

ФМ



С. В И Н О Г Р А Д
Д Ж . Д . К О У Э Н

Надежные
вычисления
при наличии
шумов



БИ



RELIABLE COMPUTATION IN THE PRESENCE OF NOISE

S. WINOGRAD and J. D. COWAN

THE M.I.T. PRESS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS
1968

Физико-
Математическая
Библиотека
Инженера

С. ВИНОГРАД
Дж. Д. КОУЭН

НАДЕЖНЫЕ ВЫЧИСЛЕНИЯ ПРИ НАЛИЧИИ ШУМОВ

Перевод с английского
Е. А. БОЧЕК и В. Г. ЧЕРНОВА

Под редакцией
А. В. ШИЛЕЙКО

ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1968

518
В 49
УДК 519.9

С. Виноград, Дж. Д. Коуэн

Надежные вычисления при наличии шумов

М., 1968 г., 112 стр. с илл.

(Серия: «Физико-математическая библиотека инженера»)

Редакторы *С. А. Беляев* и *С. А. Широкова*

Техн. редактор *Л. А. Пыжова*

Корректор *З. В. Автонеева*

Сдано в набор 2/1 1968 г. Подписано к печати 5/VI 1968 г. Бумага 84×108^{1/2}

Физ. печ. л. 3,5. Условн. печ. л. 5,88. Уч.-изд. л. 5,30.

Тираж 15000 экз. Цена книги 37 коп. Заказ № 182.

Издательство «Наука»

Главная редакция физико-математической литературы

Москва, В-71, Ленинский проспект, 15.

2-я типография Изд-ва «Наука». Москва, Шубинский пер., 10

2-2-3
71-67

ОГЛАВЛЕНИЕ

Предисловие редактора	7
Предисловие редактора английского издания	10
Из предисловия авторов	11
Основные обозначения.	13
Глава первая. Введение	15
Глава вторая. Теория автоматов	18
2.1. Абстрактные нейроны и модули	19
2.2. Модульные сети	19
2.3. Определенные и неопределенные события	20
2.4. Представление событий	21
2.5. Регулярные события	22
2.6. Конечные автоматы	23
2.7. Более сложные модули	24
Глава третья. Теория информации	25
3.1. Количество информации в сообщении	26
3.2. Теорема о кодировании при отсутствии шумов	27
3.3. Некоторые меры информации	29
3.4. Пропускная способность канала связи и теорема о кодировании при наличии шума	31
3.5. Пропускная способность двоичного симметричного канала	34
3.6. Некоторые коды с исправлением ошибок	35
Глава четвертая. Надежность автоматов	39
4.1. Работы фон Неймана	39
4.2. Другие подходы к созданию надежных автоматов	44
4.3. Сравнение с теорией информации	47
4.4. Попытки применения теории информации	48
4.5. Обсуждение	50
Глава пятая. Модульные вычислительные сети	51
5.1. Системы обработки информации	52
5.2. Вычисление с помощью ненадежных модулей	54
5.3. Модульное разбиение	57
5.4. Общий метод	59

Глава шестая. Работоспособность вычислительного канала	60
6.1. Основное неравенство	61
6.2. Некоторые примеры вычислительных каналов . . .	62
6.3. Вычисление булевых функций	64
6.4. Предельная теорема для вычислительного канала с шумом	64
Глава седьмая. Сигнальная и модульная избыточности	68
7.1. Функциональное кодирование	70
7.2. Функциональное кодирование с использованием модульной избыточности	73
Глава восьмая. Анастомотические модульные сети . . .	77
8.1. Расширение ансамбля. Основная теорема	77
8.2. Некоторые примеры надежных модульных сетей . .	83
8.3. Обсуждение теоремы. Связь с другими результатами	89
8.4. Эффект функционального кодирования	93
Глава девятая. Ошибки передач и структур	95
9.1. Синаптический шум	95
9.2. Ошибки в структурах	98
9.3. Обсуждение результатов	104
Глава десятая. Заключение	104
Приложение	107
Литература	111

ПРЕДИСЛОВИЕ РЕДАКТОРА

Проблема повышения надежности работы вычислительных систем—одна из основных проблем, определяющих развитие средств вычислительной техники. В настоящее время имеется большое количество различных методов повышения надежности как отдельных узлов, так и вычислительных систем в целом. Однако разработка общей теории надежности, к сожалению, еще далека от завершения. В своих классических работах Дж. фон Нейман показал принципиальную возможность построения из ненадежных элементов систем, обладающих сколь угодно высокой надежностью. Там же был показан и путь к повышению надежности, состоящий в так называемом «резервировании», т. е. в использовании нескольких однотипных элементов, выполняющих одну и ту же функцию и соединяемых между собой таким образом, что неправильная работа одного из элементов не вызывает неправильной работы всего комплекса. Подобный метод повышения надежности получил также название метода введения аппаратурной избыточности. Методы введения аппаратурной избыточности в различных модификациях широко используются на практике.

В теории связи разработаны также методы введения сигнальной избыточности, сводящейся в основном к тому, что, кроме сигналов, непосредственно несущих информацию, по каналу связи передаются также дополнительные сигналы, используемые на приемном конце канала для обнаружения и исправления ошибок, возникающих из-за ненадежной работы каналов. В обоих случаях повышение надежности достигается, очевидно, ценой дополнительных затрат на введение дополнительной аппаратуры либо на дополнительное время использования канала. При этом существенно важным оказывается то обстоятельство, что при использовании метода аппаратурной избыточности

такие затраты оказываются существенно больше, чем при использовании метода сигнальной избыточности.

В вычислительных системах в настоящее время используют как аппаратную, так и сигнальную избыточность, и, по всей вероятности, наибольшая эффективность может быть достигнута при комбинировании тех и других методов. Однако отсутствие законченной теории не позволяет в настоящее время ставить задачу синтеза вычислительной системы при наименьших дополнительных затратах. Авторы предлагаемой читателю монографии делают попытку сформулировать основы такой теории.

Прежде всего, по аналогии с понятием нейронной сети, введенным Маккаллоком и Питтсом, авторы вводят понятие вычислительной сети как совокупности элементов, обладающей несколькими входами и одним выходом и выполняющей определенные преобразования над входными сигналами. Наличие определенной зависимости между выходным сигналом и входными сигналами позволяет провести аналогию между вычислительной сетью и каналом связи. Однако, если в идеально работающем канале связи выходной сигнал полностью определяет входной, то в вычислительной сети выходной сигнал не содержит достаточного количества информации о всех входных сигналах. Основываясь на подобном наблюдении, авторы определяют вычислительную сеть как канал связи с частичным разрушением информации. Подобное определение в известной мере является спорным, однако оно позволяет применить к решению рассматриваемых вопросов широко развитый в настоящее время аппарат теории информации.

Развивая дальше аналогию между вычислительными сетями и каналами связи, авторы вводят понятие «работоспособности» (computation capacity) вычислительного элемента (модуля), аналогичное понятию пропускной способности канала связи. Ненадежный элемент, т. е. элемент, для которого вероятность неправильного срабатывания (сбоя) отлична от нуля, авторы рассматривают как последовательное соединение идеального вычислительного элемента и канала связи с шумом.

Переходя к рассмотрению метода введения аппаратной избыточности, авторы приходят к выводу, что в избыточной вычислительной сети за счет общего увеличения ко-

личества компонент доля участия каждого компонента в выполнении общей работы в известной мере снижается. На этом основании авторы вводят понятие «удельной» нагрузки (computation ratio) на компонент, аналогичное понятию энтропии на символ, используемому в теории связи. Опираясь с понятиями работоспособности и удельной нагрузки, авторы почти без изменений распространяют основные положения теории связи на случай вычислительных сетей. В частности, это дает им возможность доказать основную предельную теорему, аналогичную известной теореме Шеннона. Полученные результаты позволяют, хотя бы в простейших случаях сетей, реализующих булевы функции, ставить задачу формального синтеза модулей, обладающих заданной надежностью при минимальных затратах.

Работа в целом представляет собой одну из первых попыток в рассматриваемой области, содержит ряд спорных положений и полученные в ней результаты требуют дополнительной проработки. Однако сама идея проведения столь тесной аналогии между системами связи и вычислительными системами представляется нам весьма плодотворной и, безусловно, интересной. Учитывая новизну и сложность вопроса, при переводе и редактировании книги мы пытались в наибольшей возможной степени сохранить стиль и терминологию авторов. Целый ряд терминов, таких, как вычислительный канал, вычислительный модуль, работоспособность, удельная нагрузка и т. д., не являются принятыми в русской литературе и сохранены нами в переводе только для того, чтобы дать читателю наилучшее представление о стиле изложения оригинала. Главы 1—5 переведены Е. А. Бочек, а главы 6—10 — В. Г. Черновым.

А. В. Шилейко

ПРЕДИСЛОВИЕ РЕДАКТОРА АНГЛИЙСКОГО ИЗДАНИЯ

В науке и технике давно появилась необходимость систематической публикации исследовательских работ бóльшего масштаба, чем журнальные статьи, но не носящих характера законченной книги. Очень ценные работы такого рода публикуются в настоящее время полуофициальным путем, например в виде лабораторных отчетов, и поэтому не могут занять надлежащее место в литературе по соответствующей отрасли.

Настоящее издание является двадцать вторым из серии монографий, выпускаемой издательством Массачусетского технологического института (MIT Press). Мы надеемся, что эта серия сделает актуальные и важные исследовательские работы легко доступными для библиотек и отдельных читателей.

Дж. Э. Стреттон

ИЗ ПРЕДИСЛОВИЯ АВТОРОВ

«Где нет ответа,— нет и вопроса».

Людвиг Виттгенштейн

Одним из самых интересных теоретических вопросов, возникших при разработке методов создания надежных автоматов из недостаточно надежных элементов, был вопрос о возможности применения теории информации и теории кодирования к решению этой проблемы. В работе, посвященной построению надежных автоматов из избыточных схем, Дж. фон Нейман использовал очень примитивные корректирующие коды и закон больших чисел. Его результат оказался несовместимым с основной теоремой теории информации, теоремой о кодировании для канала с шумом К. Э. Шеннона. Эта теорема определяет предел избыточности, требуемый для получения любого заданного уровня надежности связи по каналу с шумом при условии, что имеются достаточно сложные устройства для кодирования и декодирования, и доказывает существование по меньшей мере одного корректирующего кода с избыточностью близкой, но не меньшей, чем предел, обеспечивающий такую надежность. Цель данной монографии заключается в том, чтобы показать возможность распространения теоремы о кодировании на случай вычислений с помощью ненадежных элементов и возможность использования корректирующих кодов при разработке надежных автоматов из менее надежных элементов, несмотря даже на случайные ошибки в схемах соединений.

Главы 1—4 содержат краткую историю вопроса, необходимый пояснительный материал по теории автоматов и теории информации и анализ предшествующих работ в этой

области, наиболее примечательными из которых являются работы Дж. фон Неймана и П. Элайса. В главе 5 заново формулируется задача и с помощью понятия меры информации выявляется различие между каналами связи и вычислительными системами. В главе 6 дается определение понятия работоспособности элементов (модулей) с шумом, аналогичное понятию пропускной способности канала связи с шумом.

Работоспособность определяет нижний предел избыточности, необходимый для выполнения заданных вычислений при наличии шумов. Здесь имеется в виду не избыточность сигнала, а модульная избыточность. Эта мысль подробно развивается в главе 7. В главе 8 доказывается первая основная теорема, представляющая собой распространение теоремы о кодировании для канала связи с шумом на вычисления с помощью «шумовых» модулей, и на соответствующих примерах рассматриваются ее следствия. Глава 9 содержит вторую основную теорему, которая определяет возможность компенсации воздействия ошибок в схеме соединений за счет модульной избыточности. Книга заканчивается главой 10, в которой отмечается, что решающим фактором во всех схемах, предназначенных для повышения надежности автоматов, является модульная сложность.

Указанные теоремы отвечают на вопрос, как заменить избыточность модульной сложностью, требуя допущений, что не существует границ сложности модуля и что ошибки в модулях не растут слишком быстро с увеличением сложности.

Работа, изложенная в монографии, проводилась в то время, когда авторы были членами группы нейрофизиологии Исследовательской лаборатории электроники Масачусетского технологического института.

С. Виноград

Исследовательский центр Томаса Ф. Уотсона,
Йорктаун Хейтс, Нью-Йорк.

Дж. Д. Коуэн

Королевский колледж, Лондонский университет,
Лондон, Англия.

Июнь 1963 г.

