Семантические меры информации
Структурный и статистический подходы к определению количества информации не учитывают её смысл, поэтому в рамках семантического подхода рассматривают три меры информации: содержательность, целесообразность и существенность. Содержательность оценивает информацию с точки зрения верности или ложности и выражается через функцию меры m(i). По аналогии с функциями вероятности, эта мера лежит в диапазоне [0;1].
Целесообразность - это изменение вероятности, достижение цели при получении некоторого сообщения; численно равна разности логарифмов априорной и апостериорной вероятностей.
Существенность - зависит от того, насколько важно знать: 1) значимость самого события; 2) значимость точки пространства события; 3) значимость времени осуществления события.
Квантование информации
 Виды сигналов
Сигналом является любой физический процесс, параметры которого изменяются в соответствии с переданным сообщением.
Аналоговый сигнал описывается непрерывной функцией x(t), при этом и аргумент, и сама функция могут принимать любые значения из интервалов tv(min)<=t<=tv(max), xv(min)<=x<=xv(max). Аналоговый сигнал обеспечивает передачу данных путём изменения во времени амплитуды, частоты, фазы.
x(t)=Asin((omega)t+(fi))
A=xv(max) - амплитуда, (omega) - угловая частота
Непрерывная функция дискретного аргумента относится к дискретно-непрерывным сигналам. При этом значение функции определяется лишь на дискретном множестве значения аргумента tvi,i=0,1,2,...;
tv(min)<=t<=tv(max). Величина x(tvi) может принимать любые значения в диапазоне [xv(min);xv(max)].
Операцию, которая переводит сигнал x(t) непрерывного аргумента t в сигнал x(tvi) дискретного аргумента tvi, называют квантованием по времени или дискретизацией.
Дискретная функция непрерывного аргумента
Значения, которые может принимать функция, образуют дискретный ряд чисел xv1, xv2, xv3 на протяжении всего интервала времени t и для любого его момента. Квантование по уровню состоит в преобразовании непрерывного множества значений x(t) в непрерывное множество значений.
Обычно используются следующие способы соотнесения значения сигнала с соответствующим уровнем квантования: 1) сигнал относится к ближайшему уровню квантования (округление); 2) сигнал относится к ближайшему меньшему уровню квантования; 3) к ближайшему большему уровню квантования (по модулю).
Цифровой сигнал описывается дискретной функций дискретного аргумента.
x(t)->xvi(tvi).
Пускай задана функция -5t^2+0.7. Построить непрерывный вид функции, выполнить её дискретизацию по времени и по уровням (тремя способами).
L\t=1
Выбор частоты отсчётов при дискретизации
Дискретизация сигнала связана с заменой промежутка изменения независимой переменной некоторым множеством значений. Такое преобразование является однозначным, а обратное преобразование неоднозначное. Отсюда возникает понятие погрешности преобразования.
По значению дискретной функции можно восстановить исходную непрерывную функцию с некоторой погрешностью. Процесс восстановления называется интерполяцией. Полученная в результате восстановления функция называется воспроизводящей. Требуемую погрешность можно получить, управляя шагом дискретизации.
Дискретизация называется равномерной, если L\tvi=tvi-tv(i-1)=L\t=const.
Неравномерная дискретизация может проводиться с интервалами L\tvi=z(delta)t, где z(delta)t - некоторый фиксированный элементарный шаг, z=1,2,3... - некоторый коэффициент. L\tv(min)<=L\tvi<=L\tv(max)
Шаг дискретизации L\t либо частота отсчётов Fv0=1/L\t выбирается на основании априорных сведений (априорной информации) о характеристика сигнала x(t). Это означает, что существует некоторый оптимум разбивки, обусловленный теоремой Котельникова.
Для любой функции можно выполнить разложение в ряд Фурье, в результате чего функция представляется в виде суммы гармоник, которые с определённой погрешностью соответствуют исходному сигналу.
f(t)=(SIGMA)v(k=0)^mAvkcos(k(omega)t+(fi)vk)
С увеличением (omega) уменьшается вклад составляющей в суммарном сигнале, поэтому всегда можно выбрать такую (omega)v(max)=2(pi)fv(max), что для всех гармоник с (omega)>(omega)v(max) их вкладом в суммарный сигнал можно пренебречь. Набор гармоник, которые входят в состав сигнала (разложение сигнала), называют спектром сигнала. Если можно ограничить спектр сигнала значением fv(max), то для оптимального выбора частоты дискретизации можно использовать теорему Котельникова.
Функция с ограниченным спектром полностью определяется дискретным множеством своих значений, взятых с частотой fv0=2*fv(max).
Таким образом, L\t=1/fv0=1/(2fv(max)).
Трактовка этой теоремы рассматривает идеальный случай, при котором сигнал начался бесконечно давно и никогда не закончится, а также не имеет в своей временной характеристике точек разрыва.
Для случая реальных сигналов полное восстановление сигнала невозможно, и из теоремы Котельникова имеем два следствия: 1) аналоговый сигнал может быть восстановлен с требуемой точностью по своим дискретным отсчётам с частотой большей, чем 2fv(max); 2) если максимальная частота в сигнале превышает половину частоты дискретизации, то способа восстановить сигнал из дискретного в аналоговый без искажений не существует.
На практике частоту отсчётов определяют по формуле:
fv0=(kappa)vз*2fv(max), где 1,5<=(kappa)vз<=6.
Основы кодирования
Кодирование - это отображение дискретного сообщения в виде определённых комбинаций символов. Совокупность правил кодирования называется кодом.
Декодирование - это процесс, обратный кодированию.
Кодовое слово - это последовательность символов, соответствующих одному сообщению.
Первичный алфавит - это символы, с помощью которых представлено сообщение на выходе источника информации.
Вторичный алфавит - это символы, с помощью которых сообщение трансформируется в кодовую комбинацию.
Коды бывают равномерные и неравномерные. Неравномерные коды содержат сообщения с разным количеством символов (напр., двоичный код).
Равномерные коды имеют одинаковую длину для всех кодовых комбинаций (напр., шестнадцатеричный код).
Полный код - это код, где любая перестановка символов даёт кодовую комбинацию того же кода.
Теория информации
\ 0 Т 22 е 1 о 2 р 3 и 4 я 23 н 11 ф 12 м 13 а 14 ц 24   21
22010203040230210401101202030130140240040
Т 01 е 02 о 03 р 0 и 10 я 11 н 12 ф 13 м 14 а 20 ц 21   22
0102030410112210121303041420211010
Рассмотрим общую схему системы передачи информации.

```
Кодер канала  Модулятор Линия связи  Демодулятор Декодер канала
+--+             +-+  X   +--+ Y      +--+           +--+
|КК+------------>|М+----->|ЛС+------->|ДМ+---------->|ДК|
++-+             +-+      +--+        +--+           ++-+
 ^                         ^                          |
 |                         |                          v
++-+                      ++-+                       ++--+
|КИ|Кодер источника       |ИП|Источник помех         |ДПр|Декодер
приёмника
+--+                      +--+                       ++--+
 |                                                    |
 |z                                                   v
++-+                                                 +-+
|ИИ|Источник информации                              |П|Приёмник
+--+                                                 +-+
```

