

КОДИРОВАНИЕ ИНФОРМАЦИИ ИЗ ПРОГРАММЫ В ПК

 Модуль 1. Занятие 4.



Тема **и цель**

Тема:

Процесс кодирования информации

Цель занятия:

Изучение процесса кодирования информации



Глоссарий

- 1. Персональный компьютер** — однопользовательская ЭВМ (электронно-вычислительная машина), имеющая эксплуатационные характеристики бытового прибора и универсальные функциональные возможности.
- 2. Прием / передача информации** — заблаговременно организованное техническое мероприятие, результатом которого становится воспроизведение информации, имеющейся в одном месте (источнике) или в другом месте (приемнике).
- 3. Кодирование информации** — процесс преобразования сигнала из формы, удобной для непосредственного использования информации, в форму, удобную для передачи, хранения или автоматической переработки.



Вопросы

1. Что такое **операционная система**?
2. Для чего используются **накопители**?
3. Как можно классифицировать память?
4. Какую роль выполняет материнская плата в персональном компьютере?
5. Что такое **персональный компьютер**?



Введение в кодирование

Информация, которую пользователи получают из окружающего их мира, поступает в виде условных знаков или сигналов в самой разной физической природы.

Кодирование — перевод информации в удобную для передачи, обработки или хранения форму с помощью специального кода. Для кодирования разных знаков при прямом преобразовании цепи сообщений в строй цепи возможно, кроме натуральных чисел, использовать любой (упорядоченный) алфавит символов достаточно большой мощности. Соответствие между исходной и закодированной таким образом последовательностями называется «совпадение с точностью до переименования».

Прямое преобразование знаковых цепей



Цифровые последовательности и диаграммы в строю цепи:

Q	R	D	&	V	Y	S	S	&	S	D	D	S	Q	S	D	&
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

×	μ	*	Δ	ζ	Ω	λ	λ	Δ	λ	*	*	λ	×	λ	*	Δ
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

W	T	C	G	H	U	R	R	G	R	C	C	R	W	R	C	G
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

В	Г	Р	К	Т	Н	С	С	К	С	Р	Р	С	В	С	Р	К
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

!	+	#	\$	H	U	*	*	S	*	#	#	*	!	*	#	\$
---	---	---	----	---	---	---	---	---	---	---	---	---	---	---	---	----

5	6	3	4	2	0	1	1	4	1	3	3	1	5	1	3	4
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

W	!	H	-	^	*	O	O	-	O	H	H	O	W	O	H	-
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

0	4	3	5	6	2	1	1	5	1	3	3	1	0	1	3	5
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

@	×	*	E	J	ζ	H	H	E	H	*	*	H	@	H	*	E
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

1	2	3	4	5	6	7	7	4	7	3	3	7	1	7	3	4
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---



Вектор **стро́я**

В соответствии с таким определением вектора, можно назвать специфически сформированным (организованным) кортежем или **вектором стро́я**. Здесь стоит отметить, что при несоблюдении ограничений на порядок расположения натуральных чисел можно получить кортеж, точнее вектор, который не представляет собой строк.

Покажем пример построения такого вектора:

1	2	3	5	4	6	7	7	5	7	3	3	7	1	7	3	5
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---



Обратное преобразование строя

В тоже время, преобразование строя в знаковую цепь называется **обратным преобразованием строя:**

1	2	3	4	5	1	5	6	7	7	1	5	2	7	5	5	7
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

&	Q	R	V	S	&	S	Y	D	D	&	S	Q	D	S	S	D
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

\$!	U	H	*	\$	*	+	#	#	\$	*	!	#	*	*	#
----	---	---	---	---	----	---	---	---	---	----	---	---	---	---	---	---

Здесь стоит отметить, что **при одинаковой мощности составов знаковых цепей их частотные распределения одинаковы, т.е. инвариантны относительно расположения элементов в цепях.**



Чтение строки по направлению

«справа налево»

Для примера, прочтем прошлый строй по направлению «**справа налево**» и получим кортеж или вектор, который по определению не является строим цепи (при чтении «слева направо»):

4	3	7	1	7	3	3	7	4	7	7	6	5	4	3	2	1
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

1	2	3	4	3	2	2	3	1	3	3	5	6	1	2	7	4
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---



Однородные последовательности и интервально-матричная модель строя цепи

Операция выявления в разных по природе информационных цепях одинаковых построений может расширить возможности междисциплинарных исследований. В тоже время, результат такой операции ограничен описанием строя в форме обычного числового кортежа, хотя и имеющего определения **«вектор строя»**.