ОСНОВНЫЕ ОБОЗНАЧЕНИЯ

m — модуль	x_i, y_i — алфавитные символы
t — время	X, Y — алфавитные ансамбли
θ — порог возбуждения модуля	$\text{Pr} ()$ — вероятность $()$
τ — задержка в модуле	$H(X)$ — количество информации в сообщении X
$y(t)$ — выходной сигнал модуля в момент времени t	k — количество символов на алфавит
1 — возбужденное состояние модуля	k — длина последовательности сообщений
0 — невозбужденное состояние модуля	η — количество различных сообщений
$P(t)$ — логическое высказывание, реализованное в момент времени t	\bar{n} — среднее количество символов на сообщение
\equiv — равнозначность	n — длина кодового слова
$\&$ — конъюнкция	I — количество взаимной информации
\neg — отрицание	C — пропускная способность канала связи
\vee — дизъюнкция	S — источник сообщений
E, F — регулярные события	
$*$ — см. текст (раздел 2.5)	
A_c, B_d — см. текст (раздел 2.6)	

P — вероятность ошибки, связанной с кодом	при заданном входе x_α
\sim — вероятность неправильного функционирования модуля	z, Z — см. текст (раздел 5.2)
Δ, ξ — см. текст (раздел 4.1)	C^* — работоспособность канала вычислений
$s(x_1, x_2) = \overline{(x_1 \& x_2)}$	f, g — функции
N — модульная избыточность	e — кодирующая функция
R — удельная нагрузка	d — декодирующая функция
\oplus — сложение по модулю 2	A, A, A — автомат
x_α — входной вектор	M — количество модулей в автомате
X — ансамбль входных векторов	h — см. текст (раздел 8.1)
$\text{Pr}(y_\beta x_\alpha)$ — вероятность получения y_β в качестве выхода	c_{ij}, c_{rn} — см. текст (раздел 9.2)
	G — см. текст (раздел 9.2)

ГЛАВА ПЕРВАЯ

ВВЕДЕНИЕ

В течение последнего десятилетия много внимания было уделено проблеме создания надежных машин из недостаточно надежных элементов. Это вызвано несколькими причинами, одна из которых связана с разработкой больших автоматических цифровых вычислительных машин. Переключательные схемы, входящие в состав арифметических устройств и устройств управления таких систем, требуют взаимного соединения большого числа переключательных элементов. Если эти элементы и их соединения не являются достаточно надежными, то вероятность появления отказов или сбоев в работе таких устройств повышается. Быстродействие, размеры и сложность вычислительных машин непрерывно растут, и в связи с этим кажется маловероятным, несмотря на последние достижения в области микроэлектроники, что надежность элементов станет настолько высокой, чтобы сделать возможным прямой синтез из них сложных вычислительных устройств, достаточно надежных для практического применения.

Другая причина связана с исследованиями организации и функционирования живых систем. Имеются доказательства того, что системы значительно более сложные, чем большие цифровые вычислительные машины, такие как нервная система животных и внутриклеточные структуры, участвующие в макромолекулярном биосинтезе, функционируют с высокой надежностью в течение долгого времени, несмотря на то, что они состоят из относительно ненадежных элементов.

Теоретические подходы к решению этой проблемы сводились к исследованиям возможности создания надежных конечных автоматов из элементов с низкой надежностью.

В работах Дж. фон Неймана [41, 42] и Мура и Шеннона [26] была показана практическая возможность построения из таких элементов схем более высокой надежности. Высокая надежность достигалась за счет введения в схемы большего, чем это теоретически необходимо для выполнения заданных вычислений, числа элементов (*избыточность*), соединение которых обеспечивало контроль сбоев. Их трактовки до некоторой степени отличались, так как Мур и Шеннон рассматривали релейно-контактные схемы, которые подвержены ошибкам только на входах, в то время как Дж. фон Нейман рассматривал вентильные схемы [20, 21], где ошибки происходят на выходе схем. Поэтому результаты их были различны, но общим в них было требование большой структурной избыточности для получения высокой надежности работы схем. В частном случае вентильных схем (в этой книге мы будем иметь дело только с ними) расчеты фон Неймана требовали избыточности по меньшей мере порядка $10^3:1$. Фон Нейман не был удовлетворен этим результатом и считал, что он является следствием зависимости его теоретических предпосылок от использованного им математического и теоретико-логического аппарата. Так, фон Нейман [40] писал:

«Теория цифровых автоматов типа «все или ничего»... является, несомненно, главой формальной логики..., технически одной из самых негибких областей математики,... (имеющей дело)... с жесткими представлениями «все или ничего» и очень слабо связанной с математическим анализом,... технически наиболее успешной областью математики ... Логика автоматов будет отличаться от нынешней системы формальной логики в двух аспектах: 1) действительная длина... цепей операций должна будет учитываться; 2) логические операции... должны будут трактоваться как процедуры, которые допускают... сбой с малой, но ненулевой вероятностью. Все это приведет к построению теории, в значительно меньшей степени базирующейся на принципе «все или ничего», чем прошлая и современная формальная логика... Эта новая система формальной логики подойдет вплотную к другой дисциплине, в прошлом мало связанной с логикой. Этой дисциплиной является термодинамика, главным образом в той ее форме, в которой она была разработана Больцманом, составляющая ту часть теоретической физи-

ки, подходящую ближе всего в некоторых аспектах к пониманию и измерению информации».

Это указывает на применимость теории информации (Габор [14], Шеннон [35]) к задачам исследования надежности. Элайс [10] впервые отметил, что результат фон Неймана (и Мура и Шеннона), по-видимому, не согласуется с основной теоремой теории информации, а именно с теоремой о кодировании, рассматривающей надежную передачу информации по каналу с шумом (Шеннон [35]).

Были предприняты попытки (Элайс [10], Питерсон и Рэбин [29], Иден [9], Коуэн [7]) распространить эту теорию на более общий случай вычислительных устройств с шумом путем использования корректирующих кодов (Хэмминг [17]) при обработке информации. Эти попытки не увенчались успехом, и сложилось мнение, что первоначальное решение фон Неймана и последующие вариации (Аллансон [1], Мурог [27], Вербеек [39], Блюм [3], Коуэн [7], Маккаллок [23]), вероятно, окажутся единственно возможными для задачи построения надежных автоматов из относительно ненадежных элементов.

В данной работе доказывается, однако, что при некоторых допущениях, не сделанных Элайсом и другими, теорема Шеннона о кодировании для канала связи с шумом может быть распространена на вычисления при наличии шумов. Точнее говоря, мы показываем, как можно определить работоспособность (соответствующую пропускной способности канала связи с шумом) ненадежных модулей, реализующих логические функции. Мы доказываем, что определенные события (события конечной длительности) могут быть реализованы с произвольно высокой надежностью в схемах, собранных из модулей и соединений низкой надежности, при условии обеспечения достаточной избыточности проектируемых схем. Если сделаны допущения, что модули достаточно сложны и сбор в модулях не зависит от их сложности, то необходимо, чтобы модульная избыточность таких схем была больше некоторого минимального значения, определяемого работоспособностью модулей, составляющих эти схемы.

Далее показано, что если частота ошибок в соединениях модулей меньше, чем некоторая функция работоспособности и модульной избыточности, то с большой вероятностью

такие ошибки могут быть исправлены, и можно построить автоматы, работающие с произвольно высокой надежностью, несмотря на ошибки в структурах.

Рассматриваемые нами автоматы не имеют четкой функциональной организации. Любая функция, вычисляемая автоматом, реализуется многими модулями, и любой модуль вычисляет одновременную композицию таких функций. Получающаяся в результате многократная одновременность функции непосредственно связана с высокой надежностью, проявляемой такими автоматами.

ГЛАВА ВТОРАЯ

ТЕОРИЯ АВТОМАТОВ

В этой и последующих главах нам неоднократно придется рассматривать некоторые аспекты взаимосвязи теории автоматов и теории информации. Для облегчения понимания последующего материала мы дадим очень краткое изложение некоторых основных положений и терминологии этих дисциплин.

Современные представления об автоматах вытекают из законов математической логики как дедуктивной науки, основанной Булем [5], развитой затем Уайтхедом, Расселом [43] и Гёделем [16]. Тьюринг [38] свел формулировки дедуктивных доказательств теорем к определению того, может или не может некоторый вычислительный автомат вычислить определенное число.

Маккаллоком и Питтс [24] использовали этот результат, чтобы показать, что любая деятельность, будь то интроспективная, бихевиористическая или физиологическая, может быть реализована вычислительной сетью, если она может быть однозначно описана конечным числом слов. Такая сеть состоит из элементов, называемых «абстрактными нейронами», свойства которых были грубым приближением известных свойств нервных клеток и процессов, происходящих в них. Использование Маккаллоком и Питтсом двузначного логического исчисления высказываний и функций высказываний (см. Уайтхед и Рассел [43]), вместе с применением Шенноном исчисления высказываний для описания релейных схем (Шеннон [34]) озаменовало

первое приложение алгебры двузначной логики к проблемам, относящимся к сложным вычислительным машинам.

Клини [18] вновь пришел к первоначальным результатам Маккаллока и Питтса в более наглядной форме, и мы, давая краткое изложение теории конечных автоматов, будем использовать его формулировки.

2.1. Абстрактные нейроны и модули

По определению Маккаллока и Питтса [24] *абстрактный нейрон* состоит из тела (сомы), откуда исходит проводник (аксон), ведущий к одному или нескольким окончаниям, каждое из которых является либо возбуждающим, либо тормозящим.

Мы будем называть такую систему *модулем*. Через равные промежутки времени каждый модуль либо возбужден, либо находится в состоянии покоя. Изменение его активности имеет природу «да — нет».

Это значит, что модуль возбужден в момент времени t тогда и только тогда, если некоторое минимальное число θ (θ называется *порогом* возбуждения модуля) контактирующих с ним возбуждающих окончаний связано и ни одно из тормозящих окончаний (абсолютное торможение) не связано с модулями, которые были возбуждены в момент $t - \tau$; величина результирующей активности модуля постоянна и не зависит от действительного числа контактирующих с ним возбуждающих окончаний. Полагают $y(t) = 1$, если модуль возбужден в момент t и $y(t) = 0$, если он находится в состоянии покоя. Будем считать также, что при работе таких модулей (рис. 2.1) ошибки не имеют места. Такие модули являются идеальными переключательными элементами, обладающими порогом срабатывания θ и временем задержки τ . В этой работе мы принимаем, что τ всегда равно единице измерения времени.

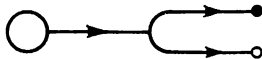


Рис. 2.1. Модуль с одним возбуждающим (черный кружок) и одним тормозящим окончанием.

2.2. Модульные сети

Модульная схема или просто *сеть* — это соединение конечного числа модулей, в котором каждое окончание любого модуля контактирует с сомой не более чем одного другого модуля. Разделительный промежуток называется

синапсом. Принимается, что передача активности через синапсы происходит без искажений. Сети могут быть *циклическими* или *ациклическими*. Сеть называется циклической, если она содержит цикл,

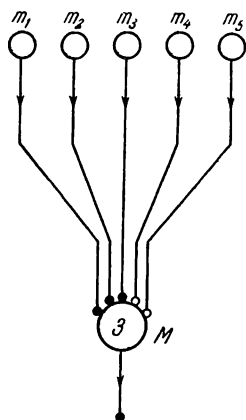


Рис. 2.2. Конъюнктивная модульная сеть, реализующая

$$P(t) \equiv y_1(t-1) \& y_2(t-1) \& y_3(t-1) \& \bar{y}_4(t-1) \& \bar{y}_5(t-1);$$

выходной модуль M имеет порог, равный 3.

т. е. если существует цепь модулей m_1, m_2, \dots, m_p , каждый элемент которой контактирует со следующим, а m_p соединяется с m_1 . В противном случае сеть называют ациклической.

Активность модулей можно выразить с помощью формул двузначной логики высказываний и функций высказываний. Так, формула

$$P(t) \equiv y_1(t-1) \& y_2(t-1) \& y_3(t-1) \& \bar{y}_4(t-1) \& \bar{y}_5(t-1),$$

где $P(t)$ — состояние «выходного» модуля M в момент времени t , а $y_i(t-1)$ — состояния «входных» модулей m_i в момент $t-1$, выражает логически тот факт, что модуль возбужден в момент t в том и только в том случае, если модули m_1, m_2, m_3 были возбуждены в момент $t-1$, а модули m_4 и m_5 находились в состоянии покоя (рис. 2.2). Символ « \equiv » означает

«тогда и только тогда», $\&$ означает «И», « $\bar{}$ » означает «НЕ». Кроме того, общепринятым является символ « \vee », обозначающий «включающее ИЛИ».

2.3. Определенные и неопределенные события

Эквивалентное представление модульной активности можно дать следующим образом. Вход в модульную сеть (осуществляемый входными модулями) за все прошедшее время до данного момента t может быть описан таблицей, у которой столбцы соответствуют входным модулям, а строки — прошедшим моментам времени $t-1, t-2, \dots$, и т. д.

Таблица 2.1

Таблица входа

	m_1	m_2	m_3	m_4	m_5
$t-1$	1	1	0	1	0
$t-2$	1	1	1	0	0
$t-3$	0	1	1	0	1

Любой подкласс класса всевозможных таблиц, описывающих вход за все прошедшее время, представляет собой *событие*, которое имеет место, если таблица, описывающая данный вход, принадлежит этому классу. Так, например, событие, соответствующее

$$P(t-1) \equiv \\ \equiv y_1(t-2) \& y_2(t-2) \& y_3(t-2) \& \bar{y}_4(t-2) \& \bar{y}_5(t-2),$$

содержится в табл. 2.1 *).

Если события относятся к фиксированному отрезку времени, состоящему из последовательных моментов $t - \xi + 1, t - \xi + 2, \dots, t - 1$ (где $\xi \geq 1$), то они называются *определенными событиями* длительности ξ . В противном случае, они называются *неопределенными событиями*.

2.4. Представление событий

Главная задача рассматриваемой теории — это представление событий, т. е. построение модульных сетей, которые будут «реализовывать» данные формулы, а также нахождение формул, реализуемых в данных модульных сетях. Мы не будем вникать в технические детали доказательств следующих теорем, поскольку непосредственный интерес представляют лишь результаты:

1. Каждому... определенному событию соответствует модульная ациклическая сеть, которая представляет это событие возбуждением некоторого внутреннего модуля в момент $t + \tau$.

*) Точнее, если имеет место событие, описываемое табл. 2.1, то функция $P(t)$ принимает значение 1 (истинно). (Прим. ред.)

2. Любое событие, которое представимо в модульной ациклической сети возбуждением данного внутреннего модуля в момент $t + \tau$ ($\tau \geq 1$) — определенное.

Таким образом, существует прямая связь между определенными событиями и ациклическими сетями.

