# Entropy

$$H = \sum_i p_i \log_2 \frac{1}{p_i}$$

# Conditional Entropy

$$H(X \mid Y) = \sum_{x,y} p(x,y) \log \frac{1}{p(x \mid y)}$$

# Mutual Information

$$I(X;Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$$

# Independence Bound on Entropy

$$H(X_1, ..., X_n) \leq \sum_i H(X_i)$$

# Fano's Inequality

$$P_e \geq \frac{1 - H(X \mid Y)}{\log |X|}$$

# Data Processing Inequality

If $X \to Y \to Z$ for some transformation $\to$, then $I(X;Y) \geq I(X;Z)$.

# Mutual Information Distance

$$D(X,Y) = H(X,Y) - I(X;Y)$$

# Kullback-Leibler Distance

$$D_{KL}(p \parallel q) = \sum_{x} p(x) \log \frac{p(x)}{q(x)}$$

# Markov Process Entropy

Average of the entropy of each state weighted by occupancy probability.