Домашнее задание

Лев Довжик, M3439

Вариант №62

Задача 1.1

Докажем три основные свойства метрики для расстояние Хэмминга:

1) Неотрицательность

Очевидно, две строки отличаются в неотрицательном числе позиций.

2) Симметричность

$\rho(a,b) = \sum\limits_{i=1}^{n} |a_i \neq b_i| = \sum\limits_{i=1}^{n} |b_i \neq a_i| = \rho(b,a)$

3) Неравенств треугольника

(1) $a_i = b_i$ В силу неторицательности $\begin{cases} |a_i \neq c_i| \geq 0 \\ |c_i \neq b_i| \geq 0 \end{cases} \Rightarrow |a_i \neq c_i| + |c_i \neq b_i| \geq 0 = |a_i \neq b_i|$

(2) $a_j \neq\ b_j$ Тогда c_j обязано отличаться хотя бы от одного из них. Значит $\left[\begin{array}{l} |a_j \neq c_j| = 1 \\ |c_j \neq b_j| = 1 \end{array}\right. \Rightarrow$

$|a_j \neq c_j| + |c_j \neq b_j| \geq 1 = |a_j \neq b_j|$

Соединив (1) и (2), получаем: $\rho(a,b) = \sum\limits_{i=1}^{n} |a_i \neq b_i| \leq \sum\limits_{i=1}^{n} (|a_i \neq c_i| + |c_i \neq b_i|) = \rho(a,c) + \rho(c,b)$

Задача 1.2

Для эквивалентность декодирования по МП и МАВ необходимо чтобы $p(c_m|y) = f(p(y|c_m))$, где $f(t)$ — неубывающая функция. И в то же время $p(y|c_m) = g(p(c_m|y))$, где $g(t)$ — также неубывающая функция. Если f обратима, что $g = f^{-1}$ и наоборот.

$$p(c_m|y) = \frac{p(c_m \wedge y)}{p(y)} = \frac{p(y|c_m)p(c_m)}{p(y)}$$

Заметим, что если $p(c_m)$ будет константой не зависящей от m, то $p(c_m|y) = \alpha p(y|c_m)$, где $\alpha = \dfrac{C}{p(y)} > 0$. А значит $f(t) = \alpha t$ не убывает и обратима. Однако такое ограничение на $p(c_m)$ означает, что все кодовые слова распределены равновероятно и $C = \dfrac{1}{M}$.

В таком случае мы получаем, что при равновероятном распределении кодовых слов декодирование по МАВ и МП эквивалентны.

Задача 1.3

Распишем $1 - P_c = \sum\limits_y p(y) - \sum\limits_m \sum\limits_{R_m} p(c_m)p(y|c_m) = \sum\limits_y \sum\limits_m p(y|c_m)p(c_m) - \sum\limits_m p(c_m) \sum\limits_{R_m} p(y|c_m) = \sum\limits_m p(c_m) \sum\limits_y p(y|c_m) - \sum\limits_m p(c_m) \sum\limits_{R_m} p(y|c_m) = \sum\limits_m p(c_m) \left(\sum\limits_y p(y|c_m) - \sum\limits_{R_m} p(y|c_m) \right) = \sum\limits_m p(c_m) \sum\limits_{T_m} p(y|c_m)$, где T_m - все слова, которые декодируются не в c_m.

Тогда получаем, что $1 - P_c = \sum\limits_m p(c_m)p(\widehat{m} \neq m|c_m) = P_e$. Значит, максимизируя P_c, мы уменьшаем P_e. Докажем, что декодирование по МАВ максимизирует P_c.

$P_c = \sum\limits_m \sum\limits_{R_m} p(c_m)p(y|c_m) = \sum\limits_m \sum\limits_{R_m} p(y \wedge c_m) = \sum\limits_m \sum\limits_{R_m} p(y)p(c_m|y)$. Так как мы максимизируем каждое $p(c_m|y)$, то мы максимизируем всю сумму.

Задача 1.4

Рассмотрим слова длины n. Пусть x и y отправленное и полученное слово соответственно. Так же пусть $\rho(x,y) = d$. Посчитаем $p(y|x) = \prod\limits_{i=1}^{n} p(y_i|x_i)$

В силу известного расстояния между словами, d множителей будут равны $\dfrac{p_0}{q-1}$, а $n-d$ равны $1 - p_0$.