2.5. Регулярные события

Существует также связь между неопределенными событиями и циклическими сетями. Для выявления этой связи необходимо пересмотреть понятия неопределенных и определенных событий и ввести новое понятие — *регулярное событие*. Регулярное событие определяется следующим образом. Пусть имеются множества таблиц E и F . Тогда $E \cup F$ (их объединение) — это множество таблиц, содержащее все без исключения таблицы, принадлежащие к E или F . Аналогично, $E \cdot F$ (их произведение) есть множество таблиц, получаемых в результате приписывания любой таблицы из F сразу под любой таблицей из E , причем входы таблиц из множества E существуют только для моментов времени $t - \xi + 1, \dots, t$. Наконец, $E * F$ (итерация E по F)

определяется как $F \cup EF \cup EEF \cup \dots$ или $\sum_{n=0}^{\infty} E^n F$. Ре-

гулярные множества таблиц — это наименьший класс множеств таблиц, который включает в себя пустое множество и множества, содержащие одну таблицу каждое (единичные множества), и является замкнутым относительно операций перехода от E и F к $E \cup F$, $E \cdot F$ и $E * F$. Событие является регулярным, если существует регулярное множество таблиц, описывающих его, т. е. событие имеет место тогда и только тогда, когда вход описывается по крайней мере одной из таблиц этого множества. Из этих определений следует, что все определенные события и некоторые (но не все) неопределенные события являются регулярными событиями. Вновь можно доказать ряд теорем, связывающих регулярные события и модульные сети. В частности, представляет интерес следующая теорема:

3. Каждому регулярному событию соответствует модульная сеть, представляющая это событие воз-

буждением некоторого внутреннего модуля в момент $t + \tau$ при условии, что все внутренние модули в начальный момент находятся в соответствующих состояниях.

2.6. Конечные автоматы

Следуя Клини, можно рассматривать модули с более общими свойствами, в которых число состояний активности больше двух. При этом входные модули имеют только два состояния, а внутренние модули могут иметь разное число состояний. Модульная сеть такого общего типа, включая, конечно, более простые сети, называется *конечным автоматом*. При a входных модулях m'_1, m'_2, \dots, m'_a ($a \geq 0$) и b внутренних модулях $m''_1, m''_2, \dots, m''_b$ ($b \geq 1$) с числом состояний n_1, n_2, \dots, n_b имеется точно $2^a \cdot n_1 \cdot n_2 \cdot \dots \cdot n_b$ возможных конечных состояний автомата. Каждое конечное состояние можно рассматривать как комбинацию какого-то внешнего состояния из 2^a возможных состояний и внутреннего состояния из $n_1 \cdot n_2 \cdot \dots \cdot n_b$ возможных вариантов. Обозначим конечные состояния через A_1, \dots, A_c ($c = 2^a \cdot n_1 \cdot \dots \cdot n_b$) и внутренние состояния через B_1, \dots, B_d ($d = n_1 \cdot \dots \cdot n_b$). Тогда можно доказать следующую теорему:

4. В любом конечном автомате (в частности, в сети Маккаллока — Питтса), находящемся в момент t_1 в некотором состоянии B_1 , событие, представленное состоянием, существующим в момент t , является регулярным.

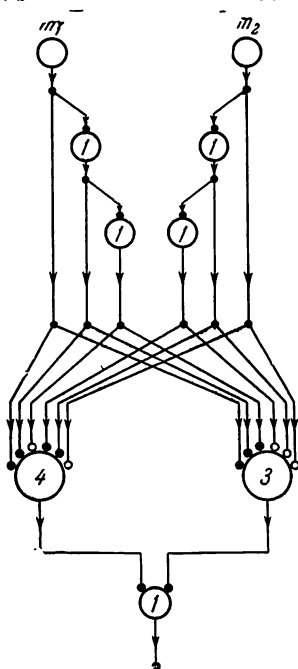


Рис. 2.3. Реализация определенного события $E \cup F$, описанного в табл. 2.2.

Следовательно, можно определить «поведение» конечного автомата как зависимость между его входами и выходами для различных моментов времени.

Далее в этой монографии мы будем иметь дело только с определенными событиями и представляющими их системами. В качестве примера к предшествующему изложению рассмотрим определенное событие $E \cup F$, заданное табл. 2.2.

Таблица 2.2
Определенное событие $E \cup F$

E	m_1	m_2
t	1	0
$t-1$	1	1
$t-2$	0	1

F	m_1	m_2
t	1	0
$t-1$	1	0
$t-2$	1	0

Эту же таблицу можно описать символически следующим образом *):

$$P(t+1) \equiv (y_1(t) \& \bar{y}_2(t) \& y_1(t-1) \& y_2(t-1) \& \bar{y}_1(t-2) \& y_2(t-2)) \vee (y_1(t) \& \bar{y}_2(t) \& y_1(t-1) \& \bar{y}_2(t-1) \& y_1(t-2) \& \bar{y}_2(t-2))$$

и реализовать в ациклической модульной сети, показанной на рис. 2.3. Таким образом, это событие реализуется в ациклической модульной сети с временем задержки, равным 2.

2.7. Более сложные модули

Маккаллоу [22] ввел понятие более гибкого и сложного модуля со следующими свойствами: а) запрет является относительным, т. е. активность модуля зависит только от того, превосходит алгебраическая сумма возбуждений и торможений порог модуля или нет; б) любое число оконча-

*) Смысл приводимой логической формулы состоит в следующем: функция $P(t+1)$ принимает значение 1 (истинно) в том и только в том случае, если имеет место событие, определяемое заданной таблицей (в данном случае, таблицей $E \cup F$). (Прим. ред.)

ний одного модуля может контактировать с любым другим модулем; в) имеет место афферентное запрещение, т. е. тормозящее окончание может блокировать передачу активности в другом контактирующем с ним окончании; г) порог модуля θ может систематически изменяться с дискретным шагом. Например, действие модульной сети; показанной на рис. 2.4, соответствует операции «сумма по модулю 2», с временем задержки, равным 1, если принято, что при афферентном запрещении задержка не имеет места. Если допустить, что при афферентном запрещении задержка существует, то модуль реализует функцию

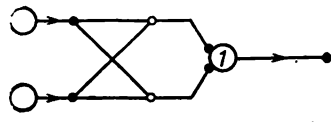


Рис. 2.4. Сложный модуль (с афферентным взаимодействием).

$$P(t+1) \equiv y_1(t-1) \& \bar{y}_2(t-1) \vee \bar{y}_1(t-1) \& y_2(t-1).$$

В дальнейшем мы примем допущение о нулевой задержке при афферентном запрещении и оставим вопрос физиологического и биологического правдоподобия открытым. Это означает, по существу, что любое высказывание, соответствующее таблице длины 1, может быть реализовано в модульной сети с временем задержки, равным 1 вместо 2. (См. Блум [3])

ГЛАВА ТРЕТЬЯ

ТЕОРИЯ ИНФОРМАЦИИ

Теория конечных автоматов, которую мы кратко рассмотрели, имеет дело с идеальными автоматами, т. е. с сетями, элементы которых (модули и соединения) совершенно свободны от ошибок любого рода. Предпосылкой к рассмотрению автоматов с неидеальными элементами является понимание роли информации и шума в таких автоматах. В соответствии с этим мы дадим краткое (по Фано [11, 12]) изложение шенноновской математической теории связи (Шеннон [35]), рассматривающей вопросы передачи и кодирования информации.

Теория Шеннона содержит элементы, уже присутствующие в статистических теориях вещества, но объектом ее изучения являются не свойства материи. Вместо этого изучается весьма специфическая система связи, состоящая из *источника* сообщений, *канала* связи и *приемника* сообщений, как это показано на рис. 3.1.



Рис. 3.1. Система связи.

Источник и приемник выбирают символы из *алфавита*. Этот алфавит состоит из вполне определенного множества различных знаков и известен как источнику, так и приемнику. Источник последовательно выбирает символы из алфавита, составляя *сообщение*, которое передается в виде *физических сигналов* по каналу к приемнику. В приемнике эти сигналы управляют выборкой символов. Таким образом, сообщение посылается и принимается.

3.1. Количество информации в сообщении

По определению Шеннона *количество информации*, содержащееся в сообщении, в среднем равно

$$H(X) = - \sum_{i=1}^k \text{Pr}(x_i) \log_2 \text{Pr}(x_i), \quad (3.1)$$

где x_i ($i = 1, 2, \dots, k$) — любой символ, выбранный из алфавита X символов (x_1, x_2, \dots, x_k) , встречающихся с вероятностями $\text{Pr}(x_i)$ ($i = 1, 2, \dots, k$) соответственно. Логарифмическая функция, следуя Больцману [4], была выбрана из-за свойства преобразования мультипликативных вероятностей в сумму логарифмов, обеспечивая, таким образом, аддитивность меры информации независимых символов.

Итак, величина $I(x_i) = -\log_2 \text{Pr}(x_i)$, называемая *собственной информацией* символа x_i , встречающегося с вероятностью $\text{Pr}(x_i)$ и взятого из алфавита (x_1, x_2, \dots, x_k) ,

была выбрана в качестве основной меры информации, и среднее значение $I(x_i)$, усредненное по ансамблю (данному алфавиту символов с их вероятностями), представляет собой среднюю собственную информацию на символ данного алфавита. Очевидно, что величина $H(X)$ имеет формальное сходство с функцией энтропии Больцмана — Гиббса (Больцман [4], Гиббс [15]), которая также определялась как среднее по ансамблю.

В самом деле, Шеннон назвал $H(X)$ *энтропией* источника (Шеннон [35]). С тех пор применялись различные термины такие, например, как энтропия связи, комэнтропия и тому подобное. Винер [44], Бриллюэн [6] и другие связывали величину $H(X)$ с понятием информации, рассмотренным в работе Сциларда (Сцилард [37]), и, в соответствии с этим, использовали термин негэнтропия для $H(X)$. Мы, однако, не будем употреблять термин «энтропия» по отношению к среднему количеству собственной информации на символ сообщения источника и приемника, а будем вместо этого употреблять термин «количество информации».

3.2. Теорема о кодировании при отсутствии шумов

Тем не менее, одно лишь использование этого количества информации не составляет теории информации. Сущность теории следует из содержания двух предельных теорем, рассматривающих результат некоторых операций *кодирования*. Термин «код» употребляется для обозначения системы сигналов, например, системы слов, произвольно используемых вместо других слов или фраз с целью обеспечения краткости или секретности. В процессе связи такие сообщения, как печатные тексты, телеграммы и т. д., обычно кодируются в некоторые стандартные формы, например, в последовательность двоичных цифр. Это делается по различным соображениям, в частности, по соображениям экономичности передачи информации в закодированном виде. Ясно, что чем меньше используется двоичных цифр, тем лучше. Первая предельная теорема Шеннона (теорема о кодировании) определяет среднее число двоичных цифр, необходимых для кодирования сообщений, выбранных из данного ансамбля.

Теорема 3.1. При заданном ансамбле X из η сообщений, содержащем количество информации $H(X)$, и алфавите, состоящем из k символов, можно закодировать сообщения таким образом, что среднее число символов на сообщение \bar{n} будет удовлетворять неравенству

$$\frac{H(X)}{\log_2 k} \leq \bar{n} < 1 + \frac{H(X)}{\log_2 k}, \quad (3.2)$$

и число \bar{n} не может быть сделано меньше, чем нижняя граница.

Относительные разности между верхними и нижними границами могут быть выбраны бесконечно малыми за счет выбора сообщений из ансамблей, содержащих большие количества информации.

В частности, если сами сообщения состоят из отрезков последовательностей независимых символов x_i , выбираемых из ансамблей, содержащих одинаковое количество информации $H(X_i)$, то $H(X) = nH(X_i)$; откуда, подставляя этот результат в уравнение (3.2) и беря предел, получаем зависимость

$$\lim_{n \rightarrow \infty} \frac{1}{n} (\log_2 k^n) = H(X_i). \quad (3.3)$$

Таким образом, мы непосредственно получаем действующее значение количества информации, содержащейся в ансамбле.

Т а б л и ц а 3.1

Код

Номер сообщения	Вероятность сообщения	Двоичное кодовое слово	$-\log_2 P_i$
0	0,25	00	2
1	0,25	01	2
2	0,125	100	3
3	0,125	101	3
4	0,0625	1100	4
5	0,0625	1101	4
6	0,0625	1110	4
7	0,0625	1111	4

как асимптотическое приближение к минимальному числу двоичных цифр, требуемых в среднем для кодирования сообщения. В качестве примера приведем код, показанный в табл. 3.1, для которого $\bar{n} = H(X)$.

3.3. Некоторые меры информации

Вторая предельная теорема Шеннона связана с передачей закодированного сообщения по физическим каналам, т. е. по системам, которые могут передавать или распространять физические сигналы. Такие каналы подвержены помехам (инженеры-связисты называют их шумом) так, что, в общем случае, сообщения, принятые из канала, каким-то образом искажены. Сообщения на входе в канал не могут быть однозначно определены по сообщениям на выходе из канала и при их определении может произойти ошибка. Для характеристики воздействия помех определяется статистическая мера информации, сообщаемой одним символом о другом символе.

В качестве такой меры принимают величину

$$I[x_i; y_j] = -\log_2 (\text{Pr}(x_i)/\text{Pr}(x_i|y_j))$$

и, очевидно,

$$I[x_i; y_j] = I(x_i) - I(x_i|y_j). \quad (3.4)$$

Другими словами, разность между собственной информацией символа x_i до и после выбора символа y_j является мерой информации, сообщаемой символом y_j о символе x_i . Эта мера называется *взаимным количеством информации символов x_i и y_j* .

