WHY LOG PROBABILITIES ?
CODING THEORY
RELATIVE ENTROPY

WANT TO SEND SYMBOL X ON CHANNEL

X	$P(X=x_i)$	$-\log P(x_i)$
x_1	$\frac{1}{2}$	1
x_2	$\frac{1}{4}$	2
x_3	$\frac{1}{8}$	3
x_4	$\frac{1}{8}$	3

BITS

MEASURE OF SURPRISE

$-\log 1 = 0$ NO SURPRISE
$-\log 0 = \infty$ INFINITE "

$-\log 2^{i} = i$ BITS

ENTROPY EQUALS EXPECTED SURPRISE

$$H(X) := \sum_i p(x_i) \log_2 \frac{1}{p(x_i)}$$

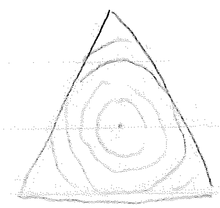
$$= \frac{1}{2}\cdot 1 + \frac{1}{4}\cdot 2 + \frac{1}{8}\cdot 3 + \frac{1}{8}\cdot 3$$

$$= 1\frac{3}{4}$$

2 DIM

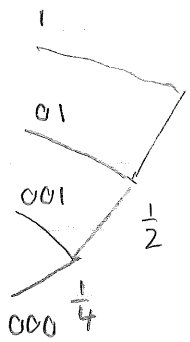
$-p\log p - (1-p)\log(1-p)$

3 DIM



HUFFMAN CODE

x_1	$\frac{1}{2}$
x_2	$\frac{1}{4}$
x_3	$\frac{1}{8}$
x_4	$\frac{1}{8}$



$\nwarrow 1$

$\swarrow 0$

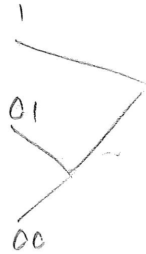
LOOP

PICK SMALLEST TWO
COMBINE BOTH INTO ONE
SUM THEIR PROBS

ENTROPY = EXPECTED CODELENGTH

X

X_1	$\frac{1}{3}$
X_2	$\frac{1}{3}$
X_3	$\frac{1}{3}$



$H(X) = \log_2 3 = 1.58$ BITS

EXPECTED CODELENGTH $= \frac{1}{3}(1+2+2)$

$= \frac{5}{3}$

$= 1.66$ BITS

CODE: ASSIGNS SYMBOLS A BITSTRING (CODEWORD)
- ANY SEQUENCE OF CODEWORDS MUST BE UNIQUELY DECODABLE

$$L(C) = \sum_i p(x_i)\, \ell_C(x_i)$$ ← EXPECTED CODE LENGTH

C: CODE

$\ell_C(x_i)$: LENGTH OF CODEWORD FOR x_i

OPTIMAL CODE C^*
- MINIMUM $L(C)$

THM: $H(X) \le L(C^*) \le H(X) + 1$

THM: HUFFMAN CODES ARE OPTIMAL

MORE INFO:
FIRST FIVE CHAPTERS OF COVER & THOMAS

RELATIVE ENTROPY

PROBABILITY VECTORS

$$\Delta(\bar{p}, \bar{q}) = \sum_i p_i \ln \frac{p_i}{q_i}$$

SYMBOL USED FOR DIVERGENCES

$$= \sum_i p_i \ln \frac{1}{q_i} - \overbrace{\sum p_i \ln \frac{1}{p_i}}^{H(\bar{p})}$$

EXPECTED CODELENGTH OF BEST CODEBOOK FOR $\bar{q}$

EXPECTED CODE LENGTH OF BEST CODE BOOK FOR $\bar{p}$

BOTH EXPECTATIONS ARE WRT $\bar{p}$

$$\Delta(\bar{p}, (\tfrac{1}{n})) = \sum_i p_i \ln \frac{p_i}{\frac{1}{n}}$$

$$= \sum_i p_i \ln p_i + \sum p_i \ln n$$

$$= \ln n - H(\bar{p})$$

$$\geq 0$$

$= 0$ AT CORNERS OF SIMPLEX

MAPLE PLOTS

$\Delta(\bar{p}, \bar{q})$ (VAR: $\bar{p}$) NOT TOO STEEP AT BOUNDARY

$\Delta(\bar{p}, \bar{q})$ (VAR: $\bar{q}$) VERY STEEP

BARRIERS FOR SIMPLEX