x)\,L(x) = \tfrac{1}{4}\cdot 2 + \tfrac{1}{8}3 + \tfrac{1}{2}\cdot 1 + \tfrac{1}{8}\cdot 3 = 1.75$$

Ideas for designing optimal question scheme?

- Find a first question that is "yes" as close to .5 a possible.
- Same as above but sort first and cut the list for a partition.

Bound: $E\,L(X) \le \lceil \log_2 |\Omega| \rceil$ ← Achieved by using question to split in half. i.e. Each object requires the same # of questions.

Question Scheme → Code

Sequence of answers:

	Code		
Apple :	01	C(apple)	← "Code word for apple"
Orange :	001	C(orange)	
Pear :	1	C(pear)	
Banana :	000	C(banana)	

0010000110 1

O B A P A

Notice that the system of questions is implicit in the code.

What are the requirements of the code?

- Unique (Necessary but not sufficient)

  Example: Apple : 0
  Orange : 1
  Pear : 00
  Banana : 01

- Prefix Free ("Prefix code"), $C(x)$ is not a prefix of $C(y)$ $\forall y \neq x$

Data Compression (lossless):

Let $X_1, X_2, \ldots$ be a sequence of information that needs to be stored.

Example: - Letters of text

- The output of a quantizer

Let $C(x) \in \{0,1\}^*$ be a code used to store the information.

Binary sequences of any length.

$$C(X_1)\,C(X_2)\,C(X_3)\ \ldots$$

1.) Non-singular

$C(x) \neq C(y)\ \forall x \neq y$

Can decode if there is punctuation.

$C(X_1), C(X_2), \ldots$

2.) Uniquely Decodable

Can decode without punctuation.

3.) Prefix Codes

Defined above.

Instantaneous.

Prefix $\Rightarrow$ Uniquely Decodable $\Rightarrow$ Non-singular

This is what we need.

We will focus on this.

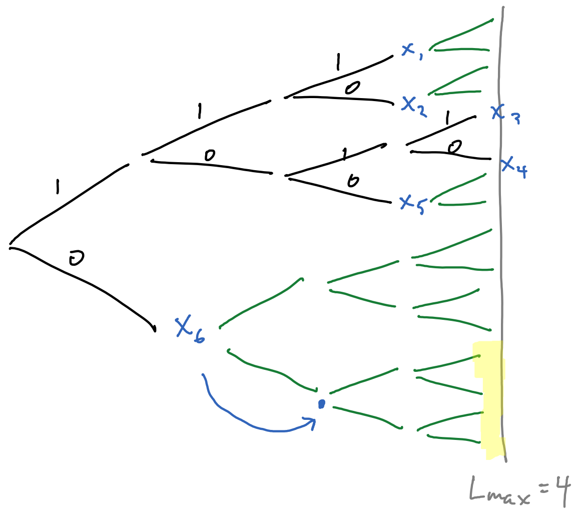
No loss of efficiency

Bounds on Prefix codes:

Kraft inequality: For prefix code: $\sum_{x \in \Omega} 2^{-L(x)} \leq 1$

Notice that any prefix code can be embedded in a binary tree.

Leaves are the objects $x \in \Omega$



$C(x_1) = 111$
$C(x_2) = 110$
$C(x_3) = 1011$
$C(x_4) = 1010$
$C(x_5) = 100$
$C(x_6) = 0$

$2^{L_{max}}$ phantom leaves.

Accounting	Phantom leaves	
x_1	2	$= 2^{L_{max} - L(x_1)}$
x_2	2	$= 2^{L_{max} - L(x_2)}$
x_3	1	$\vdots$
x_4	1	
x_5	2	
x_6	8	

$$\sum_{x \in \Omega} 2^{L_{max} - L(x)} \le 2^{L_{max}}$$

Summary:

1.) Prefix $\Rightarrow \sum_{x \in \Omega} 2^{-L(x)} \le 1$

2.) $\sum_{x \in \Omega} 2^{-L(x)} \le 1 \Rightarrow$ There exists a prefix code with lengths $L(x)$.

3.) If Kraft inequality is loose, the prefix code is inefficient for any distribution.

Entropy: Information Theory, 1948, Claude Shannon

"bit" first appears in Shannon's papers.

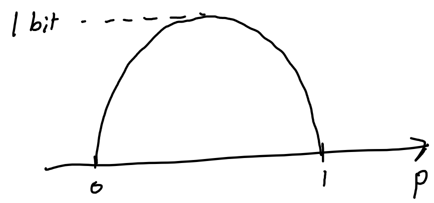
$$H(X) = E \log_2 \frac{1}{P_X(X)} = \sum_{x \in X} p_X(x) \log_2 \frac{1}{p_X(x)}$$

Units are bits.

Example 1: $X = \begin{cases} 0, & \text{w.p. } \frac{1}{2} \\ 1, & \text{w.p. } \frac{1}{2} \end{cases}$ $\qquad P_X(x) = \begin{cases} \frac{1}{2}, & x \in \{0, 1\} \\ 0, & \text{else} \end{cases}$

$$H(X) = \left(\tfrac{1}{2}\right) \log_2 \frac{1}{1/2} + \tfrac{1}{2} \cdot \log_2 \frac{1}{1/2} = \log_2 2 = 1 \text{ bit.}$$

Example 2: $p_X(x) = \begin{cases} p, & x = 1 \\ 1-p, & x = 0 \end{cases}$ $\qquad H(X) = p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p}$



Optimization Problem:

Find $L(x)$ that satisfy Kraft inequality and minimize $EL(X)$.

Integer

$\sum_{x \in \Omega} p_X(x) L(x)$

If we relax the problem, let $\hat{L}(x)$ not be integer, then easy.

Lower bound.

Can solve using legrange mult.: $\hat{L}^*(x) = \log_2 \frac{1}{p_X(x)}$ ← If we round up,
- integer
- satisfy Kraft

$$E \hat{L}^*(X) = H(X)$$

$$\Rightarrow E L(X) \geq H(X)$$

With integer constraints.

Optimal code: $E L(X) < H(X) + 1$ bits.