Можно доказать, что

$$I[x_i; y_j] = I[y_j; x_i], \quad (3.5)$$

$$I[x_i; y_j] \leq I(x_i); I[x_i; y_j] \leq I(y_j), \quad (3.6)$$

$$I[X; y_j] = \sum_{i=1}^k \text{Pr}(x_i|y_j) I[x_i; y_j] \geq 0, \quad (3.7)$$

т. е., что собственная информация любого символа x_i представляет собой максимальное количество информации, которое он может сообщить о любом символе. И если

$X = (x_1, x_2, \dots, x_{k_1})$ и $Y = (y_1, y_2, \dots, y_{k_2})$, то y_j ($j = 1, 2, \dots, k_2$) обеспечивает среднее неотрицательное количество информации относительно X .

Среднее количество информации, обеспечиваемое символами Y и X , в этом случае равно

$$\begin{aligned} I[X; Y] &= \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \Pr(x_i, y_j) I[x_i; y_j] = \\ &= \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \Pr(x_i, y_j) \log_2 \frac{\Pr(x_i, y_j)}{\Pr(x_i) \Pr(y_j)}. \end{aligned} \quad (3.8)$$

Если $H(X|Y)$ и $H(Y|X)$ определены как

$$\begin{aligned} H(X|Y) &= \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \Pr(x_i, y_j) I(x_i|y_j) = \\ &= - \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \Pr(x_i, y_j) \log_2 \Pr(x_i|y_j), \end{aligned} \quad (3.9)$$

$$\begin{aligned} H(Y|X) &= \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \Pr(x_i, y_j) I(y_j|x_i) = \\ &= - \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \Pr(x_i, y_j) \log_2 \Pr(y_j|x_i), \end{aligned} \quad (3.10)$$

то легко показать, что

$$I[X; Y] = H(X) - H(X|Y), \quad (3.11)$$

$$I[X; Y] = H(Y) - H(Y|X). \quad (3.12)$$

Итак, среднее значение взаимной информации $I[x_i; y_j]$ можно представить как разность между средними количествами информации, необходимыми для выбора x_i до и после выбора y_j . Величину $H(X|Y)$ часто называют *неопределенностью*, так как она представляет собой неопределенность относительно x_i , которая остается после выборки y_j . Иначе $I[X; Y]$ можно интерпретировать как разность между средним количеством информации, которую y_j может сообщить, и количеством информации, необходимым для определения помех в канале. Другими словами, мы считаем, что символы

x_i поступают в канал с шумом, а символы y_j выходят из него. Величина $H(Y | X)$, которая определяет помехи в канале, называется в литературе *энтропией шума*. Во всяком случае, $I[X; Y]$, несомненно, содержит элементы, необходимые для того, чтобы характеризовать канал с шумом.

3.4. Пропускная способность канала связи и теорема о кодировании при наличии шума

Пусть

$$C = \max_x I[X; Y] \quad (3.13)$$

будет максимальным средним значением взаимной информации x_i и y_j , полученным путем изменения вероятностей передачи символов. Величина C , называемая *пропускной способностью канала связи*, является верхней границей среднего количества информации, которое может быть обеспечено каждым принятым символом относительно соответствующего переданного символа и всех предшествовавших символов.

Вторая теорема Шеннона о кодировании рассматривает каналы с шумом и ее главное содержание заключается в том, что при некоторых обстоятельствах вероятность появления ошибки опознавания может быть сделана сколь угодно малой за счет надлежащего кодирования и декодирования последовательностей сообщений. Перед тем как перейти к подробному обсуждению теоремы, заметим, что в общем случае каналы имеют память, т. е. значения $P(y_j | x_i)$ в любой данный момент времени зависят от предшествовавших условных вероятностей. Мы будем рассматривать, однако, только каналы без памяти, в которых все вероятности независимы от предшествовавших. Вторая предельная теорема, таким образом, касается передачи последовательностей символов по каналу с шумом, не обладающему памятью. Эту теорему можно сформулировать по-разному, но мы приведем ее первоначальную формулировку.

Т е о р е м а 3.2. *Если задан дискретный источник Δ , выбирающий сообщения из ансамбля X , содержащего количество информации $H(X)$, и дискретный канал с шумом без памяти с пропускной способностью C , то существует по крайней мере один код, обладающий тем свойством, что*

если $H(X) \leq C$, то надлежащим образом закодированные сообщения могут быть переданы по каналу и приняты с произвольно малой частотой ошибки опознавания.

Если $H(X) > C$, то можно закодировать сообщения таким образом, что неопределенность $H(X | Y)$ (где Y — ансамбль принимаемых сообщений) будет меньше, чем $H(X) - C + \epsilon$, где ϵ сколь угодно мало. Не существует способа кодирования, дающего неопределенность меньше, чем $H(X) - C$.

Кратко изложим первоначальное неформальное доказательство этой теоремы (Шеннон [35]).

Пусть S_0 — дискретный источник, выбирающий символы из ансамбля X_0 так, что $H(X_0) - H(X_0 | Y_0) = C$, где C — пропускная способность данного дискретного «шумного» канала без памяти и $H(X_0 | Y_0)$ — соответствующая неопределенность. Мы рассматриваем последовательности этих символов длины n как сигналы на входе и выходе из канала. Тогда с большой вероятностью можно утверждать следующее:

1. Передаваемые сообщения состоят приблизительно из $2^{nH(X_0)}$ последовательностей.

2. Принимаемые сообщения состоят приблизительно из $2^{nH(Y_0)}$ последовательностей, где Y_0 — ансамбль на выходе.

3. Каждая последовательность на выходе может быть результатом около $2^{nH(X_0 | Y_0)}$ последовательностей на входе.

4. Каждая последовательность на входе может породить около $2^{nH(Y_0 | X_0)}$ последовательностей на выходе.

Теперь рассмотрим другой источник S , выбирающий символы из ансамбля X , содержащего количество информации $H(X) < C$. Тогда этот источник будет порождать около $2^{nH(X)}$ последовательностей длины n . Будем считать эти последовательности последовательностями сообщений, а последовательности из S_0 последовательностями сигналов и тем самым установим случайное соответствие между ними, кодируя сообщения из S сигналами из S_0 . Вероятность того, что некоторая последовательность сигналов не была выбрана для кодирования сообщения, в этом случае равна

$$(1 - 2^{-nH(X_0)})^{2^{nH(X)}} \cong 1 - 2^{n(H(X) - H(X_0))}.$$

Следовательно, вероятность p того, что для любого принятого сообщения ни один из $2^{nH(X_0|Y_0)}$ сигналов, принадлежащих множеству его возможных причин, не был выбран в качестве кодового слова, иного по сравнению с действительно переданным, равна

$$p = (1 - 2^{n(H(X) - H(X_0))})^{2^{nH(X_0|Y_0)}} \cong \\ \cong 1 - 2^{n(H(X) - H(X_0) + H(X_0|Y_0))} = 1 - 2^{-n(C - H(X))}. \quad (3.14)$$

Ошибка может иметь место только тогда, когда два сигнала из множества возможных причин появления данного принятого сигнала выбраны в качестве кодовых слов. Таким образом, вероятность ошибки равна

$$P = 1 - p. \quad (3.15)$$

То есть вероятность возможной ошибки опознавания дается выражением $P \cong 2^{-n(C - H(X))}$ и стремится к нулю с возрастанием n , если $H(X) \leq C$.

Если $H(X) > C$, то ясно, что C единиц информации (которые принято называть «битами») можно передать, допуская неопределенность $H(X) - C$ бит. Наконец, если бы мы могли закодировать источник с $H(X) = C + a$ битами информации так, чтобы получить неопределенность $H(X|Y) = a$, тогда было бы $H(X) - H(X|Y) > C$. Это противоречит определению пропускной способности C , данному уравнением (3.13), и, тем самым, доказывает третью часть теоремы.

Важно отметить, что вероятность ошибки P является средней (в самом деле, она является средней по всем возможным кодам), и так как она стремится к нулю, то отсюда следует, что существует по меньшей мере одна кодирующая схема, для которой $P \rightarrow 0$. Кроме того, доля всех возможных кодов, имеющих предельные значения более \sqrt{P} , не превосходит \sqrt{P} , и так как $P \rightarrow 0$, почти все кодирующие схемы в пределе сколь угодно близки к идеальной. Любой код, близкий к идеальному, обладает тем свойством, что если сигнал был искажен шумом, то оригинал все же можно опознать, т. е. искажение, вообще говоря, не будет вызывать путаницы в сигналах. Это достигается путем введения избыточности. Из уравнения (3.15) видно, что такое

кодирование уменьшает P только при увеличении n , т. е. за счет передачи более длинных последовательностей сигналов. В некотором смысле этот вывод совпадает с уравнением (3.3), по которому минимальное количество двоичных цифр, требуемых в среднем для кодирования сообщений, обычно приближается к $H(X)$ только с возрастанием n . Таким образом, кодирование сообщений как для канала с шумом, так и без шума требует использования задержек в кодирующих устройствах. Для достижения пределов, указываемых теорией, эти задержки должны быть бесконечно длинными.

3.5. Пропускная способность двоичного симметричного канала

Мы уже приводили пример кода для передачи при отсутствии шума (см. табл. 3.1). Заканчивая главу, мы дадим пример вычисления пропускной способности канала и несколько примеров кодов для каналов связи с шумом.

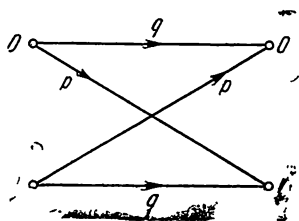


Рис. 3.2. Символическое представление двоичного симметричного канала.

Рассмотрим в качестве простого примера двоичный симметричный, не обладающий памятью канал (ДСК) с шумом, в котором двоичные цифры 1 и 0 принимаются неправильно с вероятностью p и правильно с вероятностью $q = 1 - p$:

$$\left. \begin{aligned} \Pr(y_j = 1 | x_i = 1) &= \Pr(y_j = 0 | x_i = 0) = q, \\ \Pr(y_j = 0 | x_i = 1) &= \Pr(y_j = 1 | x_i = 0) = p. \end{aligned} \right\} \quad (3.16)$$

Это можно представить схематически на рис. 3.2. Легко показать, что

$$\begin{aligned} H(Y | X) &= - \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \Pr(x_i) \Pr(y_j | x_i) \log_2 \Pr(y_j | x_i) = \\ &= -p \log_2 p - q \log_2 q. \end{aligned} \quad (3.17)$$

В силу симметрии это выражение не зависит от вероятностей передачи. Из уравнений (3.12) и (3.13) следует, что

С получается простой максимизацией $H(Y)$. Это достигается, когда $\text{Pr}(y_j = 1) = \text{Pr}(y_j = 0) = 1/2$, откуда $H(Y) = 1$, и, следовательно,

$$C_{\text{ДСК}} = 1 + p \log_2 p + q \log_2 q. \quad (3.18)$$

3.6. Некоторые коды с исправлением ошибок

Рассмотрим теперь построение кодов для передачи информации по такому каналу, как ДСК. В этом случае кодирующее и декодирующее устройства системы связи преобразуют одни последовательности двоичных цифр в другие последовательности двоичных цифр. Правильное воспроизведение всех 2^k последовательностей сообщений длины k на выходе декодирующего устройства требует передачи последовательностей сигналов длины n , где

$$\frac{k}{n} < C. \quad (3.19)$$

Будем называть k/n *скоростью передачи* и обозначать ее через R . Операция кодирования выполняется следующим образом. Входные цифры группируются в блоки по k цифр, которые последовательно преобразуются в равное количество блоков цифровых сигналов длины n . Блоки выбираются таким образом, чтобы в декодирующем устройстве давать возможность исправлять ошибки. Одним из способов является посылка блоков длины n , содержащих k цифр, соответствующих цифрам оригинала сообщения, и $n - k$ контрольных цифр, которые позволяют обнаруживать и исправлять ошибки, возникающие во время передачи. Существует много других путей получения таких кодов с исправлением ошибок (Питерсон [28]).

В качестве примера кодов с исправлением ошибок мы даем три кода Хэмминга [17], которые исправляют одиночные ошибки при скоростях передачи $1/3$, $2/5$ и $4/7$ соответственно. Эти коды приведены в табл. 3.2 и 3.3.

Мы интерпретируем эту таблицу следующим образом. Рассмотрим код (5,2). Пусть типичным кодовым словом будет $x_1 x_2 \dots x_5$. Здесь x_1 и x_2 уже определены, т. е. $x_1 x_2$ — это исходное сообщение. Остальные места в кодовом слове

Таблица 3.2
Кодирующая функция для кодов
Хэмминга (3,1), (5,2) и (7,4)

		k		
		1	2	4
n	3	21 31		
	5		312 42 51	
	7			5134 6124 7123

определяются согласно табл. 3.2 как

$$\left. \begin{aligned} x_3 &= x_1 \oplus x_2, \\ x_4 &= x_2, \\ x_5 &= x_1. \end{aligned} \right\} \quad (3.20)$$

Декодирующие функции для этих кодов задаются табл. 3.3.

Таким образом, если $y_1 y_2 \dots y_5$ — полученное кодовое слово из кода (5,2), то x_1 и x_2 вычисляются следующим образом:

$$\left. \begin{aligned} x_1 &= y_1 \oplus (y_1 \oplus y_5) \& (y_1 \oplus y_2 \oplus y_3), \\ x_2 &= y_2 \oplus (y_2 \oplus y_4) \& (y_1 \oplus y_2 \oplus y_3). \end{aligned} \right\} \quad (3.21)$$

Аналогично, в случае кода (3.1) получаем

$$x_1 = y_1 \& y_2 \oplus y_2 \& y_3 \oplus y_3 \& y_1 \quad (3.22)$$

и так далее. Эти коды работают за счет создания «резерва» множеств выходных последовательностей сигналов для каждой данной входной сигнальной последовательности.

