Теория информации

Лекция 7

Передача данных через шумящий канал.

Зависимые случайные переменные.

До сих пор мы рассматривали независимые случайные переменные когда рассматривали строку символов из некоторого алфавита A_x . Теперь рассмотрим совместные ансамбли, поскольку данные из реального мира, как правило, коррелируют между собой. Кроме того, случайная величина x на входе шумящего канала связи коррелирует со случайной величиной y на выходе канала. Иначе связь была бы вообще невозможна.

Рассмотрим некоторые определения касающиеся совместных ансамблей. Пусть X , Y и Z - ансамбли с алфавитами $A_x=\{a_1, a_2, \ldots, a_I\}$, $A_y=\{b_1, b_2, \ldots, b_J\}$ и $A_z=\{c_1, c_2, \ldots, c_K\}$ соответственно. Тогда *совместная энтропия* это

$$H(X,Y) = \sum_{xy \in A_x A_y} P(x, y) \log_2 \frac{1}{P(x, y)} \quad . \qquad (7.1)$$

Энтропия величина аддитивная для независимых величин, т.е.

$$H(X,Y) = H(X) + H(Y) \;\Leftrightarrow\; P(x, y) = P(x)P(y) \quad . \qquad (7.2)$$

Условная энтропия X *при заданном* $y=b_k$ это это энтропия условного распределения $P(x \mid y=b_k)$, т.е.

$$H(X \mid y=b_k) = \sum_{x \in A_x} P(x \mid y=b_k) \log_2 \frac{1}{P(x \mid y=b_k)} \quad . \qquad (7.3)$$

Условная энтропия X *при заданном* Y это средняя по y величина (7.3):

$$H(X \mid Y) = \sum_{y \in A_y} P(y) \left[\sum_{x \in A_x} P(x \mid y) \log_2 \frac{1}{P(x \mid y)} \right] = \sum_{xy \in A_x A_y} P(x, y) \log_2 \frac{1}{P(x \mid y)} \quad . \qquad (7.4)$$

Это мера средней неопределенности x , которая остается когда случайная величина y уже известна.

Чтобы подчеркнуть контраст с условной энтропией, энтропию $H(X)$ иногда называют *безусловной* (или *маргинальной*).
Цепное правило для количества информации следует из выражения для вероятности $P(x,y)=P(x)P(y|x)$:

$$\log_2 \frac{1}{P(x,y)} = \log_2 \frac{1}{P(x)} + \log_2 \frac{1}{P(y|x)} \tag{7.5}$$

или

$$h(x,y) = h(x) + h(y|x) \quad . \tag{7.6}$$

Аналогичное *цепное правило для энтропии*:

$$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y) \quad . \tag{7.7}$$

Взаимная информация между X *и* Y определяется как

$$I(X;Y) = H(X) - H(X|Y) \tag{7.8}$$

и удовлетворяет равенству $I(X;Y)=I(Y;X)$ и неравенству $I(X;Y)\geq 0$. Это мера того насколько в среднем уменьшается неопределенность относительно x когда становится известна y и наоборот.

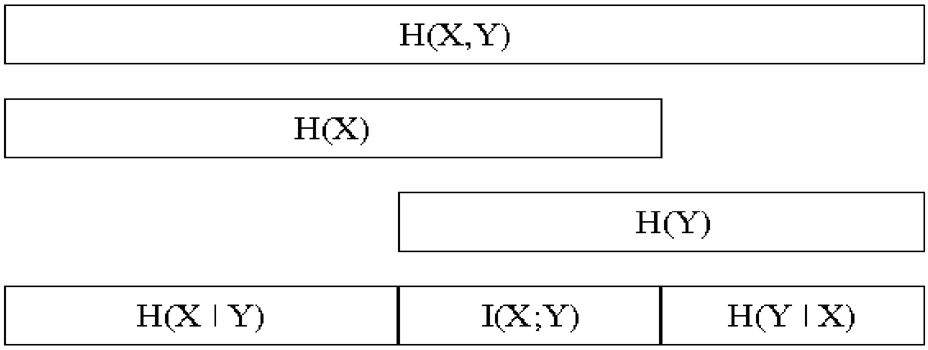
Для трех ансамблей можно ввести *условную взаимную информацию между X и Y при заданном* $z=c_k$