$\prod\limits_{i=1}^{n} p(y_i|x_i) = \left(\dfrac{p_0}{q-1}\right)^d (1-p_0)^{n-d} = \left(\dfrac{p_0}{(q-1)(1-p_0)}\right)^d (1-p_0)^n$

Так как $(1-p_0)^n$ константа, то максимизации $p(y|x)$ будет соответствовать минимизация d при $0 < \dfrac{p_0}{(q-1)(1-p_0)} < 1$. Так как при любых осмысленных параметрах канала левая часть неравенства выполняется всегда, достаточно проверить правую часть.

$\dfrac{p_0}{(q-1)(1-p_0)} < 1 \Rightarrow p_0 < q - 1 + p_0 - qp_0 \Rightarrow 1 - q < -qp_0 \Rightarrow p_0 < \dfrac{q-1}{q}$

Задача 1.5

Пусть мы хотим гарантированно исправлять t ошибок. Рассмотрим декодирование по минимуму расстояния Хэмминга. Перейдём в метрическое пространство всех слов построенной на метрике Хэмминга.

Пусть $R_r(x) = \{y|\rho(x,y) \leq r\}$. Чтобы добиться желаемого подобные шары радиуса t, построенные вокруг каждого кодового слова, должны не пересекаться. Это возможно, если расстояние между любыми кодовыми словами было больше $2t$, то есть хотя бы $2t+1$, так как расстояние Хэмминга является натуральным числом.

То есть $d \geq 2t+1$ или же $t \leq \left\lfloor \dfrac{d-1}{2} \right\rfloor$ так как мы работаем в натуральных числах.

Теперь рассмотрим шары радиуса $d-1$, построенные вокруг каждого кодового слова. Очевидно, что в каждом таком шаре будет только одного кодовое слово — его центр. Но так же в каждом сфере будут все слова содержащие до $d-1$ возможных ошибок включительно относительно соответствующего кодового слова, а значит, совершив не более чем столько ошибок мы попадёт подобный шар соответствующего исходного слова и сможем сообщить об ошибке т.к других кодовых слов в этом шаре нет.

Задача 1.8

Рассмотрим слова длины n. Пусть x и y отправленное и полученное слово соответственно. Пусть k позиций было стёрто, а m — исказились. m и будет нашим расстоянием Хэмминга.

Если вероятность стирания ε, а замены p_0. То $p(y|x) = \varepsilon^k p_0^m (1-\varepsilon-p_0)^{n-k-m}$. Немного преобразовав, получаем: $p(y|x) = \left(\dfrac{p_0}{1-\varepsilon-p_0}\right)^m \varepsilon^k (1-\varepsilon-p_0)^{n-k}$

Так как $\varepsilon^k(1-\varepsilon-p_0)^{n-k}$ константа при заданном y, то максимизации $p(y|x)$ будет соответствовать минимизация d при $0 < \dfrac{p_0}{1-\varepsilon-p_0} < 1$.

При любых осмысленных параметрах канала левая часть неравенства выполняется всегда, достаточно проверить правую часть.

$$\frac{p_0}{1-\varepsilon-p_0} < 1 \Rightarrow p_0 < 1-\varepsilon-p_0 \Rightarrow p_0 < \frac{1-\varepsilon}{2}$$

Зададимся вопросом количества исправляемых стираний. Очевидно, что если произошло d стираний, то они могли задеть все позиции в которых два каких-то кодовых слова отличались между собой и код не сможет их различить.

С другой стороны, если стираний было $d-1$. То так как остальные $n-d+1$ у y совпадают с x, то среди них найдётся такая позиция i, что $x_i \neq \widehat{x}_i$, для любого $\widehat{x} \neq x$. Значит такой код и может исправить не более $d-1$ стираний.

Теперь зафиксируем количество стираний s. Если $s \geq d$, то по вышеупомянутым соображениям код не сможет исправить ошибку. В случае же $s < d$ минимальное расстояние по нестёртым символам будет $d - s$ в худшем случае, а как известно из задания 1.5 такой код может исправить до $\left\lfloor \dfrac{d-s-1}{2} \right\rfloor$ ошибок.