Просеминар по математической логике и информатике 20 ноября 2009 года

Кодирование

1. Требуется отгадать задуманное число от 1 до 1000, задавая вопросы, на которые можно отвечать «да» и «нет».

а) За какое минимальное число вопросов можно гарантировать отгадывание?

б) …если список вопросов нужно составить заранее?

в) …если на один из вопросов разрешается дать неверный ответ?

г) …если на один из вопросов из заранее составленного списка разрешается дать неверный ответ?

Во всех случаях предлагается указать (возможно более близкие) верхние и нижние оценки (за сколько вопросов заведомо можно гарантировать отгадывание и за сколько точно нет гарантированного способа).

2. а) Каждой букве некоторого алфавита соответствует последовательность из N нулей и единиц, её *кодовое слово*, причём разные буквы закодированы по-разному. Какое максимальное число букв может быть в алфавите?

б) Пусть мы дополнительно хотим, чтобы любые два кодовых слова отличались по крайней мере в двух позициях. (Это значит, что ошибка в одном бите не может остаться незамеченной.) Какое максимальное число таких кодовых слов можно найти (при данном N)?

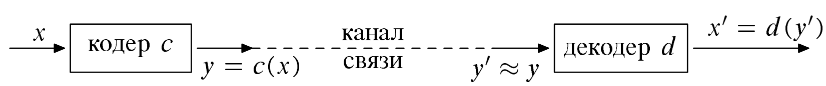
3. Определим *расстояние Хэмминга* на двоичных словах длины n как число позиций, в которых различаются биты.

а) Выполнено ли «неравенство треугольника» для расстояния Хэмминга?

б) Булевый куб можно рассматривать как граф: рёбра соединяют вершины на расстоянии 1. Сколько существует кратчайших путей в этом графе, соединяющих две вершины на расстоянии k?

в) Сколько существует вершин, отстоящих от данной на расстояние k?

г) Существует ли в этом графе гамильтонов цикл (проходящий по всем вершинам по одному разу)?



4. Кодер c отображает k-битовое слово x в его код, N-битовое слово $y = c(x)$, который передаётся (возможно, с искажениями) по каналу связи; декодер d должен из пришедшего слова y' восстановить исходное слово x.

а) Каким свойством должен обладать кодер c, чтобы существовал декодер, исправляющий одиночные ошибки (в y' один из битов может отличаться от y).

б) Какое максимальное число кодовых слов можно найти с таким свойством при $N = 3$?

в) Докажите, что при любом N максимальное число кодовых слов с таким свойством не больше $2^N/(N+1)$.

5. Докажите, что если $N + 1$ есть степень двойки, то оценка точна.

[Указание. Соответствующая конструкция называется *кодом Хэмминга* и выглядит так. Пусть, скажем, $N = 7$, и кодовое слово состоит из семи битов $x_1, \ldots, x_7$. Запишем в таблицу двоичные числа от 1 до 7; каждая строка таблицы соответствует одной из переменных x_i:

$$\begin{array}{cccc} x_1 & 0 & 0 & 1 \\ x_2 & 0 & 1 & 0 \\ x_3 & 0 & 1 & 1 \\ x_4 & 1 & 0 & 0 \\ x_5 & 1 & 0 & 1 \\ x_6 & 1 & 1 & 0 \\ x_7 & 1 & 1 & 1 \end{array}$$

В качестве кодовых будем рассматривать те слова $x_1 \ldots x_7$, которые после умножения справа на эту матрицу дают строку из трёх нулей. Другими словами, должны быть равны нулю «контрольные суммы» $x_4 \oplus x_5 \oplus x_6 \oplus x_7 = 0$, $x_2 \oplus x_3 \oplus x_6 \oplus x_7 = 0$ и $x_1 \oplus x_3 \oplus x_5 \oplus x_7 = 0$ (здесь $\oplus$ — сложение по модулю 2).]

6. а) Пусть имеются три N-битовых слова x, y, z. Докажите, что можно выбрать два из них, которые отличаются не более чем в $(2/3)N$ битов.

б) Что можно сказать о максимальном количестве кодовых слов, если мы хотим иметь возможность декодировании, разрешая портить до 90% битов? до 50% битов? до 49% битов?

7. Пусть имеются N-битовые слова $x_1, \ldots, x_k$, любые два из которых отличаются более чем в половине позиций. Докажите, что $k \leq N + 1$. (Мораль: коррекция более чем 25% ошибок невозможна, если только кодовых слов не очень мало.)

8. Покажите, что существует код, удлиняющий все слова не более чем в 100 раз и исправляющий до 1% ошибок.

9. Будем рассматривать слова в p-ичной системе счисления (вместо двоичных; здесь p-некоторое простое число). Покажите, что при $k < p$ можно кодировать слова длины k словами длины p, причём расстояние между кодовыми словами будет не меньше $p - k$. (Указание. Число корней многочлена не превосходит его степени.)