* CONDITIONAL ENTROPY

X, Y

Entropy of Y conditioned on X

$$H(Y|X) = \sum_{i=0}^{M-1} P(X = s_i) \cdot H(Y|X = s_i)$$

$$= \sum_{i=0}^{M-1} Pr(X=s_i) \sum_{j=0}^{N-1} Pr(Y=r_j \mid X=s_i) \times \log_2 \frac{1}{P(Y=r_j \mid X=s_i)}$$

$$H(Y \mid X=s_i)$$

$$= \sum_{i=0}^{M-1} Pr(s_i) \sum_{j=0}^{N-1} Pr(r_j \mid s_i) \cdot \log_2 \frac{1}{Pr(r_j \mid s_i)}$$

$$= \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \frac{Pr(s_i, r_j)}{Pr(s_i)\, Pr(r_j \mid s_i)} \cdot \log_2 \left(\frac{1}{Pr(r_j \mid s_i)} \right)$$

$$= \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \Pr(s_i, r_j) \log_2 \frac{1}{\Pr(r_j | s_i)}$$

$H(Y|X)$
Conditional Entropy of Y given X

$H(Y|X)$ = uncertainty remaining in Y having observed X

$$H(Y|X) + H(X)$$

$$= \underbrace{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} Pr(s_i, r_j) \log_2 \frac{1}{Pr(r_j|s_i)}}_{H(Y|X)} + \underbrace{\sum_{i=0}^{M-1} Pr(s_i) \log_2 \frac{1}{Pr(s_i)}}_{H(X)}$$

$$\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} Pr(r_j, s_i) \log_2 \left(\frac{1}{Pr(s_i)} \right)$$

$$\Rightarrow H(Y|X) + H(X)$$

$$= \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} Pr(s_i, r_j) \log_2 \frac{1}{Pr(r_j|s_i)}$$

$$+ \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} Pr(s_i, r_j) \log_2 \frac{1}{Pr(s_i)}$$

$$= \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} Pr(s_i, r_j) \log_2 \left(\frac{1}{\underbrace{Pr(r_j|s_i) Pr(s_i)}_{= Pr(s_i, r_j)}} \right)$$

$$= \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P_r(s_i, r_j) \log_2 \left(\frac{1}{P_r(s_i, r_j)} \right)$$

$$H(X,Y)$$

uncertainty in Y given X

uncertainty in X

joint uncertainty of X & Y

$$H(Y|X) + H(X) = H(X,Y)$$

$$H(X|Y) + H(Y) = H(X,Y)$$

Example:

$P_{ij} = P(X = s_i, Y = r_j)$

		r_0	r_1	r_2	r_3
		$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$
$\frac{1}{4}$	s_0	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$
$\frac{1}{4}$	s_1	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$
$\frac{1}{4}$	s_2	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
$\frac{1}{4}$	s_3	$\frac{1}{4}$	0	0	0

$P_{03} = P(X = s_0, Y = r_3) = \frac{1}{32}$

$$P_r(Y = r_1 \mid X = s_0) = \frac{P_r(s_0, r_1)}{P_r(s_0)} = \frac{1/16}{1/4} = \frac{1}{4}.$$

$P(Y|X)$.

$$P(Y \mid X = s_0)$$

P_{0j}

$$P(Y = r_j \mid X = s_0)$$

$$= \frac{Pr(Y = r_j, X = s_0)}{Pr(X = s_0)}$$

$$Pr(Y = r_0 \mid X = s_0) = \frac{1/8}{1/4} = \frac{1}{2}$$

$$Pr(Y = r_1 \mid X = s_0) = \frac{1/16}{1/4} = \frac{1}{4}$$

$$Pr(Y = r_2 \mid X = s_0) = \frac{1}{8}$$

$$= Pr(Y = r_3 \mid X = s_0).$$

$$H(Y|X=s_0) = H\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\right)$$

$$= \frac{1}{2}\log_2 2 + \frac{1}{4}\log_2 4 + \left(\frac{1}{8}\log_2 8\right) \times 2$$

$$= 1.75 \text{ bits}$$

$$H(Y|X=s_1) = H\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{8}, 0\right)$$

$$= 1.75$$

$$H(Y|X=s_2) = H\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$$

$$= \frac{1}{4}(\log_2 4) \times 4$$

$$= 2$$

$$H(Y|X=s_3) = H(1, 0, 0, 0)$$

$$= 0$$

$$H(Y|X) = \sum_{i=0}^{M-1} \Pr(X=s_i) \, H(Y|X=s_i)$$

$$= \frac{1}{4} \times 1.75 + \frac{1}{4} \times 1.75$$

$$+ \frac{1}{4} \times 2 + \frac{1}{4} \times 0$$

$$= \frac{22}{16} = \frac{11}{8} = 1.375 \text{ bits}$$

$H(Y|X)$

$H(X) = 2 \text{ bits} = H\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$

$H(X) + H(Y|X)$

$= 2 + 1.375 = 3.375$

Joint Entropy of X, Y

$= H(X, Y)$

$$H(Y) + H(X|Y) = H(X,Y)$$

check

$$\boxed{H(X) + H(Y|X) = H(X,Y)}$$

information in X

Conditional Entropy

information in Y given X

Joint information in X, Y.