$$I(X;Y|z=c_k) = H(X|z=c_k) - H(X|Y,z=c_k) \tag{7.9}$$

и условную взаимную информацию между X и Y при заданном Z

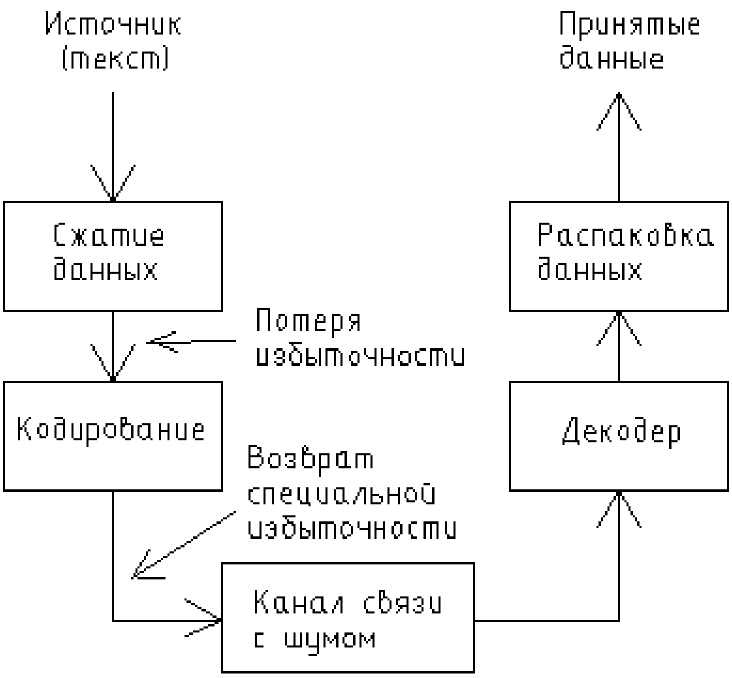
$$I(X;Y|Z) = H(X|Z) - H(X|Y,Z) \quad . \tag{7.10}$$

Здесь термин *взаимная* относится только к двум ансамблям и взаимность обозначается точкой с запятой. Поэтому обозначения типа $I(X;Y;Z)$ или $I(X|X;Y)$ неверные и ничего не значат, однако вполне можно записать $I(A,B;C,D|E,F)$. Ниже приведена диаграмма, показывающая соотношение между определенными выше соотношениями.



Передача данных через шумящий канал

До сих пор, когда мы рассматривали сжатие и распаковку данных, то неявно предполагали, что канал между упаковщиком и распаковщиком идеален. Реальные же каналы обязательно добавляют некоторый шум. Поэтому, чтобы сделать поведение канала подобным поведению канала без шума, необходимо добавить кодер и декодер. Ниже представлена схема такой передачи данных.



Предположим мы передаем со скоростью 1000 бит/с двоичные данные с вероятностями нулей и единиц $p_0 = p_1 = 1/2$ через канал с вероятностью ошибки $f = 0.1$. Если задаться вопросом, какова же реальная скорость передачи, то на первый взгляд может показаться, что $1000(1-0.1)=900$ бит/с. Однако это предположение является неверным, потому что принимающая сторона не знает когда именно произошла ошибка. Рассмотрим случай, когда шум настолько велик, что символы на принимающей стороне канала не зависят от символов на стороне передачи. Это соответствует уровню шума с $f = 0.5$. В этом случае информация через канал не передается вообще. Резонно предположить, что мера переданной информации определяется взаимной информацией между источником и принятыми на другом конце канала данными.