Т а б л и ц а 3.3

Декодирующие функции для кодов Хемминга
(3,1), (5,2) и (7,4)

		n		
		3	5	7
k	1	11 & 22 & 33 & 1		
	2		11 15 & 123 22 24 & 123	
	4			11 1345 & 1246 & 1237 22 1345 & 1246 & 1237 33 1345 & 1246 & 1237 44 1345 & 1246 & 1237

Упомянутые множества состоят из последовательностей, «сгруппированных» вокруг выходной последовательности, которая соответствует, при отсутствии шума, данной входной последовательности, т. е., эти множества состоят из всех последовательностей, которые отличаются некоторыми неопределенными цифрами от данной входной последовательности. Тогда любая частная выходная последовательность декодируется согласно этой классификации. Например, код (3,1), правило образования которого показано в табл. 3.3, резервирует выходные последовательности 111, 110, 101, 011 для входной последовательности 111 и последовательности 000, 001, 010, 100 для входной последовательности 000. Таким образом, одиночные ошибки автоматически обнаруживаются и исправляются декодирующим устройством.

Вероятность отсутствия ошибки при декодировании Q дается для этих кодов формулой

$$Q = \sum_i \alpha_i p^i q^{n-i}, \quad (3.23)$$

где p и q те же, что в уравнении (3.16). Коэффициент α_i приведен в табл. 3.4.

Т а б л и ц а 3.4
Вероятность отсутствия ошибки
для кодов Хемминга
(3,1), (5,2) и (7,4)

		k				
		i	$\binom{n}{i}$	1	2	4
n	3	0	1	1		
		1	3	3		
	5	0	1		1	
		1	5		5	
		2	10		2	
	7	0	1			1
		1	7			7
		2	21			
		3	35			

Так, для кода (5,2) имеем

$$Q = q^5 + 5q^4p + 2q^3p^2 \quad (3.24)$$

и так далее.

На рис. 3.3 показана система связи, использующая такие коды.



Рис. 3.3. Система связи в действии.

На этом завершается наше изложение теории автоматов и теории информации и мы можем приступить к рассмотрению сущности этой работы, т. е. к синтезу конечных автоматов, надежно функционирующих, несмотря на сбои в составляющих их элементах.

ГЛАВА ЧЕТВЕРТАЯ

НАДЕЖНОСТЬ АВТОМАТОВ

4.1. Работы фон Неймана

Фон Нейман был, вероятно, первым, кто подробно рассмотрел проблему надежности автоматов [41, 42]. Питтс и Маккаллок [31], пытаясь применить теорию автоматов к задаче, связанной со слуховым и зрительным восприятием и функционированием коры головного мозга, отметили, что модульные сети, предназначенные для моделирования этих сторон деятельности сложных биологических систем, должны быть построены таким образом, чтобы их функции не нарушались в результате малых отклонений в возбуждениях и торможениях, в порогах модулей и в отдельных деталях модульных соединений. Винер [45] также рассмотрел один из аспектов этой проблемы — неправильное функционирование вычислительных машин, вызываемое сбоями или отказами составляющих их модулей. Он отметил, что в существующих вычислительных машинах использование мажоритарной (пороговой) логики (по принципу: «То, что я вам говорю, верно два раза из трех») вместе с поиском методов создания новых модулей привело к повышению надежности вычислений.

На подобных идеях основывались и работы фон Неймана. Его основное решение было получено следующим образом. Желательно получить надежный выход конечного автомата, описываемого некоторой заданной булевой функцией. Принимается, что такой автомат строится из модулей, которые вычисляют одну из двух функций — либо штрих Шеффера $s(x_1, x_2)$, либо мажоритарную функцию $m(x_1, x_2, x_3)$. Все модули вычисляют либо одну либо другую функцию. Сеть, состоящая из таких модулей, может вычислить произвольную булеву функцию, если при этом прибегнуть к некоторым ухищрениям. Сигналы поступают от одного модуля к другому по соединениям, передающим только двоичные импульсы.

Предполагается также, что модули перестают нормально функционировать с (точной) вероятностью ϵ и сбой в модулях статистически не зависят от общего состояния сети и от других сбоев. Делая, в частности, это допущение, фон

Нейман отмечал, что более реальным было бы допущение: «сбой статистически зависим от общего состояния сети и от наличия других сбоев. В любом данном (конкретном) состоянии, однако, сбой в рассматриваемом базовом модуле имеет вероятность равную ε ». Однако для простоты анализа он принял более простое допущение.

Для уменьшения воздействия сбоев в модулях сеть составляется из значительно большего числа модулей и соединений, чем теоретически необходимо для вычисления требуемой булевой функции, т. е. сеть проектируется *избыточной*. Так, конечный автомат (рис. 4.1), реализующий без

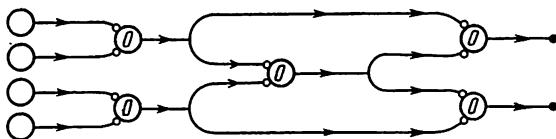


Рис. 4.1. Неизбыточная модульная сеть, состоящая из модулей, реализующих функцию штрих Шеффера.

избыточности данное определенное событие, используется в качестве прототипа для избыточного автомата, изображенного на рис. 4.2. Эта сеть имеет следующую структуру. Каждый модуль сети прототипа заменен блоком модулей того же самого типа, а каждое соединение заменено пучком. Структура блока и пучка не является полностью упорядоченной — допускаются некоторые случайности на микроуровне. Однако организация связей между блоками точно соответствует организации сети прототипа, т. е. на макроуровне случайностей не имеется. Каждый блок и пучок выполняют только то, что выполняли один модуль и одно соединение прототипа. Таким образом, множество модулей и соединений сети является избыточным. Именно эта структурная избыточность позволяет уменьшить воздействия сбоев в модулях, так как каждый блок работает по многократному сигналу, передаваемому пучком, и принимает решение только на основе большинства.

Каждый пучок состоит из n соединений, каждое из которых передает только двоичные импульсы. Таким образом,

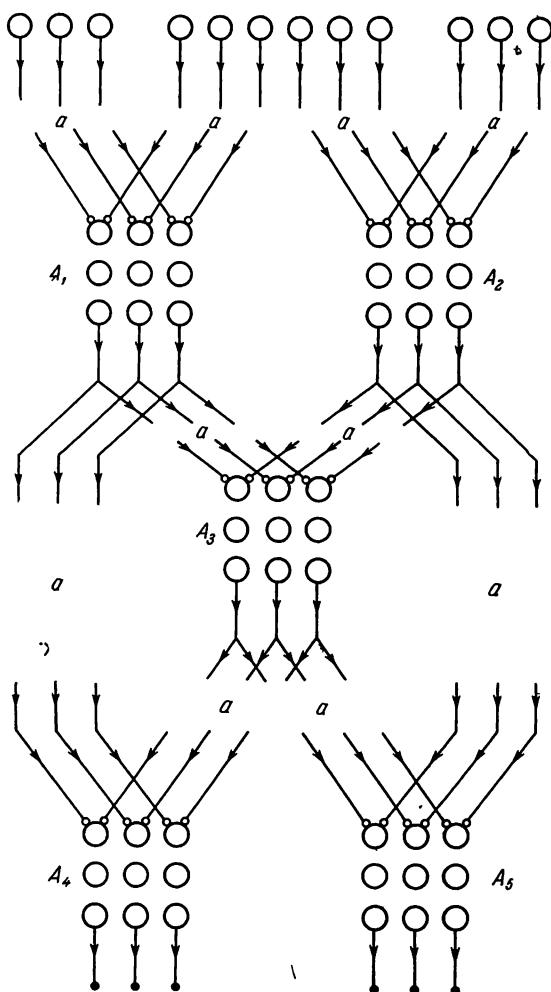
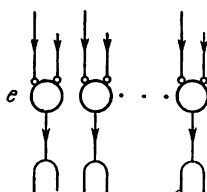


Рис. 4.2. Избыточный вариант сети, показанной на рис. 4.1. A_i — избыточные блоки; a — перестановки из соединений 1, 2, 3.

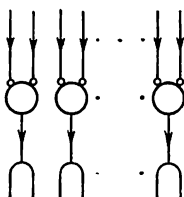
для пучка имеется 2^n различных картин возбуждения или сигнальных конфигураций от (111 ... 1) до (000 ... 0). Число единиц в конфигурации (уровень возбуждения) обозначается ξ , а доверительный уровень (порог) — Δ , так что

$$\left. \begin{array}{ll} 1. n \geq \xi \geq (1 - \Delta)n & \text{сигналы 1.} \\ 2. 0 \leq \xi \leq \Delta n & \text{сигналы 0.} \\ 3. \Delta n < \xi < (1 - \Delta)n & \text{сигналы} \end{array} \right\} (4.1)$$

сбоя.



a



a

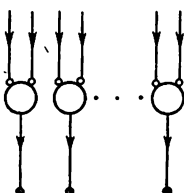


Рис. 4.3. Подробная схема избыточного блока: e — «исполнительный» орган; r — «восстанавливающий» орган.

Каждый избыточный блок работает, реагируя на свои входные пучки, следующим образом.

Его первый ряд модулей вычисляет n копий $s(x_1, x_2)$, т. е. этот ряд вычисляет $s(x_{1a_1}, x_{2b_1}), s(x_{1a_2}, x_{2b_2}), \dots, s(x_{1a_n}, x_{2b_n})$, где (a_1, a_2, \dots, a_n) и (b_1, b_2, \dots, b_n) — любые возможные перестановки из $(1, 2, \dots, n)$. Выходной пучок этого ряда затем расщепляется, перемешивается, и его сигнальная конфигурация обрабатывается вторым рядом модулей. Этот второй процесс затем повторяется, и результирующий пучок несет сигнальные конфигурации к другим блокам. Первый ряд называется, следуя фон Нейману, *исполнительным органом*, т. е. он «выполняет» данную функцию блока. Множество последующих рядов называется *восстанавливающим органом*, так как его задача, как мы это покажем, заключается в том, чтобы бороться с воздействиями сбоев в модулях исполнительного ряда и, таким образом, восстанавливать сигнальные конфигурации, искаженные шумом (рис. 4.3).

Это выполняется следующим образом: любая сигнальная конфигурация с уровнем возбуждения ξ , превосходящим некоторое начальное значение ξ_c , увеличивает этот уровень за счет восстанавливающего органа и, наоборот, любая сигнальная конфигурация с уровнем возбуждения

ξ , меньшим ξ_c , уменьшает этот уровень (рис. 4.4). В результате сбои, представляемые уровнями возбуждений, заключенными между Δn и $(1 - \Delta)n$, постепенно корректируются, так как уровни этих частных возбуждений постепенно заменяются уровнями, принимающими значения вне границ, указанных выше.

Сам блок, как отмечалось, избыточен, и эта избыточность может быть определена как

$$N = 3 \cdot n. \quad (4.2)$$

Для данной вероятности неправильного срабатывания ε фон Нейман показал (изменяя Δ и n), что существует оптимальное Δ , такое, что никакие сбои не передаются через избыточную сеть. Для вероятности P неправильной работы сети в целом при заданных модулях, которые функционируют неправильно с вероятностью $\varepsilon = 5 \cdot 10^{-3}$, он получил зависимость

$$P \cong \frac{6,4}{n} \cdot 10^{-8n/10\,000}. \quad (4.3)$$

Это выражение таково, что для получения разумно малых P необходимо иметь довольно большие значения n (см. табл. 4.1).

Таблица 4.1

Вероятность неправильного функционирования сети
в зависимости от избыточности

n	P	n	P
1000	$2,7 \cdot 10^{-2}$	10 000	$1,6 \cdot 10^{-10}$
2000	$2,6 \cdot 10^{-3}$	20 000	$2,8 \cdot 10^{-19}$
3000	$2,5 \cdot 10^{-4}$	25 000	$1,2 \cdot 10^{-23}$
5000	$4,0 \cdot 10^{-6}$		

Фон Нейман не был удовлетворен этим результатом и считал, что



Рис. 4.4. Передаточная функция восстанавливающего органа.

«ошибку следует трактовать термодинамическими методами, и что она должна быть объектом термодинамической теории, какой была информация в работах Л. Сциларда и К. Э. Шеннона».

4.2. Другие подходы к созданию надежных автоматов

Другие подходы к созданию надежных автоматов, включая использование более сложных модулей с иным характером возникновения ошибок по сравнению с ранее упомянутыми, также приводили к избыточным

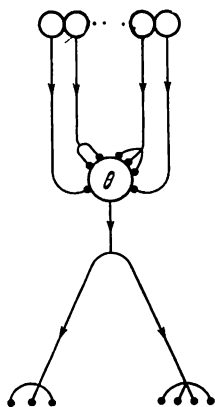


Рис. 4.5. Сложный модуль с синаптической репликацией.

автоматам в основном с теми же характеристиками, определяемыми уравнением (4.3). Так, Аллансон [1] рассмотрел синтез сетей из модулей, которые вычисляют мажоритарную функцию многих входов. Модуль такого типа изображен на рис. 4.5.

Предполагается, что ошибки происходят как в самих модулях, так и при синаптической передаче, т. е. при передаче сигналов между модулями. Предполагается также, что эти синаптические ошибки могут быть двух родов: ошибки, которые уничтожают входной сигнал, и ошибки, которые самопроизвольно порождают фиктивный входной сигнал. Мы можем описать возникновение ошибок в синапсе, используя терминологию теории информации и представив синапс как двоичный асимметричный канал, в котором p_1 и p_2 — соответственно «положительная» и «отрицательная» синаптические ошибки. Аллансон показал, что повышение надежности синаптической передачи может быть достигнуто за счет увеличения числа окончаний, контактирующих с каким-либо модулем, т. е. за счет синаптической репликации.