z - исходный сигнал
Кодер источника предназначен для устранения избыточности первичного алфавита, то есть для уменьшения среднего числа символов на одну букву сообщения. При отсутствии помех за счёт кодирования сокращается время передачи и объём переданной информации, что повышает эффективность системы. Такое кодирование называют эффективным, или оптимальным.
Кодер канала обеспечивает заданную достоверность приёма и передачи информации путём введения искусственной избыточности. Такое кодирование называется помехоустойчивым.
Модулятор предназначен для формирования физического сигнала.
Линия связи - физическая среда, по которой распространяется сигнал.
Источник помех - всегда существует в реальных условиях.
Демодулятор - выполняет преобразование, обратное тому, которое выполнил модулятор.
Декодер канала - предназначен для выявления и коррекции ошибок передачи. Он обеспечивает помехоустойчивость.
Декодер приёмника - предназначен для представления сообщения в алфавите, понятном для приёмника.
Способы кодирования и представления кода
Табличное представление (перечисление пар), задающих соответствие "символ-код". Любому дискретному сообщению либо букве сообщения можно приписать некоторый порядковый номер. Передача или хранение сообщения при этом сводятся к передаче либо хранению чисел, представленных в требуемой системе исчисления.
А - 00000
Б - 00001
В - 00010
Представление в виде многочлена
Для любой системы счисления с основой x в присутствии n различных цифр любое число может быть представлено в виде многочлена.
F(x)=av0+av1x^1+av2x^2+...+av(n-1)x^(n-1)=(SIGMA)v(i=0)^(n-1)avix^i
   210
П - 105v6
F(x)=5+x^2
Представление в виде геометрической модели
Кодовые комбинации n-значного кода рассматриваются как точки n-мерного пространства. Число градаций по каждой оси соответствует основе системы счисления. Для построения кодового пространства необходимо в строго фиксированной последовательности осуществлять проекцию каждой из n точек на оси n-мерного пространства. При этом значение кодовых комбинаций являются координатами в кодовом пространстве.

```
       ^3             321
       |              ---
       +-+            000
      /|/|            001
     +-+-+---------->2   110
     |/|/             111
     +-+
    /
   /
  /
 v1
    ^3             321
    |              ---
    |              123v4
    |              ---
   .201v4          201
   |+-+-+-------->2
   + / /
  / / /
 / / /
v1/ /
```

