

第一章 补充习题及参考答案

(题号标*号者为必做题)

1.试述信息、消息、信号及其相互关系。

解答

消息(message)是通信系统传输的具体对象，其终极目的在于通过消息的传递使受信者获知信息。

信息(information)则是指受信者在收到消息前对消息的不确定性。消息之中包含信息，但不一定就是信息。

消息是具体的，信息是抽象的。

不同的消息含有不同数量的信息，同一消息对不同的接受对象来说信息的多少也不同。在数据通信中，Message 被译成报文，也称为 Data(数据)。

信号是消息的载体和具体表现形式，消息(信息)需转化为与具体传输媒体特定的信道特性相适应的信号形式方能传输。

信号是随时间变化的某种物理量，通常是电压、电流或者光波，还可以是电荷或磁通。通信系统一般使用电或光形式的电磁波信号。

信息分为模拟信息和数字信息，转换成相应的信号也分为模拟信号与数字信号。模拟量是连续的，数字量是离散的。无论模拟信号还是数字信号，在传输过程中根据信道特性的不同(也有模拟信道与数字信道之分)，是可以相互转换的。

2.叙述信息量的定义及单位。什么是平均信息量？信息速率的定义是什么？

解答

香农的信息量公式表述为，一个消息所载的信息量 I ，等于它所表示的事件发生的概率 P 的倒数的对数。

也就是说，若一个消息 x_i 出现的概率为 $P(x_i)$ ，则这一消息所含的信息量为

$$I(x_i) = \log \frac{1}{P(x_i)} = -\log P(x_i)$$

信息量满足相加性。

信息量公式中对数以2为底时，信息量单位为 bit(比特)；对数以e为底时，单位为 nat(奈特)；对数以10为底时，单位为 Hartley(哈特莱)。

平均信息量被称为熵，是每个符号所含信息量的统计平均值。M 种符号的平均信息量为

$$H(x) = -\sum_{i=1}^m P(x_i) \log P(x_i)$$

信息速率是单位时间传输的信息量，它被定义为信息源每秒钟发出的消息所包含的符号数目与符号平均信息量的乘积。

3*. 已知英文字母 e 和 q 的出现概率分别为 $P(e)=0.105$ ， $P(q)=0.001$ ，分别求它们的信息量。

解答

已知 e 和 q 的出现概率分别为 $P(e)=0.105$ ， $P(q)=0.001$ ，则它们的信息量分别是

$$I(e) = -\log_2 P(e) = -\log_2 0.105 = 3.25 \text{ (bit)}$$

$$I(o) = -\log_2 P(q) = -\log_2 0.001 = 9.97 \text{ (bit)}$$

4. 设某信源产生 a、b、c、d 四个符号。若各符号的出现是相互独立的，且其出现概率分别为 1/2、1/4、1/8、1/8，试求该信源的平均信息量。

解答

a、b、c、d 的平均信息量

$$\begin{aligned} H &= P(a) \log_2 \frac{1}{P(a)} + P(b) \log_2 \frac{1}{P(b)} + P(c) \log_2 \frac{1}{P(c)} + P(d) \log_2 \frac{1}{P(d)} \\ &= \frac{1}{2} \log_2 2 + \frac{1}{4} \log_2 4 + \frac{1}{8} \log_2 8 + \frac{1}{8} \log_2 8 \\ &= \frac{1}{2} + \frac{1}{2} + \frac{3}{8} + \frac{3}{8} = 1.75 \text{ bit} \end{aligned}$$

6*. 一个离散信号源每毫秒发出 4 种符号中的一个，各相互独立符号出现的概率分别为 0.4、0.3、0.2、0.1，求该信号源的平均信息量与信息速率。

解答

该信号源的平均信息量

$$\begin{aligned} H &= -P_1 \log_2 P_1 - P_2 \log_2 P_2 - P_3 \log_2 P_3 - P_4 \log_2 P_4 \\ &= -0.4 \log_2 0.4 - 0.3 \log_2 0.3 - 0.2 \log_2 0.2 - 0.1 \log_2 0.1 \\ &= 0.53 + 0.52 + 0.46 + 0.33 = 1.84 \text{ bit} \end{aligned}$$

信息速率

$$R = r \cdot H = 1000 \times 1.84 = 1840 \text{ bps}$$

7*. 看懂下列计算信息量的例题的两种解法，并思考为何存在误差，什么情况下误差会趋于零？

例题：一消息由 0、1、2、3 四种符号组成，各符号出现概率分别为 3/8、1/4、1/4 和 1/8。消息总长 57 个符号，其中 0 出现 23 次，1 出现 14 次，2 出现 13 次，3 出现 7 次。用上述二种方法求该消息的信息量。

解法一：（直接求和）

$$\begin{aligned} I &= -\sum_{i=1}^4 n_i \log P_i = -23 \log_2 \frac{3}{8} - 14 \log_2 \frac{1}{4} - 13 \log_2 \frac{1}{4} - 7 \log_2 \frac{1}{8} \\ &= 32.55 + 28 + 26 + 21 = 108.55(\text{bit}) \end{aligned}$$

解法二：（H 是平均信息量，N 为消息所含符号总数）

$$\begin{aligned} I &= H \cdot N = -\sum_{i=1}^4 P_i \log P_i \cdot 57 \\ &= \left(-\frac{3}{8} \log_2 \frac{3}{8} - \frac{1}{4} \log_2 \frac{1}{4} - \frac{1}{4} \log_2 \frac{1}{4} - \frac{1}{8} \log_2 \frac{1}{8} \right) \times 57 \\ &= 1.9056 \times 57 = 108.62(\text{bit}) \end{aligned}$$

解答

两种解法的结果存在误差，原因是该消息的序列还不够长，各符号出现的频次与所给出的概率并不完全一致。随着序列长度的增加，误差将趋于零。

8. 讲解数据通信过程时，提到了**低通滤波**这个术语，请分别解释相关的频谱、频带、通带、阻带、低通滤波、高通滤波、带通滤波、带阻滤波、梳齿滤波等概念（其中部分术语会在第三章、第四章、第五章、第八章中出现）。

解答

频谱：正如第三章所介绍的，对于一个复杂信号，可用傅立叶分析将它分解为许多不同频率的正弦分量，而每一正弦分量则以它的振幅和相位来表征。将各正弦分量的振幅与相位分别按频率高低次序排列成频谱。

频带：复杂信号频谱中各分量的频率理论上可扩展至无限，但因原始信号的能量一般集中在频率较低范围内，在工程应用上一般忽略高于某一频率的分量。频谱中该有效频率范围称为该信号的频带。

通带：衰减很小的频带。

阻带：衰减很大的频带。

低通滤波器：通带与阻带交界处的频率称为截止频率。低通滤波器的通带由零频率起一直到截止频率为止，其阻带在截止频率高的一侧。

高通滤波器：起始频率根据设计要求决定，其通带在截止频率的高的一侧。。

带通滤波器：是以两个有限截止频率之间的频段为通带。

带阻滤波器：其通带正好与带通滤波器的通带相反。

梳齿滤波器：它有许多按一定频率间隔相间排列通带与阻带。