2.9.3 马尔可夫信源的信息熵

一、在状态S_j下发出一个符号的条件熵：

$$H(X|u=S_j)=-\sum_{k=1}^{q}P(a_k|S_j)\log P(a_k|S_j)$$

二、一般马尔可夫信源熵：平均符号熵的极限值

$$H_\infty=H_\infty(X)=\lim_{N\to\infty}H_N(X)$$

$$=\lim_{N\to\infty}\frac{1}{N}H(X_1X_2\cdots X_N)$$

三、m阶马尔可夫（记忆长度为m）极限熵

=

m阶条件熵

极限熵定理：

对任意离散平稳信源，若$H_1(X)<\infty$，则有

$$\lim_{N\to\infty} H_N(X) = \lim_{N\to\infty} H(X_N \mid X_1 X_2 \cdots X_{N-1})$$

$$= H(X_{m+1} \mid X_1 X_2 \cdots X_m)$$

$$= H_{m+1}$$

该定理证明参见：姜丹，《信息论与编码》

四、齐次、遍历的m阶马尔可夫信源熵

对于齐次遍历的马尔可夫链，其状态S_j可由（$a_{k1},\ldots,a_{km}$）唯一确定

$$P(a_{k_{m+1}} \mid a_{k_m}, \cdots a_{k_1}) = P(a_{k_{m+1}} \mid S_j)$$

对上式两端同时取对数，并求统计平均再取负，得

$$左端=-\sum_{k_1,\cdots k_{m+1},S_j} p(a_{k_{m+1}},a_{k_m},\cdots a_{k_1};S_j)\log P(a_{k_{m+1}}\mid a_{k_m},\cdots a_{k_1})$$

$$=-\sum_{k_1,\cdots k_{m+1},S_j} p(a_{k_{m+1}},a_{k_m},\cdots a_{k_1})\log P(a_{k_{m+1}}\mid a_{k_m},\cdots a_{k_1})$$

$$=H(a_{k_{m+1}}\mid a_{k_m},\cdots a_{k_1})=H_{m+1}$$

$$右端=-\sum_{k_1,\cdots k_{m+1},S_j} p(a_{k_{m+1}},a_{k_m},\cdots a_{k_1};S_j)\log P(a_{k_{m+1}}\mid S_j)$$

$$=-\sum_{k_1,\cdots k_{m+1},S_j} p(a_{k_m},\cdots a_{k_1};S_j)p(a_{k_{m+1}}\mid S_j)\log P(a_{k_{m+1}}\mid S_j)$$

$$=\sum_{k_{m+1}}\sum_{S_j} p(S_j)p(a_{k_{m+1}}\mid S_j)\log P(a_{k_{m+1}}\mid S_j)$$

$$=\sum_{S_j} p(S_j)H(X\mid S_j)$$

齐次、遍历的m阶马尔可夫信源熵

$$H_{m+1} = \sum_{k_{m+1}} \sum_{S_j} p(S_j) p(a_{k_{m+1}} \mid S_j) \log P(a_{k_{m+1}} \mid S_j)$$

$$= \sum_{S_j} p(S_j) H(X \mid S_j)$$

说明：

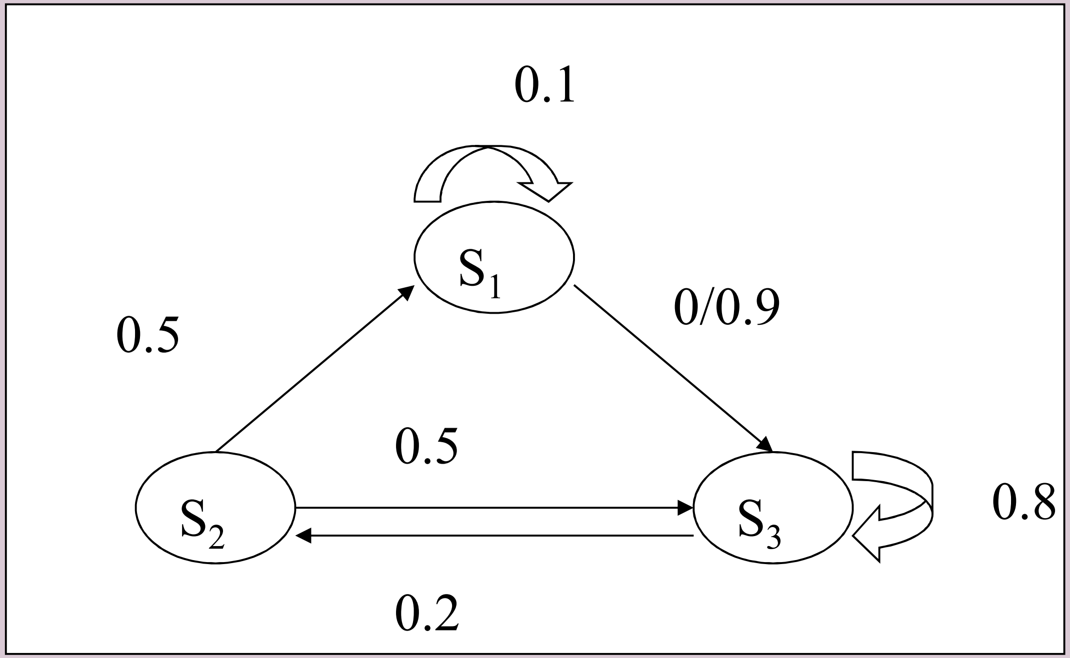
$H(X|S_j)$：信源处于状态S_j时发出一个符号的平均不确定性；

$p(S_j)$：马尔可夫链的平稳分布，满足

$$p(S_j)=\sum_{S_k\in S}p(S_k)p(S_j|S_k)\qquad \sum_{S_k\in S}p(S_j)=1$$

m阶马尔可夫信源的极限熵表示信源稳定后每发一个符号所能提供的平均信息量

例：如下页图所示为三状态马尔可夫信源，计算信息熵



解：转移概率矩阵为

$$P = \begin{array}{c} S_1 \\ S_2 \\ S_3 \end{array} \begin{bmatrix} 0.1 & 0 & 0.9 \\ 0.5 & 0 & 0.5 \\ 0 & 0.2 & 0.8 \end{bmatrix}$$

(1) 求平稳分布P(S1)、P(S2)、P(S3)

$$\begin{cases} P(S_1)=0.1P(S_1)+0.5P(S_2) \\ P(S_2)=\ 0.2P(S_3) \\ P(S_3)=0.9P(S_1)+0.5P(S_2)+0.8P(S_3) \\ P(S_1)+P(S_2)+P(S_3)=1 \end{cases}$$

解 得 ：
$$\begin{cases} P(S_1)=\dfrac{5}{59} \\ P(S_2)=\dfrac{9}{59} \\ P(S_3)=\dfrac{45}{59} \end{cases}$$

(2) 计算S_j状态下每输出一个符号的平均信息量

$$H(X|S_1)$$

$$= 0.1\log\frac{1}{0.1} + 0.9\log\frac{1}{0.9} = 0.468996 \text{ 比特/符号}$$

$$H(X|S_2)$$

$$= 0.5\log\frac{1}{0.5} + 0.5\log\frac{1}{0.5} = 1 \text{比特/符号}$$

$$H(X|S_3)$$

$$= 0.2\log\frac{1}{0.2} + 0.8\log\frac{1}{0.8} = 0.721928 \text{比特/符号}$$

(3) 求信息熵

$$H_\infty = \sum_{j=1}^{3} p(S_j) H(X \mid S_j)$$

＝0.742910　比特/符号