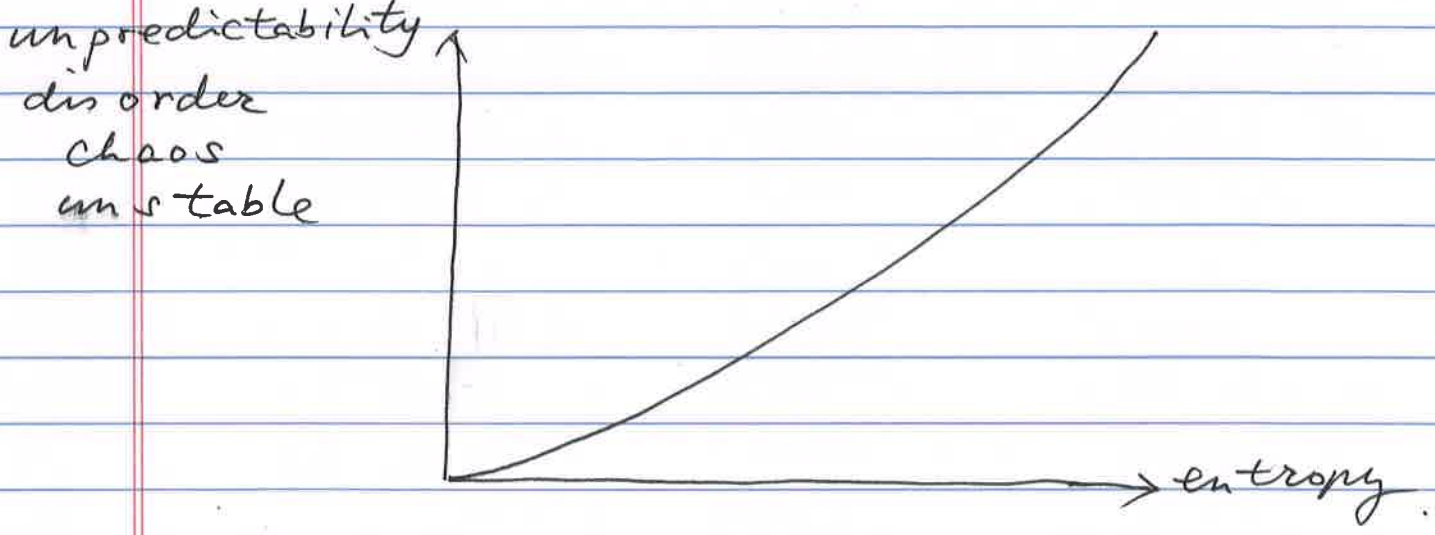
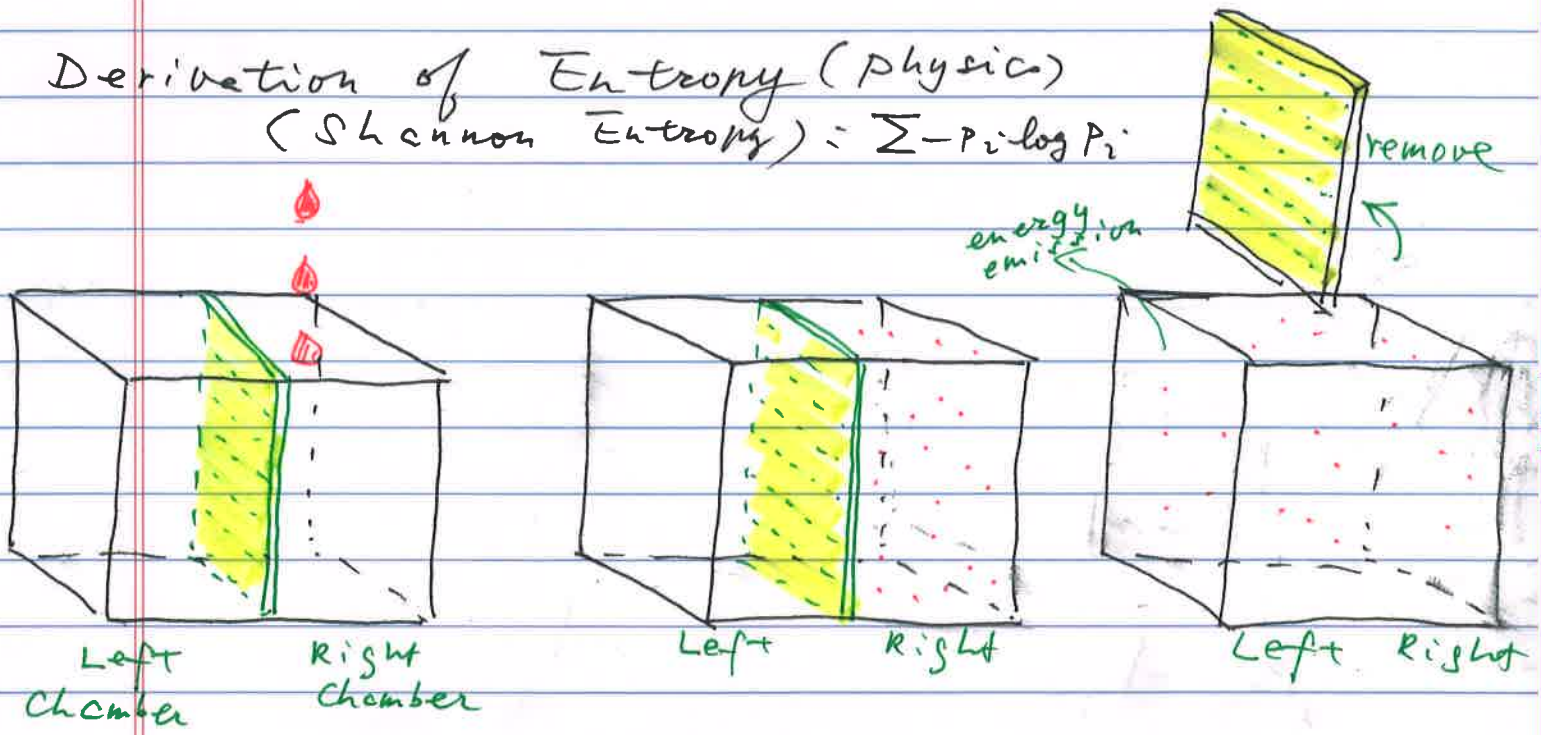
Lesson 045 : Information Science.



