§ 3. Упаковка шаров

Минимальное расстояние кода C в A^n есть

$$d_{\min}(C) = d(C) = \min_{x,y \in C} d_H(x, y).$$

Примеры. 1. Минимальное расстояние кода повторения длины n, очевидно $= n$.

2. Код с проверкой четности имеет минимальное расстояние $= 2$.

3. Для обобщенного кода Рида–Соломона длины 27 минимальное расстояние, как было показано, $= 21$.

4. Для $[7, 4]_2$-кода Хэмминга минимальное расстояние $= 3$, для его расширения $= 4$.

5. $[4, 2]_3$-код Хэмминга имеет минимальное расстояние $= 3$.

Теорема 1. Для кода C в A^n эквивалентны следующие утверждения:

1) алгоритм CB_e всегда правильно исправляет e или менее ошибочных символов;

2) для любых различных $x, y \in C \quad B_e(x) \cap B_e(y) = \varnothing$;

3) минимальное расстояние кода C удовлетворяет неравенству $d_{\min}(C) \geqslant 2e + 1$.

Доказательство. (1→2) Пусть $z \in B_e(x)$ для некоторого $x \in C$; по предположению, z декодируется x; следовательно, z не может принадлежать $B_e(y)$ для какого-либо другого $y \in C$.

(2→3) Докажем, что если 3) не выполняется, то и 2) не выполняется. Пусть $d_{\min}(C) < 2e + 1$, т.е. $d_{\min}(C) = d \leqslant 2e$. Выберем $x, y \in C$ так, чтобы $d_H(x, y) = d$. Если $d \leqslant e$, то уже $x \in B_e(x) \cap B_e(y)$; если же $d > e$, то выберем z так, чтобы он совпадал с x в e координатных позициях, в которых x и y различны (это возможно, так как $e < d$), и с y в остальных позициях. Тогда $d_H(y, z) = e$, $d_H(x, z) = d - e \leqslant e$. Таким образом $z \in B_e(x) \cap B_e(y)$.

(3→2) очевидно: если бы $z \in B_e(x) \cap B_e(y)$, то по неравенству треугольника $d_H(x, y) \leqslant d_H(x, z) + d_H(y, z) \leqslant 2e$.

(2→1) также очевидно: в силу 2) любой шар радиуса e: $B_e(z)$ может содержать лишь одно кодовое слово, что и доказывает корректность алгоритма CB_e.

Код C, удовлетворяющий условиям теоремы 1, называется *кодом, исправляющим e ошибок*. Можно заметить, что в теореме 1-е условие алгоритмическое, 2-е геометрическое, а 3-е алгебраическое,

так что можно легко переключаться с одной точки зрения на другую.

Задача. Пусть $f \geqslant e$. Доказать, что следующие утверждения эквивалентны для кода C в A^n:

1) алгоритм CB_e всегда правильно исправляет e меньшее число вхождений ошибочных символов и не дает ошибки декодирования при f или меньшем числе вхождений ошибочных символов;

2) для любых различных $x, y \in C \quad B_f(x) \cap B_e(y) = \varnothing$;

3) минимальное расстояние кода C, $\quad d_{\min}(C) \geqslant e + f + 1$.

Код C, удовлетворяющим условиям, сформулированным выше, называется *кодом, исправляющим e и обнаруживающим f ошибок*

Теорема 2. (Граница Хэмминга). Если C — m-ричный код длины n, исправляющий e ошибок, то его размер

$$K \leqslant m^n \Big/ \sum_{i=0}^{e} \binom{n}{i} (m-1)^i.$$

Доказательство. Согласно теореме 1, шары радиуса e с центрами в C не пересекаются; таким образом справедливо неравенство между общим объемом этих шаров и размером всего кодового пространства A^n:

$$|C| \cdot |B_e(*)| \leqslant |A^n|;$$

так как $|C| = K$, $|B_e(*)| = \mathrm{Vol}_m^n(e)$, $|A^n| = m^n$, получаем

$$K \cdot \mathrm{Vol}_m^n(e) \leqslant m^n .$$

Граница Хэмминга также называется границей сферической упаковки. Код, на котором достигается эта граница, называется совершенным.

Теорема 3. (Граница Гилберта-Варшамова). Существует m-ричный код C, исправляющий e ошибок, длины n, такой, что

$$K \geqslant m^n \Big/ \sum_{i=0}^{2e} \binom{n}{i} (m-1)^i.$$

Доказательство. Пусть A^n — кодовое пространство. Будем строить коды $C_1, \ldots, C_i, \ldots$ последовательно, с помощью алгоритма:

1) $C_1 := \{x_1\}$ для произвольного $x_1 :\in A^n$;

i) $S_i := \bigcup_{j=1}^{i-1} B_{d-1}(x_j)$, и если $S_i \neq A^n$, то $x_i :\in A^n \setminus S_i$, $C_i := C_{i-1} \cup \{x_i\}$.

Пока выполнено неравенство

$$|C_i| \cdot |B_{d-1}(*)| < |A^n|,$$

шаг i можно выполнять, так что алгоритм завершится, когда это неравенство станет ложным.

Примеры. 1. Для двоичного кода длины 90, исправляющего 2 ошибки, $m = 2,\ e = 2,\ n = 90;$ граница Хэмминга:

$$K \leqslant \frac{2^{90}}{\mathrm{Vol}_2^{90}(2)} = \frac{2^{90}}{4096} = 2^{78};$$

граница Гилберта-Варшамова:

$$K \geqslant \frac{2^{90}}{\mathrm{Vol}_2^{90}(4)} = \frac{2^{90}}{2676766} \approx 4.62 \cdot 10^{20}.$$

2. Для троичного кода длины 8, исправляющего 2 ошибки, $m = 3,\ e = 2,\ n = 8;$ граница Хэмминга:

$$K \leqslant \frac{3^8}{\mathrm{Vol}_3^8(2)} = \frac{6561}{129} \approx 50.9;$$

граница Гилберта-Варшамова:

$$K \geqslant \frac{3^8}{\mathrm{Vol}_3^8(4)} = \frac{6561}{1697} \approx 3.9.$$

Задача. Определить, существует ли код со следующими параметрами:

1) $m = 2,\ n = 5,\ K = 6;$
2) $m = 2,\ n = 6,\ K = 9;$
3) $m = 3,\ n = 4,\ K = 9;$
4) $m = 3,\ n = 8,\ K = 51.$