Кодовое расстояние - это расстояние между точками кодового пространства. Кодовое расстояние показывает, через какое количество расстояний должны пройти качественные признаки кодовой комбинации для того, чтобы перейти в требуемое состояние.
Минимальное кодовое расстояние - это расстояние между самыми ближними точками кодового пространства. Это основной параметр, который определяет помехоустойчивость дискретного кода с произвольной основой.
Представление в виде матрицы
Для полных равномерных кодов несложно заметить, что если взять соответствующую единичную матрицу, то сложение её строк по модулю 2 позволяет получить все комбинации данного кода. Такая матрица называется определяющей либо порождающей. Этот способ используется для построения корректирующих кодов.
Представление кода в виде дерева
В общем виде кодовое дерево может быть представлено как граф. Каждый уровень графа содержит n^m узлов, где n - нормы уровня, а m - значность кода. Для равномерного двоичного кода число узлов на каждом уровне равно 2^n.

```
         ( )
          /   \
      (0)   (1)
     /   \  |  \
   (00) (01)(10)(11)
  /   \
(000)(001)
```

Метод Хаффмана
Метод Хаффмана является процедурой построения префиксного кодового дерева.
1. Первичный алфавит располагается в порядке уменьшения вероятностей.
2. Последние два символа объединяются в новый символ с вероятностью, равной сумме вероятностей образовавших его символов.
3. Новый алфавит упорядочивается по убыванию вероятностей.
4. Процедура продолжается до тех пор, пока вероятность нового символа не будет равна единице. Эти шаги ведут к построению дерева.
5. Веткам построенного дерева присваиваются символы вторичного алфавита и делается это всегда в одинаковой последовательности.
6. Кодовыми комбинациями являются последовательности символов вторичного алфавита, которые расположены на пути от корня до вершин дерева.
Построить оптимальный неравномерный код по методу Хаффмана для символов первичного алфавита с вероятностями: 0.15, 0.1, 0.2, 0.05, 0.05, 0.1,

```
0.05, 0.3.
№ литеры|Вероятность   |A1       |A2      |A3      |A4    |A5    |A6   |
A7
--------+--------------+---------+--------+--------+------+------+-----
+----
   1    |    0.3  00   |0.3  00  |0.3  00 |0.3  00 |0.3 00>0.4 1 >0.6 0|
1
   2    |    0.2  10   |0.2  10  |0.2  10 |0.2  10 >0.3 01|0.3 00]0.4 1|
   3    |    0.15 010  |0.15 010 |0.15 010>0.2  11 |0.2 10]0.3 01]      |
   4    |    0.1  110  |0.1  110 >0.15 011|0.15 010]0.2 11]      |      |
   5    |    0.1  111  |0.1  111 |0.1  110]0.15 011]     |      |      |
   6    |    0.05 0111 >0.1  0110]0.1  111]        |      |      |      |
   7    |    0.05 01000]0.05 0111]        |        |      |      |      |
   8    |    0.05 01101]         |        |        |      |      |      |
```