Для определения формализма строя отдельной цепи при обычном способе ее чтения «поэлементно» можно ввести две нумерации:

- 1. Первая нумерует** элементы собственного алфавита (словаря) данной знаковой последовательности по мере их первой встречи
- 2. Вторая дает сквозную нумерацию** всех компонентов кортеже от начала до конца.



Декомпозиция строя

Теперь разложим однородную символьную последовательность на X неполных однородных кортежей, на позициях которых заняты одинаковыми знаками только некоторые места:

<i>V</i>	<i>N</i>	<i>A</i>	<i>B</i>	<i>J</i>	<i>K</i>	<i>T</i>	<i>T</i>	<i>B</i>	<i>T</i>	<i>A</i>	<i>A</i>	<i>T</i>	<i>V</i>	<i>T</i>	<i>A</i>	<i>B</i>	знаковая цепь	матрица интервалов				
1	2	3	4	5	6	7	7	4	7	3	3	7	1	7	3	4	строй цепи					
1													1				однородные цепи и соответствующие им цепи интервалов	0	0	0	13	4
	2																	0	0	0	0	16
		3								3	3				3			0	8	1	4	2
			4					4								4		0	0	5	8	1
				5														0	0	0	0	13
					6													0	0	0	0	12
						7	7		7			7	7					1	2	3	2	3

\updownarrow X
 $\leftarrow n_{jmax} \rightarrow$



Декомпозиция строа

Декомпозиция строа полной неоднородной знаковой цепи на неполные однородные представлена выше, **разнородные (неполные не однородные) цепи** - ниже

<i>V</i>	<i>N</i>	<i>A</i>	<i>B</i>	<i>J</i>	<i>K</i>	<i>T</i>	<i>T</i>	<i>B</i>	<i>T</i>	<i>A</i>	<i>A</i>	<i>T</i>	<i>V</i>	<i>T</i>	<i>A</i>	<i>B</i>	знаковая цепь	матрица интервалов						
1	2	3	4	5	6	7	7	4	7	3	3	7	1	7	3	4	строи цепи							
1	2	3	4	5	6	7												1	1	1	1	1	1	11
							7	4		3			1					0	0	0	1	2	3	4
									7		3					4		0	0	0	0	2	5	1
												7			3			0	0	0	0	0	3	2
													7					0	0	0	0	0	0	3
														7										

n_{jmax}
 m



Декомпозиция **СТРОЯ**

В случае правильного выполнения декомпозиций, полученные множества однородных (разнородных) последовательностей будут **несовместимыми**:

T	T	G	G	G	T	T	C	C	G	G	G	G	G	G	<i>Cricetulus griseus</i>
G	G	A	A	A	G	G	T	T	A	A	A	A	A	A	<i>Homo sapiens</i>
1	1	2	2	2	1	1	3	3	2	2	2	2	2	2	строй, общий для обоих фрагментов

G	-	-	-	-	G	G	-	-	-	-	-	-	-	-	G	G	-	-	G
A	-	-	-	-	A	A	-	-	-	-	-	-	-	-	A	A	-	-	A
1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1	0	0	1

Строй знаковой последовательности



Здесь следует отметить, что описание знаковых цепей и текстов графами также возможно. Обычно, это взвешенный граф типа «дерево», узлы которого представляют выделенные по определенным признакам символы или слова, а ребра - интервалы между ними.

Для исследования построения реальной информационной цепи вводится формальное понятие - **строй знаковой последовательности**, который представляет только определенный порядок следования, взаимное расположение различных и одинаковых его компонентов без учета их обозначений и содержимого.

Функции характеристик строя



Для формального определения функции характеристики строя используются некоторые понятия:

1. **Место** - элементарная ячейка, предназначенная для хранения одного компонента цепи
2. **Позиция** - это упорядоченное множество мест
3. **Фрагмент** - участок полной цепи
4. **Окно** - позиция фрагмента (участок позиции полной цепи)
5. **Размер окна** - количество мест для позиции окна
6. **Шаг** - смещение окна на позиции полной цепи, позволяющее выделить следующий фрагмент цепи
7. **Размер шага** - размер смещения окна, измеряемый числом мест
8. **Функция характеристики строя цепи** - упорядоченное множество значений характеристик строя, вычисленных для всех фрагментов, последовательно взятых на позиции полной цепи
9. **Отсчетное значение функции характеристики строя** - значение функции характеристики строя, вычисленное для отдельного фрагмента, задаваемого его номером, длиной, и размером шага.

