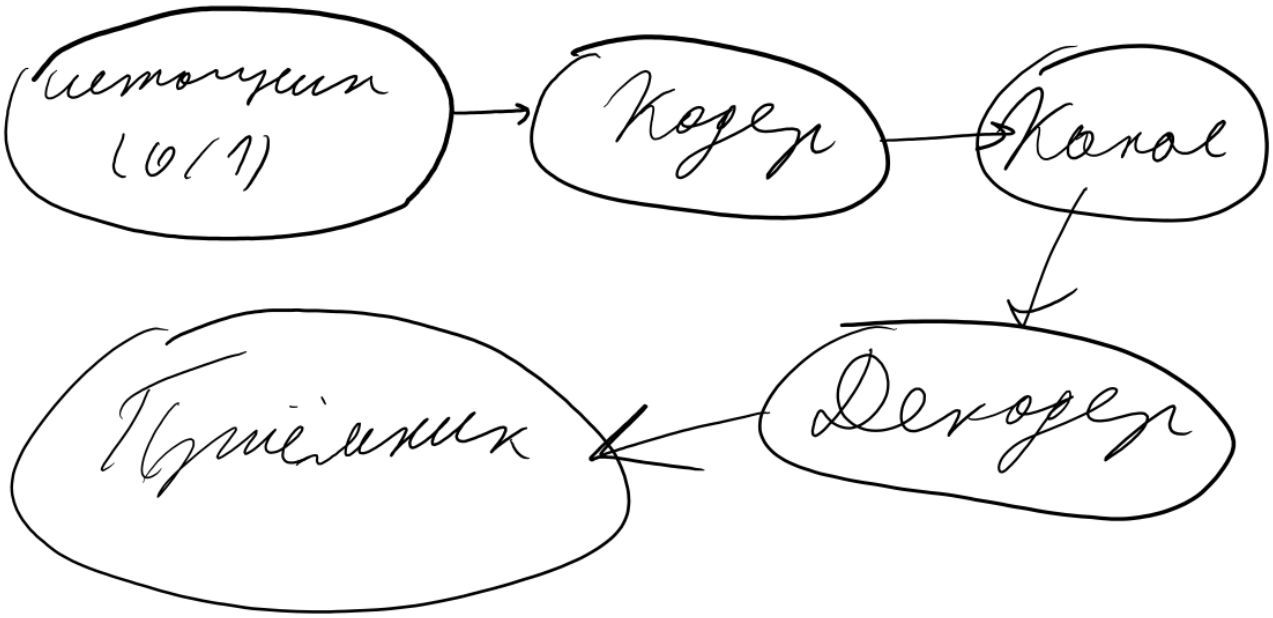
1 Лекция 1

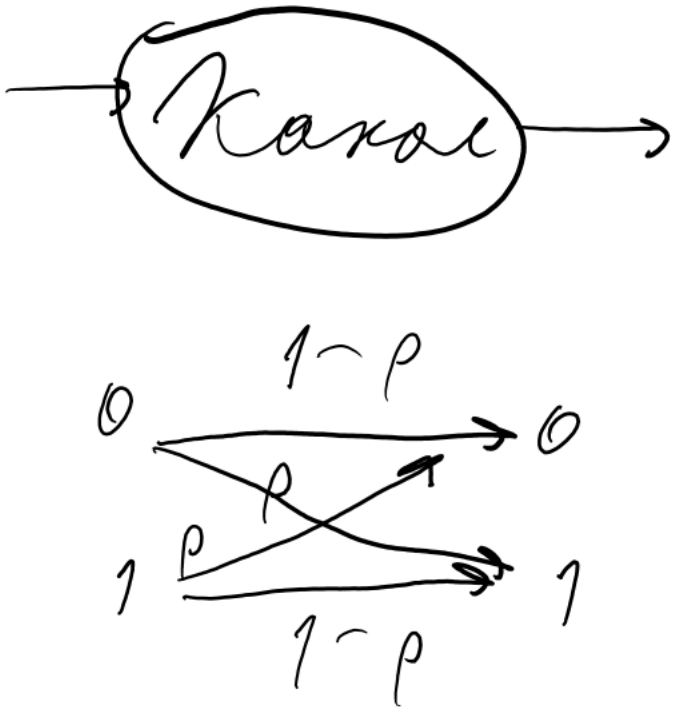
Теорема кодирования, также известная как теорема Шеннона о кодировании для канала с шумом, устанавливает фундаментальный предел возможностей надежной передачи информации по каналу связи с помехами. Она утверждает, что для любого дискретного канала с пропускной способностью C и источника с энтропией H при скорости передачи $R < C$ существуют такие коды, которые позволяют достичь сколь угодно малой вероятности ошибки декодирования. Обратная сторона теоремы говорит, что при $R > C$ надежная передача невозможна. Эта теорема заложила основу теории информации и стимулировала развитие методов помехоустойчивого кодирования.

Рассмотрим классическую схему цифровой связи «источник — кодер — канал — декодер — приемник».



Источник порождает сообщение, которое может быть непрерывным (аналоговым) или дискретным. Это может быть речь, изображение, текст и т.д. Кодер выполняет два основных преобразования: сначала исходное сообщение сжимается (кодирование источника) для устранения избыточности, а затем к нему добавляется специально введенная избыточность для защиты от ошибок (канальное кодирование). Результатом является последовательность символов, готовая к передаче. Канал — это физическая среда, по которой передается сигнал (например, радиоволны, оптическое волокно, медный провод). В канале сигнал неизбежно искажается под воздействием шумов, помех и затухания. Декодер выполняет обратные операции: по принятому искаженному сигналу он восстанавливает переданное кодовое слово (или оценивает его) и затем восстанавливает исходное сообщение, удаляя канальную избыточность. Приемник — это конечное устройство или система, которая интерпретирует восстановленное сообщение и использует его по назначению (например, выводит звук на динамик, отображает изображение на экране).