Распространение этих соображений на сети с «шумовыми» модулями и с «шумовыми» соединениями было впервые проведено Мурогой [27] и Вербееком, Блюмом, Коуэном и Маккаллоком [39, 3, 7, 22, 23]. Они использовали модули,

аналогичные модулям, изображенным на рис. 4.6, за исключением того, что с каждым модулем связывалась вероятность ошибки ε . Кроме этого, рассматривались синаптические ошибки, имеющие место с вероятностью p_s . Эти ошибки исправлялись в основном путем сочетания модульной итерации с синаптической репликацией. Например, функция $P(t) \equiv \bar{y}_1(t-1) \vee \bar{y}_2(t-1)$ надежно вычислялась модульной сетью, приведенной на рис. 4.6. Каждая переменная,

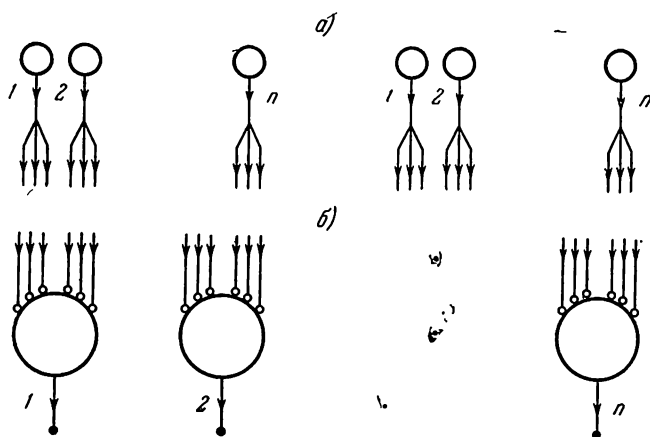


Рис. 4.6. а) Модульная сеть с исправлением ошибок;
б) схема соединений «все \pm со \pm всеми».

как и раньше, представлена пучком из n линий, и уровень возбуждения ξ , удовлетворяющий неравенству $\Delta n < \xi < (1 - \Delta)n$, свидетельствует о неправильном функционировании. Порог θ каждого из n модулей установлен таким, что входные ошибки исправляются и остаются только модульные ошибки. Для этого нужно выполнить условие

$$-2n(1 - \Delta) < \theta \leq -n(1 + \Delta), \quad (4.4)$$

которое приводит к уравнению

$$\theta = -(n + [n\Delta] + 0,5), \quad (4.5)$$

где $\Delta < 0,33$ и $[n\Delta]$ — наименьшее целое число, большее, чем $n\Delta$. Отметим, что каждая входная линия контактирует

со всеми n модулями, каждый из которых вычисляет пороговую функцию, заданную уравнением (4.5).

Верхняя граница вероятности неправильного функционирования сети из-за входных и синаптических ошибок может быть вычислена как функция биномиального распределения вероятностей. Объединяя ее с вероятностью неправильного функционирования сети из-за модульных ошибок, получаем верхнюю границу вероятности неправильного функционирования выходного пучка. Отсюда, в соответствии с биномиальным законом распределения, можно вычислить вероятность P неправильного функционирования сети. Это вычисление усложняется тем, что ошибки на входе и синаптические ошибки могут оказывать или не оказывать независимые воздействия на выход. Верхняя граница P может быть, однако, получена, если вычислить вероятности неправильного функционирования сети для обоих случаев и принять наибольшее из двух значений за P . При этом получаются следующие результаты.

Пусть ε_i — вероятности ошибок на входе (в примере, приведенном на рис. 4.6, это было бы эквивалентно утверждению, что входные модули функционируют неправильно с вероятностями ε_i). Тогда

$$\eta = \varepsilon_i(1 - p_s) + (1 - \varepsilon_i) p_s \quad (4.6)$$

есть вероятность неправильного функционирования из-за воздействия как входных, так и синаптических ошибок. Табл. 4.2 показывает изменение величины P в зависимости

Т а б л и ц а 4.2
Вероятность неправильного функционирования модуля $\max P$ в зависимости от избыточности

n	$\eta = \varepsilon = 0,005$	$\eta = \varepsilon = 0,01$
5	$6,7 \cdot 10^{-2}$	$2,4 \cdot 10^{-1}$
10	$3,1 \cdot 10^{-3}$	$3,8 \cdot 10^{-2}$
20	$1,0 \cdot 10^{-4}$	$7,2 \cdot 10^{-3}$
30	$3,6 \cdot 10^{-6}$	$1,5 \cdot 10^{-3}$
40	$\sim 10^{-7}$	$5,6 \cdot 10^{-5}$

от n , η и ε в случае, когда $\eta = \varepsilon$. Сравнение таблицы 4.2 с таблицей 4.1 показывает, в частности, возросшую эффективность исправления ошибок за счет использования более сложных модулей.

В самом деле, возросшая эффективность такова, что ошибки исправляются сетями единичной глубины, и отпадает необходимость иметь восстанавливающие органы. С другой стороны, можно рассмотреть более сложные модули, которые реализуют как вычисление, так и восстановление, все в одном месте (Коуэн [8]). Однако вероятность неправильного функционирования P стремится к нулю, как и в схеме фон Неймана, при n , стремящемся к бесконечности, с той лишь разницей, что для данного P это стремление к бесконечности более медленное. Таким образом, хотя и была достигнута возросшая эффективность по сравнению со схемой фон Неймана, все еще требуется произвольно высокая избыточность для достижения произвольно низкой вероятности неправильного функционирования.

4.3. Сравнение с теорией информации

Если мы сравним избыточные модульные сети, разработанные фон Нейманом и другими, с избыточной системой связи Шеннона [35], то сразу станут очевидными некоторые существенные различия. Сравнивая модульную избыточность с избыточностью сигнальных последовательностей, применяемой для борьбы с воздействиями «шумового» канала, и определяя *модульную избыточность* сети N как отношение числа модулей в избыточной сети к числу модулей в неизбыточной сети, имеем

$$\left. \begin{aligned} N &= 3n \text{ (фон Нейман),} \\ N &= n \text{ (Мурога и др.).} \end{aligned} \right\} \quad (4.7)$$

Аналогично, удельная нагрузка на модуль R определяется как $1/N$, так что

$$\left. \begin{aligned} R &= \frac{1}{3n} \text{ (фон Нейман),} \\ R &= \frac{1}{n} \text{ (Мурога и др.).} \end{aligned} \right\} \quad (4.8)$$

Из табл. 4.1 и табл. 4.2 видно, что вероятность неправильного функционирования P приблизительно пропорциональна $\exp(-c_1/R)$, т. е.

$$\left. \begin{aligned} P &\cong d_1 \exp\left(-\frac{c_1}{R}\right) \quad (\text{фон Нейман}), \\ P &\cong d_2 \exp\left(-\frac{c_2}{R}\right) \quad (\text{Мурога и др.}), \end{aligned} \right\} \quad (4.9)$$

где d_1 и d_2 — медленно меняющиеся функции n , а c_1 и c_2 — константы.

Если сравнить уравнение (4.9) с выражением вероятности ошибки опознавания P , задаваемым уравнением (3.15), т. е.

$$P \cong 2^{-n(c-R)}, \quad (4.10)$$

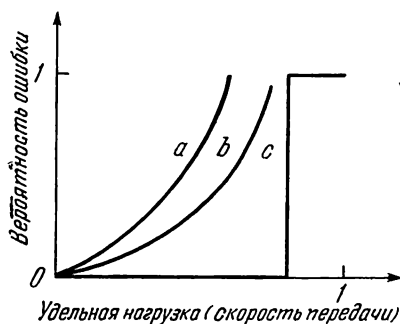


Рис. 4.7. Сравнение различных решений проблемы повышения надежности. а) Фон Нейман; б) Мурога и др.; в) Шеннон.

то можно заметить, что различие заключается в том, что P стремится к нулю с ростом n независимо от R , при условии: $R \leq C$, в то время как P стремится к нулю вместе с R . Это показано на рис. 4.7.

Важным следствием этого является то, что характеристики избыточных автоматов, сконструированных в соответствии с теорией фон Неймана, Муроги и др.,

не дают возможности определить такое понятие работоспособности, которое допускало бы непосредственную оценку.

4.4. Попытки применения теории информации

Отсутствие понятия работоспособности канала вычислений применительно к избыточным автоматам, рассмотренным в разделах 4.1 и 4.2, наводит на мысль, что более прямое применение теории кодирования к задаче вычислений может привести к получению возможности описания работы избыточных автоматов в терминах способности выпол-

нять операции с заданной вероятностью за заданное время, или, короче, в терминах работоспособности. В соответствии с этим Элайс [10] предложил модель вычислительной системы, состоящей из модульной сети, подверженной шумам, и идеальных кодирующих и декодирующих устройств. Структура этой системы, которую мы несколько изменили, не меняя ее основных черт, изображена на рис. 4.8. Требуется, чтобы модульная сеть реализовала событие E , соответствующее некоторой логической функции входов x_1 и x_2 . Элементами этой сети являются модули, которые вычисляют булеву функцию одной или двух переменных, например, \neg , $\&$, \vee , \equiv , \oplus и т. д., с малой, но ненулевой вероятностью ошибки. Эти вероятности статистически независимы от общего состояния сети и от наличия других сбоев; т. е. делается такое же допущение относительно ошибки, как у фон Неймана. Однако модульная сеть избыточна и избыточность системы не является схемной или канальной избыточностью — это просто избыточность сигнальной последовательности, как в случае связи. Последовательности длины k входных сигналов x_1 и x_2 кодируются в сигнальные последовательности длины n . Эти последовательности являются входными для сети. Выходные последовательности длины n декодируются в последовательности длины k , которые, надо надеяться, соответствуют реализуемому событию. Важно отметить, однако, что сеть не имеет памяти, и вся обработка ведется поразрядно, так что в пределах последовательности операции не имеют места.

Предполагается, что кодирование по каждому входу x_1 и x_2 независимо, и декодирующее устройство однозначно отображает сигнальные последовательности в выходные при отсутствии шумов в модульной сети. Эти допущения обеспечивают то, что заданное множество событий целиком реализуется в модульной сети с шумом и не реализуется (ни

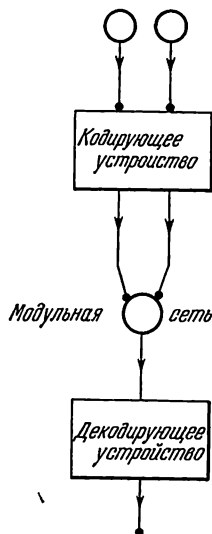


Рис. 4.8. Вычислительная система.

частично, ни полностью) в кодирующих и декодирующих устройствах без шума.

Получены следующие результаты: из шестнадцати возможных булевых функций двух переменных x_1 и x_2 имеется только восемь функций, для которых можно использовать (n, k) коды, дающие положительные скорости передачи информации через сеть при бесконечно малых частотах ошибки. Из этих восьми функций шесть не представляют большого интереса, так как являются функциями 1 (тавтология), 0 (противоречие), x_1 , x_2 , \bar{x}_1 и \bar{x}_2 . Оставшиеся две функции $x_1 \oplus x_2$ и $x_1 \equiv x_2$ таковы, что каждая отображает множество $2^k n$ -разрядных последовательностей в себя таким образом, что сохраняются метрические свойства этого множества, т. е. последовательности, отличающиеся одна от другой большим количеством разрядов на входе, отображаются в такие же последовательности на выходе. Это значит, что при кодировании входных последовательностей для этих функций можно использовать коды Хэмминга [17] или любые другие группы кодов (Слепян [36]). Однако $x_1 \oplus x_2$ и $x_1 \equiv x_2$ не являются универсальными функциями, и поэтому все множество восьми функций неполно, т. е. сети, реализующие даже определенные события, нельзя построить только из элементов этого множества.

Элайс показал, что для другого множества восьми функций, включающего $x_1 \& x_2$, $x_1 \& \bar{x}_2$, $\bar{x}_1 \& x_2$, $\bar{x}_1 \& \bar{x}_2$, $x_1 \vee x_2$, $x_1 \vee \bar{x}_2$, $\bar{x}_1 \vee x_2$ и $\bar{x}_1 \vee \bar{x}_2$, наилучшим является код, при котором поразрядные кодирующие устройства повторяют каждую цифру n раз и принимают большинство выходов в качестве правильного выхода. Это значит, что для этих восьми функций произвольно высокая надежность вычисления может быть достигнута только за счет произвольно низких скоростей вычисления. Анализ Элайса был повторен Питерсоном и Рэбином [29], которые получили, по существу, аналогичный результат.

4.5. Обсуждение

Мы можем разделить рассмотренные соображения на ряд различных классов. Первый класс состоит из соображений фон Неймана, в которых изучаются простые, по существу, модули, функционирующие неправильно с вероят-

ностью ϵ . Избыточные блоки и последовательности состоят, по существу, из множества копий простых модулей или же цифр сообщений, т. е. в процессе кодирования не происходит никакого увеличения модульной сложности (мы используем в качестве меры модульной сложности количество входов модуля). Результирующие сети, несомненно, обнаруживают и исправляют ошибки, но делают это, вообще говоря, неэффективно. Аллансон, Муруга, Вербеек и др. [1, 27, 39, 3, 7, 23] предложили более эффективные схемы, в которых вместо простых модулей использовались блоки из сложных модулей.

Ни в одном из этих классов не обеспечивается надежная обработка информации с положительной скоростью в том смысле, как мы определили выше. Однако очевидно, что повышенная эффективность при кодировании избыточных автоматов может быть получена за счет увеличения сложности модулей, составляющих такие автоматы, и это приводит к выводу, что только автоматы, состоящие из избыточных блоков сложных модулей, могут обеспечить ненулевую удельную нагрузку при надежных вычислениях.