Характеристики зависимости однородных цепей



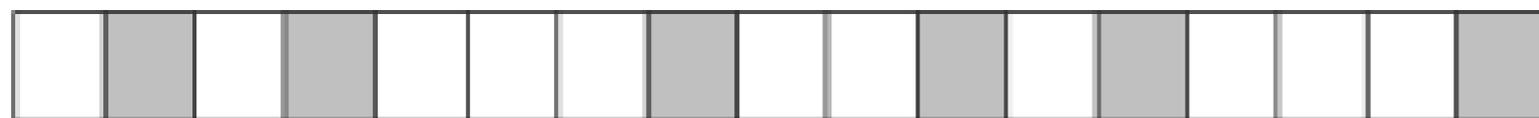
На основе наглядных графических образов (диаграмм) могут быть представлены **характеристики пространственной зависимости** пары однородных цепей, взятых из состава полной неоднородной цепи. Это позволяет соотносить субъективное суждение о степени «пространственной зависимости» элементов сопоставляемых цепей с формальной числовой характеристикой, представляющей такую зависимость:



полная неоднородная цепь



j -ая однородная цепь



l -ая однородная цепь



бинарно-однородная цепь

Элементы j-й цепи

На диаграммах ниже элементы j-й цепи обозначены черными квадратами, а элементы l-й - серыми:

Диаграмма 1



Диаграмма 2



Диаграмма 3



Диаграмма 4



Диаграмма 5



Завис-ть	K_1	K_2	K_3
От	-11	-11	0
От	0	0	

от	0,733	0,44	0,282
от	0,129	0,181	

от	0,939	0,939	0,712
от	0,539	0,539	

от	0,5	0,5	0,353
от	0,25	0,25	

от	0,996	0,996	0
от	-0,6	-0,6	

Неустановленная или неопределенная характеристика соответствия

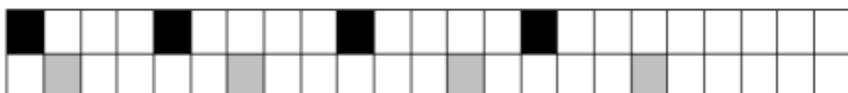


Диаграмма 7



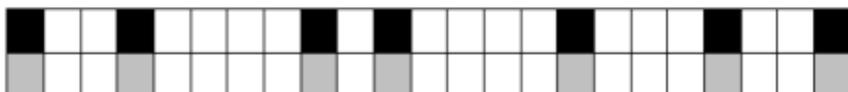
Соответствие		a_1	a_2
от	от	0,889	0,842
от	от	0,797	

Диаграмма 8



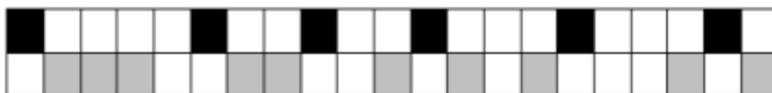
от	от	0,992	0,856
от	от	0,738	

Диаграмма 9



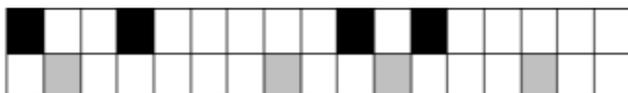
от	от	1	1
от	от	1	

Диаграмма 10



от	от	0,674	0,645
от	от	0,617	

Диаграмма 11



от	от	0,942	0,835
от	от	0,739	

Диаграмма 12



от	от	0,911	0,812
от	от	0,723	

Диаграмма 13



от	от	0,872	0,807
от	от	0,747	

Неустановленная или неопределенная характеристика соответствия отмечается для такого случая, когда размеры сопоставляемых интервалов не задаются в форме определенного преобразования:



Факторы и средства, обеспечивающие сжатие данных

Абсолютная избыточность эффективного кода определяется разностью средней длины кодового слова и энтропии.

Если при кодировании отдельных сообщений исходного текста избыточность кода оказывается велика, то следует кодировать в исходной последовательности не отдельные сообщения, а целые их группы – блоки.

Основной эффект сжатия массива данных, записанного кодовыми словами, получается за счет того, что более частые сообщения кодируются короткими словами, а редкие – длинными, так что средняя длина кодового слова получается минимальной.

Статистически зависимые сообщения



Однако очевидны ситуации, когда на вход приемника поступают статистически зависимые сообщения. Реальные тексты, файлы – это последовательности вовсе не случайных сообщений. Однако оптимальный код предназначен для кодирования только статистически независимых сообщений. В отмеченных ситуациях требуется предварительно осуществлять де-корреляцию последовательности сообщений.