Помехоустойчивость, эффективность и надёжность систем передачи информации

Помехоустойчивость - это способность системы осуществлять приём информации в условиях присутствия помех, т. е. сторонних воздействий.
Помехоустойчивость бывает статическая и динамическая.
Статическая помехоустойчивость - это защищённость от неправильного срабатывания при отсутствии информации.
Динамическая помехоустойчивость - это способность выделять полезный сигнал.
Для повышения статической помехоустойчивости разрабатывают коды, имеющие большое количество сигналов на сообщение, так как чем сложнее сообщение, тем меньше вероятности сконструировать его из случайных помех.
Одним из путей повышения динамической помехоустойчивости является формирование специальной стартовой комбинации, открывающей приёмник. Эта стартовая комбинация является намного более сложной, чем информационные коды.
Другой путь - это кодирование информации блоками, что сокращает передаваемое сообщение, т. к. чем короче сообщение, тем меньше вероятность его искажения.
Эффективность вводится для оценки степени целесообразности усложнения кодов при получении заданной помехоустойчивости. При одинаковой полосе частот и мощности на сообщение более эффективной считается та система, которая передаёт сообщения за более короткий промежуток времени.
Надёжность - это способность безотказной работы на протяжении определённого промежутка времени.

Методы повышения надёжности и помехоустойчивости

1. Приёмник Котельникова.
 Суть метода заключается в увеличении времени передачи полезного сигнала.
2. Метод Бодо-Вердена (мажоритарное декодирование)
 Суть метода заключается в том, что сообщения передаются многократно и приёмник накапливает сигнал каждого вида. При нечётном числе повторений правильным считается сигнал, который повторяется чаще.
3. Метод решающей обратной связи (квитирование)
 Суть метода заключается в том, что приёмник возвращает принятое сообщение источнику. Источник сравнивает исходное сообщение и полученное и делает выводы о качестве передачи информации.
4. Проверка на чётность.
 Суть метода заключается в том, что множество кодовых комбинаций разбивается на подмножество допустимых и подмножество недопустимых кодовых комбинаций. В подмножество допустимых комбинаций относят те, которые имеют чётное количество единиц. При этом любая одиночная ошибка переводит разрешённую кодовую комбинацию в неразрешённую.

Кодовое расстояние (d) - это параметр, характеризующий помехоустойчивость кода и заложенную в нём избыточность. Минимальное кодовое расстояние между разрешёнными и неразрешёнными кодами dv0=1, следовательно, минимальное кодовое расстояние между разрешёнными комбинациями dv0=2. Достигается это введением дополнительного контрольного разряда.
Таким образом, при помехоустойчивом кодировании общая длина кодовой комбинации n складывается из количества информационных и контрольных разрядов.
Для кодового расстояния dv0=3 в задачах исправления одиночных ошибок имеют место следующие формулы:
nvк=[logv2(n+1)]
nvк=[logv2((nvu+1)+logv2(nvu+1)])]

Основная теорема Шеннона про кодирование в присутствии шумов

1) Если источник с энтропией H(A) создаёт информацию на входе шумового канала без памяти со скоростью меньшей, чем пропускная способность канала, то существует такой код, при котором вероятность ошибки на приёмнике сколь угодно мала.
2) Если H(A) ограничено некоторым пределом, то при кодировании достаточно длинными блоками среди кодов, обеспечивающих сколь угодно малую вероятность ошибки, существует код, при котором скорость передачи информации максимально близка к скорости образования информации.