ГЛАВА ПЯТАЯ

МОДУЛЬНЫЕ ВЫЧИСЛИТЕЛЬНЫЕ СЕТИ

В предыдущих главах мы не определили явно, что подразумевается под вычислением, а просто считали, что модульные сети, реализующие определенные события, являются (неявно) примерами вычислительных сетей. Ясно, однако, что мы должны связывать вычисление с актом выбора и необходимого разрушения информации в этом процессе. Именно это разрушение и используется для определения понятия вычисления*).

Рассмотрим, например, модуль, реализующий любую булеву функцию двух переменных x_1 и x_2 . Из шестнадцати различных возможных булевых функций двух переменных

*) Другими словами, авторы определяют вычислительные процессы как процессы переработки информации, при которых количество переработанной информации существенно меньше количества исходной информации. (Прим.ред.)

некоторые просто идентичны x_1 и x_2 или их отрицаниям \bar{x}_1 и \bar{x}_2 , а другие являются такими функциями, как $x_1 \& x_2$, $x_1 \vee x_2$, $\bar{x}_1 \& \bar{x}_2$, $\bar{x}_1 \vee \bar{x}_2$ и т. д. Определяя понятие вычисления, мы должны разграничивать эти два множества, так как элементы первого таковы, что функция определяет аргумент, в то время как элементы последнего таковы, что функция неоднозначно определяет аргумент. Мы сделаем это ниже для общего случая модуля со входами x_1, x_2, \dots, x_s и выходом y .

5.1. Системы обработки информации

Следуя модели системы связи, рассмотренной в главе 3 (см. рис. 3.1), будем считать, что такой модуль является частью вычислительной системы или системы для обработки информации, как это показано на рис. 5.1. Предположим, что в момент t на каждом входе независимо выбирается один из k_1 возможных символов ансамбля $X_r (r=1, 2, \dots, s)$ и что каждый выходной сигнал выбран в момент

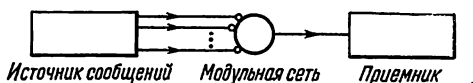


Рис. 5.1. Обобщенная система обработки информации.

$t + 1$ из ансамбля Y (k_2 символов). Тогда общее количество информации, связанной со входами x_1, x_2, \dots, x_s , равно

$$H(X) = \sum_{r=1}^s H(X_r) \quad (5.1)$$

и количество информации, связанной с выходом, равно $H(Y)$. Зависимость между этими величинами определяется структурой модульной функции. Таким образом, каждая конфигурация входов $x_\alpha(t) = x_1(t)x_2(t)\dots x_s(t)$ определяет единственный выход $y_\beta(t + 1)$, и поэтому поведение системы может быть представлено множеством переходных вероятностей

$$\text{Pr}(y_\beta(t + 1) | x_\alpha(t)) \quad (\alpha = 1, \dots, k_1^s; \beta = 1, \dots, k_2)$$

таких, что

$$\Pr(y_\beta(t+1)|x_\alpha(t)) = \begin{cases} 1, & y_\beta(t+1) = f(x_\alpha(t)), \\ 0, & y_\beta(t+1) \neq f(x_\alpha(t)), \end{cases} \quad (5.2)$$

где f — заданная функция, реализуемая модулем.

Средняя неопределенность относительно y_β , обеспечиваемая x_α , может быть определена как

$$\begin{aligned} H(Y|X) &= \\ &= - \sum_{\alpha=1}^{k_1^S} \sum_{\beta=1}^{k_2} \Pr(x_\alpha(t), y_\beta(t+1)) \log_2 \Pr(y_\beta(t+1)|x_\alpha(t)). \end{aligned} \quad (5.3)$$

Теперь можно доказать следующую лемму:

Л е м м а 5.1.

$$H(Y|X) \equiv 0.$$

Д о к а з а т е л ь с т в о .

$$\begin{aligned} H(Y|X) &= - \sum_{\alpha=1}^{k_1^S} \sum_{\beta=1}^{k_2} \Pr(x_\alpha(t), y_\beta(t+1)) \log_2 \Pr(y_\beta(t+1)|x_\alpha(t)) = \\ &= - \sum_{\alpha=1}^{k_1^S} \sum_{\beta=1}^{k_2} \Pr(y_\beta(t+1)|x_\alpha(t)) \times \\ &\quad \times \Pr(x_\alpha(t)) \log_2 \Pr(y_\beta(t+1)|x_\alpha(t)) = \\ &= 1 \cdot \Pr(x_\alpha(t)) \cdot \log_2 1 + 0 \cdot \Pr(x_\alpha(t)) \cdot \log_2 0 \equiv 0, \end{aligned}$$

что и требовалось доказать.

По аналогии с соотношениями (3.11) и (3.12), можно утверждать, что

$$H(X) - H(X|Y) = H(Y) \quad (5.4)$$

или что

$$H(X) - H(Y) = H(X|Y). \quad (5.5)$$

То есть количество информации, обеспечиваемое входными символами x_α , минус количество информации,

обеспечиваемое выходными символами y_β модульной сети без шума, равно неопределенности относительно входов, даваемой выходами.

Будем считать, что имеет место вычислительный процесс, если

$$H(X|Y) > 0, \quad (5.6)$$

т. е., если выходные символы не полностью определяют входные конфигурации; будем считать, что имеет место процесс передачи информации (связь), если

$$H(X|Y) = 0, \quad (5.7)$$

т. е. выходные символы полностью определяют входные конфигурации. Непосредственно из уравнения (5.5) следует, что вычисление имеет место, если $H(X) > H(Y)$, т. е. если при переходе от X к Y информация теряется *).

5.2. Вычисление с помощью ненадежных модулей **)

До сих пор мы пренебрегали воздействием шума в системе и принимали, что входы модуля однозначно определяют его выход. Рассмотрим теперь ненадежный модуль (или модульную сеть), выбирающий, как и прежде, входные символы в момент t из ансамбля X . Однако выходные символы выбираются в момент $t + 1$ не из ансамбля Y выходных символов, рассмотренного выше в случае с идеальным модулем, а из некоторого ансамбля Z . Поведение этой системы представляется множеством переходных вероятностей

$$\text{Pr}\{z_\gamma(t+1)|x_\alpha(t)\} \quad (\gamma = 1, \dots, k_2).$$

Предположим, что эти вероятности удовлетворяют условию

$$\text{Pr}\{z_\gamma(t+1)|x_\alpha(t)\} = \text{Pr}\{z_\gamma(t+1)|x'_\alpha(t)\}$$

*) Здесь следует подчеркнуть, что причиной подобного уменьшения количества информации является сама вычислительная операция, а не ненадежность работы элементов. (Прим. ред.)

**) Здесь и далее под ненадежным модулем (модульной сетью) понимается модуль (модульная сеть), подверженный воздействию шума. (Прим. ред.)

для всех $z_\gamma \in Z$ всякий раз, когда

$$f(x_\alpha(t)) = f(x'_\alpha(t)). \quad (5.8)$$

Это значит, что $\Pr(z_\gamma(t+1) | x_\alpha(t))$ можно представить как

$$\begin{aligned} \Pr(z_\gamma(t+1) | x_\alpha(t)) &= \\ &= \sum_{\beta=1}^{k_\beta} \Pr(z_\gamma(t+1) | y_\beta(t+1)) \Pr(y_\beta(t+1) | x_\alpha(t)), \end{aligned} \quad (5.9)$$

т. е. модульная сеть с шумом, определяемая множеством переходных вероятностей $\Pr(z_\gamma(t+1) | x_\alpha(t))$, может быть

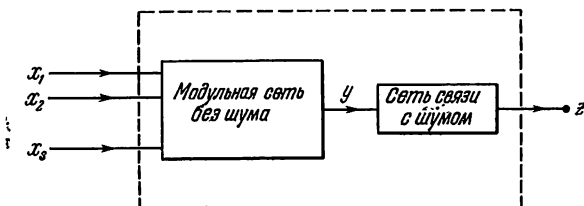


Рис. 5.2. Разбиение модуля с шумом.

представлена как сочленение двух сетей, которые соответственно описываются множествами переходных вероятностей $\Pr(y_\beta(t+1) | x_\alpha(t))$ и $\Pr(z_\gamma(t+1) | y_\beta(t+1))$. Первое из этих множеств представляет сеть, реализующую функцию $f(x_\alpha(t))$, в то время как второе представляет сеть связи с шумом и с нулевой задержкой (рис. 5.2).

Уравнение (5.9) позволяет доказать две следующие леммы.

Л е м м а 5.2.

$$H(Z | X) = H(Z | Y).$$

Д о к а з а т е л ь с т в о.

Опустим аргумент t , так как он не фигурирует в этом доказательстве. Из уравнения (5.9)

$$\Pr(z_\gamma | x_\alpha) = \Pr(z_\gamma | y_\beta) \Pr(y_\beta | x_\alpha) = \Pr(z_\gamma | y_\beta) \quad (y_\beta = f(x_\alpha))$$

и имеется только одно y_β такое, что $y_\beta = f(x_\alpha)$. Отсюда

$$\begin{aligned} H(Z|X) &= - \sum_{\alpha=1}^{k_1^s} \sum_{\gamma=1}^{k_2} \Pr(x_\alpha, z_\gamma) \log_2 \Pr(z_\gamma | x_\alpha) = \\ &= - \sum_{\alpha=1}^{k_1^s} \sum_{\gamma=1}^{k_2} \Pr(z_\gamma | x_\alpha) \Pr(x_\alpha) \log_2 \Pr(z_\gamma | x_\alpha) = \\ &= - \sum_{\gamma=1}^{k_2} \sum_{\alpha \in \gamma_{\alpha\beta}} \sum_{\beta=1}^{k_2} \Pr(z_\gamma | y_\beta) \Pr(x_\alpha) \log_2 \Pr(z_\gamma | y_\beta), \end{aligned}$$

где $\gamma_{\alpha\beta}$ есть множество всех x_α таких, что $y_\beta = f(x_\alpha)$ для данного β .

Отсюда следует

$$\begin{aligned} H(Z|X) &= - \sum_{\alpha \in \gamma_{\alpha\beta}} \Pr(x_\alpha) \sum_{\beta=1}^{k_2} \sum_{\gamma=1}^{k_2} \Pr(z_\gamma | y_\beta) \log_2 \Pr(z_\gamma | y_\beta) = \\ &= - \sum_{\beta=1}^{k_2} \sum_{\gamma=1}^{k_2} \Pr(z_\gamma | y_\beta) \Pr(y_\beta) \log_2 \Pr(z_\gamma | y_\beta) = H(Z|Y), \end{aligned}$$

что и требовалось доказать.

Л е м м а 5.3.

$$H(X|Z) = H(X|Y) + H(Y|Z).$$

Д о к а з а т е л ь с т в о.

Из уравнений (3.11) и (3.12) имеем

$$H(X) - H(X|Z) = H(Z) - H(Z|X), \quad (5.10)$$

$$H(Z) - H(Z|Y) = H(Y) - H(Y|Z). \quad (5.11)$$

Из леммы (5.2) и уравнения (5.10) имеем

$$\begin{aligned} H(X|Z) &= H(X) - H(Z) + H(Z|Y) = \\ &= H(X) - H(Y) + H(Y|Z) = H(X|Y) + H(Y|Z), \end{aligned}$$

что и требовалось доказать.

Леммы 5.1 и 5.3 позволяют нам установить очень важную зависимость между средним количеством информации, обеспечиваемым z_γ относительно x_α , т. е. $I[X; Z]$, и средним

количеством информации, обеспечиваемым z_γ относительно y_β , т. е. $I[Y; Z]$.

Т е о р е м а 5.1.

$$I[X; Z] = I[Y; Z].$$

Д о к а з а т е л ь с т в о.

$$\begin{aligned} I[X; Z] &= H(Z) - H(Z | X) \quad \text{по (3.12),} \\ &= H(Z) - H(Z | Y) \quad \text{по лемме (5.2),} \\ &= I[Y; Z] \quad \text{по (3.12),} \end{aligned}$$

что и требовалось доказать.

То есть среднее количество информации, обеспечиваемое z_γ относительно x_α , равно среднему количеству информации, обеспечиваемому z_γ относительно y_β . Это означает, что из выходных сигналов z_γ ненадежной модульной сети, вычисляющей функцию $y_\beta = f(x_\alpha)$, нельзя, в среднем, извлечь больше информации относительно x_α , чем из выходных сигналов гипотетической идеальной сети, вычисляющей ту же функцию.

Это следует из того факта, что, в общем, при переходе от X к Z информация теряется не только из-за воздействия шума, но и в результате выполнения вычисления. Положение станет очевидным, если записать $I[X; Z]$ в форме

$$I[X; Z] = H(X) - H(X | Y) - H[Y | Z]. \quad (5.12)$$

Это выражение показывает, что количество информации, обеспечиваемое z_γ относительно x_α , равно количеству информации, обеспечиваемому x_α , минус количество информации, потерянное при выполнении вычислений, минус количество информации, потерянное из-за воздействия шума.

5.3. Модульное разбиение

Уравнение (5.10), указывающее место потерь информации в процессе вычисления при наличии шумов, строго зависит от принятого разбиения множества $\Pr(z_\gamma(t+1)|x_\alpha(t))$.

Такое непосредственное разбиение вообще невозможно, тем более, что множество $\Pr(y_\beta(t+1)|x_\alpha(t))$ состоит только из единиц и нулей. В последующем обсуждении мы будем рассматривать только модули, поддающиеся

разбиению. В тех случаях, когда разбиение невозможно, мы будем рассматривать такие модули, как если бы они были еще более ненадежными, и заменять множество вероятностей $\Pr(z_\gamma(t+1)|x_\alpha(t))$, описывающее их поведение, другим

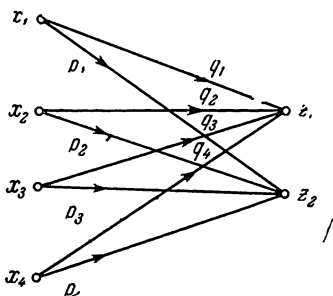


Рис. 5.3. Символическое представление вычислительного канала с шумом.