Декорреляция и кодирование L - грамм по четыре сообщения имеют следующий вид:

«Пусть имеется информация...»

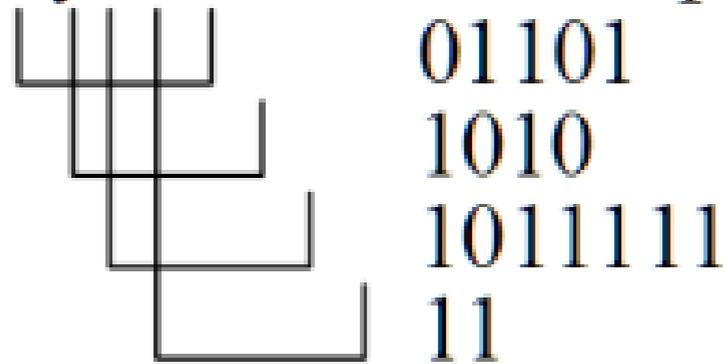
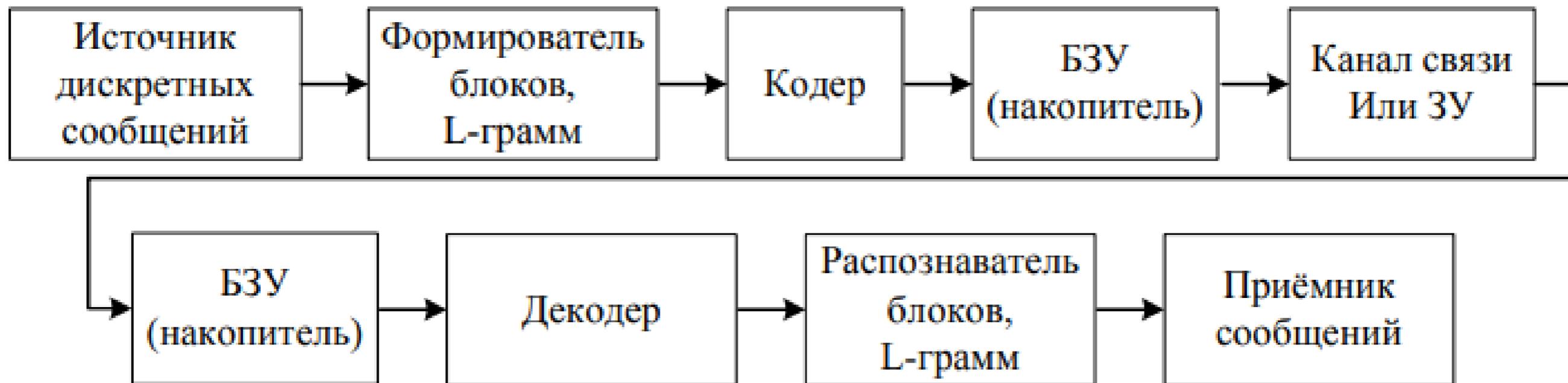


Схема системы эффективного

КОДИРОВАНИЯ



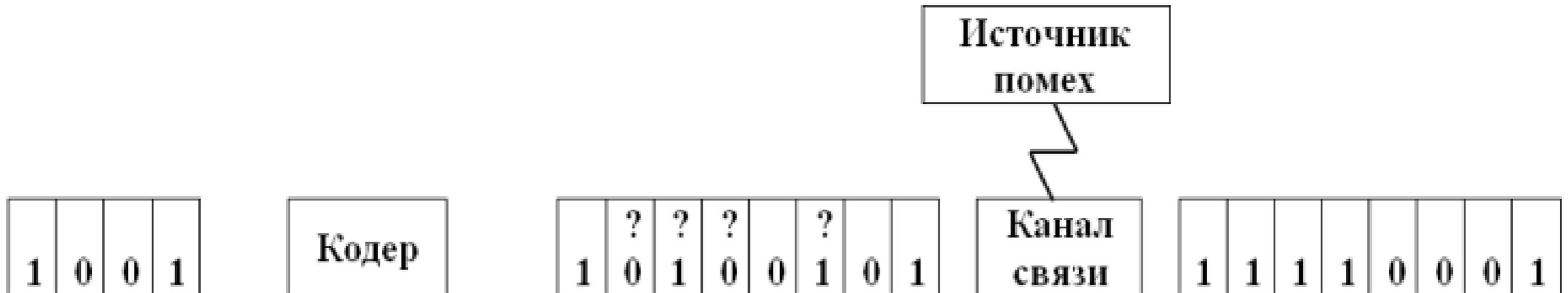
Полная структурная схема системы эффективного кодирования, передачи и декодирования сообщений представлена ниже:



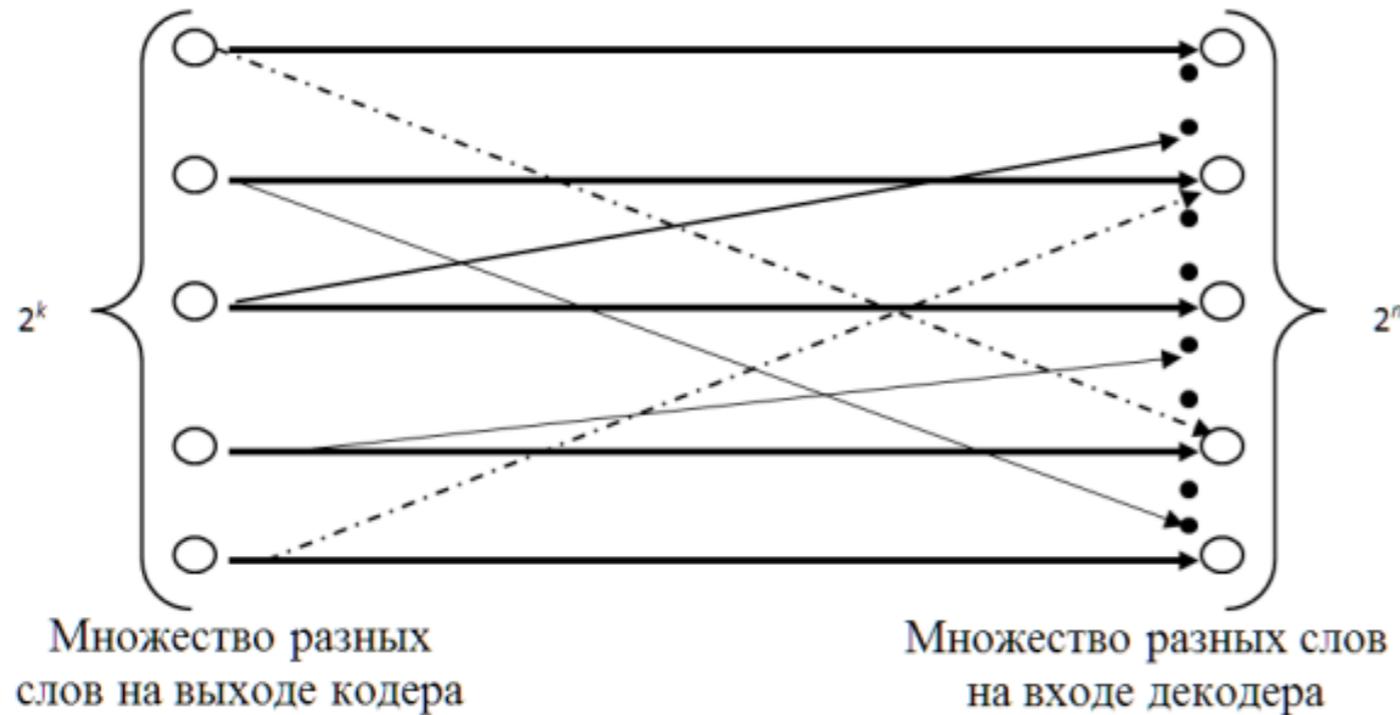
Помехоустойчивое кодирование



Идея помехоустойчивого кодирования состоит во внесении кодером избыточной информации в виде алгоритма (правил) кодирования с помощью дополнительных разрядов помехоустойчивого кодового слова, с последующей проверкой декодером этого слова на соответствие принятому **алгоритму кодирования**:



Передача кодовых слов от кодера к декодеру



Здесь представлены варианты передачи кодовых слов от кодера к декодеру:

- - разрешенные слова;
- - запрещенные слова;
- - варианты правильной передачи;
- - - - - варианты передачи с необнаруживаемой ошибкой;
- · · · · варианты передачи с обнаруживаемой ошибкой.



Вопросы

1. Что такое **кодирование**?
2. Какие компоненты есть у ПК?
3. Что такое **ввод/вывод данных**?
4. Какими бывают программы?
5. Что такое **программа**?