Одной из простейших моделей канала с ошибками является двоичный симметричный канал (ДСК).



В этой модели входные и выходные символы принадлежат двоичному алфавиту $\{0, 1\}$. Канал характеризуется переходной вероятностью p, которая определяет вероятность того, что переданный символ 0 будет принят как 1 или переданный 1 будет принят как 0. С вероятностью $1 - p$ символ передается без искажений. ДСК часто используется как базовая модель для анализа эффективности кодов, так как он симметричен и стационарен.

Коды, используемые для защиты от ошибок, можно классифицировать по способу обработки информации. Блоковые коды разбивают непрерывный поток данных на блоки фиксированной длины k (информационных символов) и преобразуют каждый блок в кодовое слово длины n ($n > k$), добавляя избыточные символы. Кодирование каждого блока происходит независимо. В отличие от них, сверточные коды не делят поток на независимые блоки; они обрабатывают информацию непрерывно, используя скользящее окно. Выходные символы зависят не только от текущего входного блока, но и от нескольких предыдущих, что обеспечивается памятью кодера. Это позволяет добиться высокой эффективности, но усложняет декодирование.

Для анализа и сравнения корректирующих свойств кодов используются метрики, связанные с именем Ричарда Хэмминга. Вес Хэмминга $w(\mathbf{v})$ двоичного вектора (кодового слова) $\mathbf{v}$ определяется как количество ненулевых позиций в этом векторе. Проще говоря, это число единиц в двоичной последовательности. Расстояние Хэмминга $d(\mathbf{u}, \mathbf{v})$ между двумя векторами $\mathbf{u}$ и $\mathbf{v}$ равно числу позиций, в которых эти векторы различаются. Для двоичных векторов расстояние Хэмминга можно также выразить как вес их поразрядной суммы по модулю два: $d(\mathbf{u}, \mathbf{v}) = w(\mathbf{u} \oplus \mathbf{v})$.

Ключевой характеристикой блокового кода является его минимальное расстояние $d_{\min}$. Оно определяется как наименьшее расстояние Хэмминга между любыми двумя различными кодовыми словами данного кода. Минимальное расстояние определяет корректирующую способность кода. Справедлива следующая фундаментальная теорема: код с минимальным расстоянием d гарантированно исправляет любые комбинации ошибок кратности t (где t — максимальное число ошибок в кодовом слове), если выполняется условие $t \leq \left\lfloor \frac{d-1}{2} \right\rfloor$. Здесь $\lfloor \cdot \rfloor$ обозначает целую часть числа. Это означает, что кодовые слова расположены в пространстве так, что шары радиуса t вокруг них не пересекаются, и любое искаженное слово попадает в шар только одного переданного слова, что позволяет однозначно его восстановить.

Среди блоковых кодов особое место занимают линейные коды. В линейном двоичном коде длины n множество всех кодовых слов является линейным подпространством размерности k векторного пространства $\mathbb{F}_2^n$ над полем Галуа из двух элементов. Это свойство позволяет использовать аппарат линейной алгебры для описания кода. Любой линейный код может быть задан с помощью порождающей матрицы $\mathbf{G}$ размера $k \times n$. Строки этой матрицы образуют базис подпространства кодовых слов, то есть любое кодовое слово $\mathbf{c}$ может быть получено как линейная комбинация строк порождающей матрицы: $\mathbf{c} = \mathbf{u}\mathbf{G}$, где $\mathbf{u}$ — информационный вектор длины k. Таким образом, порождающая матрица дает компактный способ кодирования.

Другим способом описания линейного кода является использование проверочной матрицы $\mathbf{H}$ размера $(n-k) \times n$. Эта матрица задает подпространство, ортогональное к коду: для любого кодового слова $\mathbf{c}$ выполняется соотношение $\mathbf{c}\mathbf{H}^\mathsf{T} = \mathbf{0}$ (в некоторых учебниках используется транспонирование иначе: $\mathbf{H}\mathbf{c}^\mathsf{T} = \mathbf{0}$). Столбцы проверочной матрицы часто используются для обнаружения и исправления ошибок — синдром принятого слова вычисляется как $\mathbf{s} = \mathbf{r}\mathbf{H}^\mathsf{T}$, и по нему можно определить наличие и местоположение ошибки. Существует также дуальное представление: если ввести вектор $\mathbf{h} = (h_1, h_2, \ldots, h_n)$, являющийся одной из строк проверочной матрицы, то для любого кодового слова $\mathbf{c} = (c_1, \ldots, c_n)$ выполняется тождество скалярного произведения $\sum_{i=1}^{n} c_i h_i = 0$ (в поле $\mathbb{F}_2$), что можно записать как $(\mathbf{c}, \mathbf{h}) = 0$. Это условие означает, что все кодовые слова ортогональны каждой строке проверочной матрицы. Связь между порождающей и проверочной матрицами выражается соотношением $\mathbf{G}\mathbf{H}^\mathsf{T} = \mathbf{0}$, которое является следствием ортогональности подпространств.