Пример

Совместное распределение и маргинальные вероятности приведены в таблице:

P(x,y)		*x*				*P(y)*
		1	2	3	4	
y	1	1/8	1/16	1/32	1/32	1/4
	2	1/16	1/8	1/32	1/32	1/4
	3	1/16	1/16	1/16	1/16	1/4
	4	1/4	0	0	0	1/4
P(x)		1/2	1/4	1/8	1/8	

Совместная энтропия этого ансамбля $H(X,Y)=27/8=3.375$ бит. Маргинальные энтропии $H(X)=7/4$ и $H(Y)=2$ бит. Можно также вычислить условные распределения величины *x* для каждого значения *y* и энтропию каждого условного распределения:

P(x ǀ *y)*		*x*				*H(X* ǀ *y)*
		1	2	3	4	
y	1	1/2	1/4	1/8	1/8	7/4
	2	1/4	1/2	1/8	1/8	7/4
	3	1/4	1/4	1/4	1/4	2
	4	1	0	0	0	0

H(X ǀ *Y) = 11/8*

Заметим, что величина $H(X|y)$ иногда больше, а иногда меньше чем $H(X)$, что означает, что знание y может как увеличить, так и уменьшить неопределенность относительно x. Однако в среднем $H(X|Y) < H(X)$. Взаимная информация в данном случае $I(X;Y) = H(X) - H(X|Y) = 3/8$ бит.

Модели шумящих каналов

Дискретный канал без памяти Q характеризуется входным алфавитом A_x, выходным алфавитом A_y и набором условных распределений вероятностей $P(y|x)$, по одному распределению на каждый $x \in A_x$. Эти переходные вероятности могут быть записаны в виде матрицы

$$Q_{ji} = P(y=b_j \mid x=a_i) \quad . \tag{7.11}$$

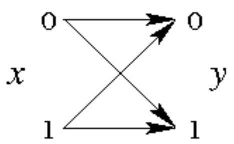
Эта матрица ориентирована так, что можно записать

$$\boldsymbol{p}_y = \boldsymbol{Q}\,\boldsymbol{p}_x \quad . \tag{7.12}$$

Ниже несколько полезных моделей каналов.

Двоичный симметричный канал.

$$A_x = \{0,1\}, \qquad A_y = \{0,1\}$$



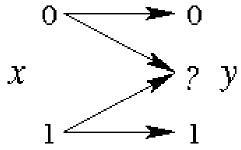
$$P(y=0|x=0) = 1-f \qquad\qquad P(y=0|x=1) = f$$

$$P(y=1|x=0) = f \qquad\qquad P(y=1|x=1) = 1-f$$

$$\boldsymbol{Q} = \begin{pmatrix} 1-f & f \\ f & 1-f \end{pmatrix}$$

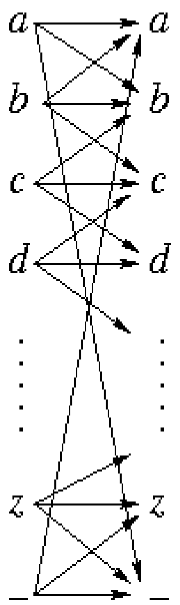
Двоичный стирающий канал

$$A_x = \{0,1\}, \qquad A_y = \{0,?,1\}$$



$$\boldsymbol{Q} = \begin{pmatrix} 1-f & 0 \\ f & f \\ 0 & 1-f \end{pmatrix}$$

Неточная печатная машинка



....

$$P(y=b \mid x=c) = 1/3$$

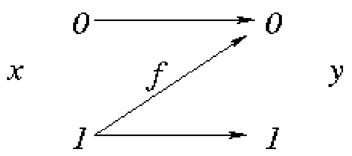
$$P(y=c \mid x=c) = 1/3$$

$$P(y=d \mid x=c) = 1/3$$

....

$$\boldsymbol{Q} = \begin{pmatrix} 1/3 & 1/3 & 0 & ... & 0 & 1/3 \\ 1/3 & 1/3 & 1/3 & 0 & ... & ... \\ 0 & 1/3 & 1/3 & 1/3 & 0 & ... \\ ... & ... & ... & ... & ... & ... \\ 1/3 & 0 & ... & ... & 1/3 & 1/3 \end{pmatrix}$$

Z-канал



$$\boldsymbol{Q} = \begin{pmatrix} 1 & f \\ 0 & 1-f \end{pmatrix}$$