Derivation of Entropy (physics)
(Shannon Entropy) : $\sum -p_i \log p_i$



thermodynamics.

$P(X)$: the probability that a Pigment molecule appears in Right Chamber

Entropy : $-1 \cdot \log_2 1 - 0 \cdot \log_2 0 = 0.$

$\Rightarrow \quad -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \cdot \log_2 \frac{1}{2} = 1.$

Information Science

(gain : loss) Broker	Entropy	Deterministic
100% : 0%	0	✓
75% : 25%	$-\frac{3}{4}\log_2\frac{3}{4} - \frac{1}{4}\log_2\frac{1}{4}$	✗
25% : 75%	$-\frac{1}{4}\log_2\frac{1}{4} - \frac{3}{4}\log_2\frac{3}{4}$	✗
0% : 100%	0	✓
50% : 50%	1	✗

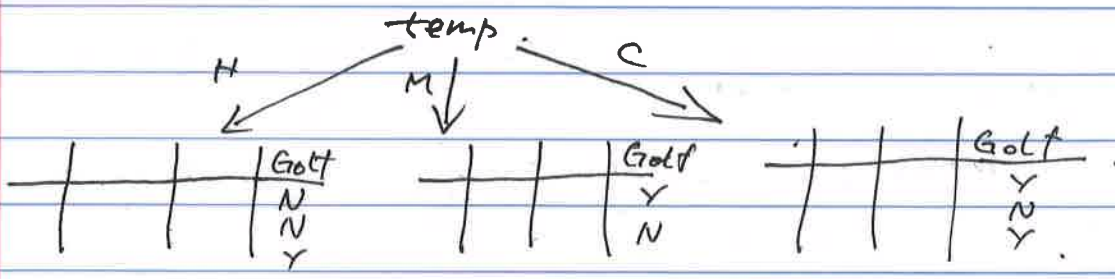
$$\text{Predictability} \propto \frac{1}{\text{Entropy}}.$$

Entropy and Information Gain (IG).

Outlook	Temp	Humidity	Windy	play golf
Rainy	Hot	High	F	No
Rainy	Hot	H	T	No
Overcast	Hot	H	F	Yes
Sunny	Mild	H	F	Yes
Sunny	Cool	Normal	F	Yes
Overcast	Cool	N	T	No
Rainy	Cool	N	T	Yes
Rainy	Mild	H	F	No

$$E(\text{play Golf}) = \left(-\frac{1}{2}\cdot\log_2\left(\frac{4}{8}\right) - \frac{1}{2}\cdot\log_2\frac{4}{8}\right)$$

$$= 1.$$



$$E(\text{temp, Golf}) = \frac{3}{8}\cdot E(\text{temp} = \text{Hot, Golf}) +$$
$$\frac{2}{8}\,E(\text{temp} = \text{Mild, Golf}) +$$
$$\frac{3}{8}\,E(\text{—— cool, Golf}).$$

$$IG = E(\text{golf}) - E(\text{temp, golf}).$$

Cross — Entropy

$$H(p, q) = -\sum_i p_i \log_2(q_i)$$

p: target, q: prediction

Cross-entropy is used to measure the difference between two probability distribution.

Cross-Entropy is widely used as a Loss-function when optimizing classification model.

Cross - Entropy vs.

Kullback - Leibler (KL)
Divergence (Relative Entropy)

$$KL(P/Q) = \sum_i p_i \log \frac{p_i}{q_i}$$

$$= \sum_i (-p_i \log q_i + p_i \log p_i)$$

$$= \sum_i p_i \log \frac{1}{q_i} - \sum p_i \log \frac{1}{p_i}$$

$$= H(P,Q) - H(P)$$

$$\Rightarrow \underbrace{H(P,Q)}_{\text{cross entropy}} = \underbrace{H(P)}_{\text{shannon entropy}} + \underbrace{KL(P/Q)}_{\text{KL entropy}}.$$