Линейные групповые коды

Линейными называются такие коды, у которых контрольные символы являются линейными комбинациями информационных символов. Для двоичных кодов в качестве линейной операции использую сумму по модулю 2.
Линейные коды всегда можно представить в систематической форме, при которой информационные и корректирующие символы расположены в строго определённых позициях кодовой комбинации.
Свойства линейных кодов:
1) сумма по модулю 2 кодовых комбинаций некоторого кода является кодовой комбинацией того же кода, т. е. линейные коды образуют группу относительно операции "сумма по модулю 2", поэтому коды называют групповыми.
Наиболее распространённым среди таких кодов является код Хемминга, предназначенный для исправления одиночной ошибки. Для формирования кода Хемминга необходимо задать:
1) количество информационных разрядов nvi;
2) выявляющие свойства кода r;
3) корректирующие свойства кода S.
Необходимо вычислить n и nvk и приступить непосредственно к формированию кода.
1) Номера контрольных символов выбираем по закону 2^i, где i=1,2,3,..., то есть контрольными разрядами являются 1-й, 2-й, 8-й, 16-й, 32-й и т. д.
2) Каждый контрольный разряд следит за определённым набором проверяемых символов. Для определения этого набора составим таблицу, число строк в которой равняется n=nvi+nvk. Первой строке соответствует проверочный символ 1, второй - 2 и т. д.

```
0 0 0 0 | 1
0 0 1 0 | 2
0 0 1 1 | 3
0 1 0 0 | 4
0 1 0 1 | 5
0 1 1 0 | 6
0 1 1 1 | 7
```

Принцип формирования наборов проверочных символов: в первую проверку входят символы, содержащие единицу в младшем разряде, во вторую проверку - символы, содержащие единицу во втором разряде и т. д.

```
Проверяемые разряды    Контр. ряд
1-я проверка -> 1,3,5,7,9,11,... -> 1
2-я проверка -> 2,3,6,7,10,11,14,15,... -> 2
3-я проверка -> 4,5,6,7,12,13,14,15,... -> 4
```

3) Значение каждого контрольного символа выбирается так, чтобы его сумма по модулю 2 со всеми проверяемыми символами равнялась нулю.
4) Выявление ошибки: факт ошибки устанавливается при получении единицы при суммировании по модулю 2 разрядов в соответствующей проверке.
5) Коррекция ошибок: последовательное проведение проверок даёт вектор из нулей и единиц. Он называется синдромом. Если синдром равен нулю, то ошибки нет, иначе значение синдрома даёт номер ошибочного разряда.
Пример: исправить любую одиночную ошибку при передаче кодовой комбинации 0101.

```
 1  2  3  4  5  6  7
kv1kv2 0 kv3 1  0  1
1-я пров.: 1(+)3(+)5(+)7
kv1(+)0(+)1(+)1=0
kv1=0
2-я пров.: 2(+)3(+)6(+)7
kv2(+)0(+)0(+)1=0
kv2=1
3-я пров.: 4(+)5(+)6(+)7
kv3(+)1(+)0(+)1=0
kv3=0
0 1 0 0 1 0 1
```

Внесём ошибку и исправим её.

```
0 1 1 0 1 0 1
1-я пров.: 0(+)1(+)1(+)1=1
2-я пров.: 1(+)1(+)0(+)1=1
3-я пров.: 0(+)1(+)0(+)1=0
011v2=3-й разряд
```