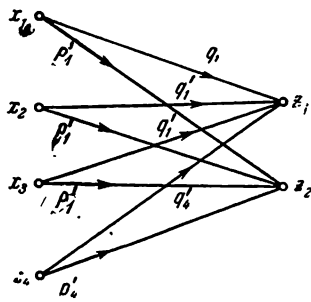


Рис. 5.4. Вероятностная функция $g(x_\alpha(t))$.

множеством $\Pr^*(z_\gamma(t+1)|x_\alpha(t))$, допускающим разбиение и выбранным так, чтобы воздействие шума не уменьшалось. Рассмотрим следующий пример подобной процедуры. Пусть имеется ненадежный модуль, который должен вычислять заданную функцию $x_1(t)$ & $x_2(t)$, где x_1 и x_2 — символы, выбранные из двоичного алфавита. Здесь можно воспользоваться шенноновским символическим изображением канала с шумом (Шеннон [35]). Другими словами, символическое изображение, показанное на рис. 5.3, представляет собой модуль, который вычисляет вероятностную функцию*) $g(x_\alpha(t))$ с переходными вероятностями

$$\Pr(z_1(t+1)|x_\alpha(t)) = q_\alpha, \quad \Pr(z_2(t+1)|x_\alpha(t)) = p_\alpha;$$

$$q_\alpha + p_\alpha = 1.$$

В рассматриваемом примере q_1 , q_2 и q_3 различны, но больше или равны 0,5, а q_4 меньше или равно 0,5.

*) В дальнейшем авторы приняли следующую систему обозначений. Через x_α обозначается одно из четырех возможных состояний входа (одна из четырех возможных комбинаций символов x_1 и x_2). z_1 и z_2 — два возможных значения выходного сигнала. Для рассматриваемой булевой функции $x_1 = \{0,0\}$, $x_2 = \{0,1\}$, $x_3 = \{1,0\}$, $x_4 = \{1,1\}$, $z_1 = 0$ и $z_2 = 1$. (Прим. ред.)

Такое множество переходных вероятностей, в общем, не допускает разбиения. Однако если q_1 , q_2 и q_3 заменить на

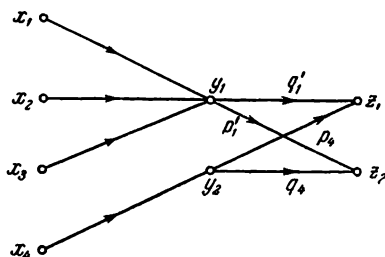


Рис. 5.5. Разбиение $g(x_\alpha(t))$ на последовательное соединение «безошибочной» функции $f(x_\alpha(t)) = x_1(t) \& x_2(t)$ и канала с шумом, выполняющего однозначное преобразование.

q'_1 , где $q'_1 = \min(q_1, q_2, q_3)$, то результирующая схема является менее надежной и допускает разбиение. На рис. 5.4 и 5.5 показана такая схема и ее разбиение.

5.4. Общий метод

В общем случае пусть ненадежный модуль, который должен вычислить функцию $f(x_\alpha(t))$, будет представлен множеством переходных вероятностей

$$\Pr(z_\beta(t+1) | x_\alpha(t)).$$

Заменим множество переходных вероятностей другим множеством $\Pr^*(z_\beta(t+1) | x_\alpha(t))$ таким, что

$$\Pr^*(z_\beta(t+1) | x_\alpha(t)) = \begin{cases} \max_{\gamma_{\alpha\delta}} \Pr(z_\beta(t+1) | x_\alpha(t)), & \text{если } y_\beta(t+1) \neq f(x_\alpha(t)), \\ 1 - \sum_{\gamma_{\delta\alpha}} \Pr^*(y_\delta(t+1) | x_\alpha(t)), & \text{если } y_\beta(t+1) = f(x_\alpha(t)), \end{cases} \quad (5.13)$$

где $\gamma_{\alpha\delta}$ — множество всех x_δ таких, что $f(x_\delta(t)) = f(x_\alpha(t))$ и $\gamma_{\delta\alpha}$ — множество всех y_δ таких, что $y_\delta \neq f(x_\alpha(t))$. Этот метод получения разложимого модуля ограничивает воздействие шума, однако он может дать слишком грубую оценку

или даже вообще быть непригодным в том случае, когда

$$\sum_{\gamma \delta \alpha} \text{Pr}^*(y_\delta(t+1) | x_\alpha(t)) \geq 1.$$

В некоторых случаях можно использовать другой метод, который сводит к минимуму указанные трудности. Снова рассмотрим канал, изображенный на рис. 5.3. Предположим, что q_1, q_2, q_3 расположены по порядку

$$q_1 \leq q_2 \leq q_3. \quad (5.14)$$

Допустим, что вычислительный канал с шумом, определенный этим множеством (и q_4), используется разумно в том смысле, что менее надежные переходы используются реже, чем более надежные («согласование»), так что

$$\text{Pr}(x_1) \leq \text{Pr}(x_2) \leq \text{Pr}(x_3). \quad (5.15)$$

При этих условиях можно легко показать, что решение

$$q'_1 = 1/3(q_1 + q_2 + q_3), \quad q'_2 = q_4 \quad (5.16)$$

удовлетворяет следующему неравенству:

$$I[X; Z'] \leq I[X; Z], \quad (5.17)$$

где Z' — ансамбль, состоящий из алфавита Z и соответствующих вероятностей $\text{Pr}^*(z_\gamma)$, равных

$$\text{Pr}^*(z_\gamma) = \sum_{\alpha=1}^{k_1^s} \text{Pr}^*(z_\gamma(t+1) | x_\alpha(t)) \text{Pr}(x_\alpha(t)). \quad (5.18)$$

То есть выбор *среднего арифметического* q_1, q_2 и q_3 для q'_1 ни недооценивает, ни переоценивает шум до такой степени, как это было бы при выборе $q'_1 = q_1$, и, более того, q_1 не зависит от входных вероятностей. Однако общие решения уравнения (5.18) требуют дальнейших исследований.

ГЛАВА ШЕСТАЯ

РАБОТОСПОСОБНОСТЬ ВЫЧИСЛИТЕЛЬНОГО КАНАЛА

В предыдущей главе для характеристики вычислительного канала и канала связи были использованы меры информации $I[X; Z]$ и $I[Y; Z]$ соответственно. В этом разделе мы

попытаемся оправдать использование этой меры, приведя теорему о кодировании, подобную теореме 3.2 из раздела 3.4.

Эта теорема содержит в себе понятие пропускной способности канала, определяемой как максимальная скорость передачи символов y_β , отнесенная к единице количества передаваемой информации по каналу связи с шумом, т. е.

$$C = \max_Y I[Y; Z], \quad (6.1)$$

где максимизация достигается за счет надлежащего выбора множества вероятностей $\text{Pr}(y_\beta)$ при ограничении $\sum_{\beta=1}^{k_z} \text{Pr}(y_\beta) = 1$. Напоминаем, что C определяет максимальное количество информации относительно входов, которое может быть извлечено из выходных сигналов канала связи с шумом или, равносильно, из ненадежной модульной сети связи. Мы распространим теперь это понятие пропускной способности канала на ненадежную модульную сеть, вычисляющую заданные функции.

6.1. Основное неравенство

Определим величину C^* , называемую *работоспособностью* такой сети, как

$$C^* = \max_X I[X; Z], \quad (6.2)$$

где z_γ — выходы ненадежной модульной сети, вычисляющей данную функцию $f(x_\alpha(t))$. Мы предположили, что такую сеть можно представить как последовательное соединение идеальной вычислительной сети и сети связи с шумом. Пусть C — пропускная способность последней. Тогда имеем следующую теорему.

Т е о р е м а 6.1.

$$C^* \leq C.$$

Д о к а з а т е л ь с т в о.

$$C^* = \max_X I[X; Z] = \max_X I[Y; Z] \leq \max_Y I[Y; Z] = C,$$

(по теореме 5.1)

что и требовалось доказать.

Напомним, что $\max_X I[X; Z]$ является сокращенным обозначением для $\max_X I[X_1, X_2, \dots, X_s; Z]$. Предполагается, что модульные входы X_1, X_2, \dots, X_s выбираются каждый независимо, так что

$$\Pr(x_\alpha) = \prod_{r=1}^s \Pr(x_{ra}). \quad (6.3)$$

Пусть $\Pr^*(y_\beta)$ — вероятности, которые максимизируют $I[Y; Z]$. Тогда необходимым и достаточным условием того, что C^* равно C , является возможность нахождения множества вероятностей, удовлетворяющих уравнению (6.3), такого, что

$$\Pr^*(y_\beta) = \sum_{\alpha} \Pr(x_\alpha). \quad (6.4)$$

6.2. Некоторые примеры вычислительных каналов

Приведем два примера вычислительных каналов, одного с $C^* = C$ и другого с $C^* < C$.

Пример а).

Рассмотрим вычислительный канал, показанный на рис. 5.4 и 5.5. Пусть $\Pr(y_1 = 1) = r$ и $\Pr(y_2 = 0) = 1 - r$ вероятности, которые максимизируют $I[Y; Z]$. Множество вероятностей $\Pr(x_1 = 1) = r$, $\Pr(x_1 = 0) = 1 - r$, $\Pr(x_2 = 1) = 1$ и $\Pr(x_2 = 0) = 0$ удовлетворяет уравнению (6.3) и, таким образом, C^* равно C .

Пример б).

Рассмотрим вычислительный канал с шумом, показанный на рис. 6.1.

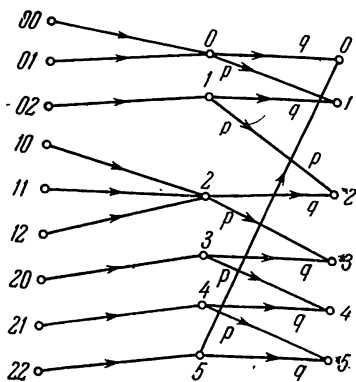


Рис. 6.1. Вычислительный канал с шумом при разбиении с $C^* < C$.

Этот канал получен для модуля с двумя входами, каждый из которых может находиться в одном из трех различных

состояний (0, 1, 2), и одного выхода, который может находиться в одном из шести различных состояний (0, 1, 2, 3, 4, 5). Значение C для канала связи, показанного на рис. 6.1, равно (Шеннон [35])

$$C = \log_2 6 + p \log_2 p + q \log_2 q, \quad (6.5)$$

где распределение вероятностей Y дается формулой

$$\Pr(y = \beta) = 1/6 \quad (\beta = 0, 1, \dots, 5). \quad (6.6)$$

Таким образом, для получения в этом случае работоспособности канала вычислений C^* , равной C , требуется такое распределение вероятностей X , которое вызывает распределение вероятностей Y , заданное уравнением (6.6).

Пусть $\Pr(x_\alpha = j) = \Pr_{\alpha j}$ ($\alpha = 1, 2$; $j = 0, 1, 2$). При этом $C^* = C$ тогда и только тогда, когда следующая система уравнений:

$$\left. \begin{aligned} \Pr_{10} \cdot \Pr_{20} + \Pr_{10} \cdot \Pr_{21} &= 1/6 & (a) \\ \Pr_{10} \cdot \Pr_{22} &= 1/6 & (b) \\ \Pr_{11} \cdot (\Pr_{20} + \Pr_{21} + \Pr_{22}) &= 1/6 & (c) \\ \Pr_{12} \cdot \Pr_{20} &= 1/6 & (d) \\ \Pr_{12} \cdot \Pr_{21} &= 1/6 & (e) \\ \Pr_{12} \cdot \Pr_{22} &= 1/6 & (f) \end{aligned} \right\} \quad (6.7)$$

имеет решение, удовлетворяющее условиям

$$\Pr_{\alpha j} \geq 0, \quad \sum_{j=0}^2 \Pr_{\alpha j} = 1. \quad (6.8)$$

Исключая \Pr_{10} из уравнений (6.7a) и (6.7b) и используя условие, что $\Pr_{20} + \Pr_{21} + \Pr_{22} = 1$, получаем решение $\Pr_{22} = 1/2$. Подставляя это значение \Pr_{22} в уравнение (6.7 f), получаем решение $\Pr_{12} = 1/3$. Последующее использование этой величины \Pr_{12} в уравнениях (6.7d) и (6.7e) приводит к уравнению $\Pr_{20} + \Pr_{21} + \Pr_{22} = 1/2$, которое противоречит уравнению (6.8). Следовательно, для данного примера C^* всегда меньше, чем C .

6.3. Вычисление булевых функций

В оставшейся части этой монографии мы будем рассматривать только булевы функции. Следующая теорема является поэтому, несомненно, важной:

Т е о р е м а 6.2. Пусть t — разложимый модуль, предназначенный для вычисления булевой функции $f(x_1, \dots, x_s)$. Тогда во всех случаях $C^* = C$.

Д о к а з а т е л ь с т в о.

Так как f — булева функция, то она может принимать только одно из двух значений, т. е. $y = 1$ или 0 . Пусть $\text{Pr}(y = 1) = p_1$, $\text{Pr}(y = 0) = q_1 = 1 - p_1$ будет распределением вероятностей y , которое реализует C . Так как $f(x_1, \dots, x_s)$ — булева функция (в общем случае не постоянная), то существует одна переменная x_r и $s - 1$ значений для других $s - 1$ переменных из x_1, \dots, x_s , которые таковы, что

$$\left. \begin{aligned} f(x_1, x_2, \dots, x_s) \Big|_{\substack{x_r=r \\ x_i=x_i^* (i \neq r)}} &= 0, \\ f(x_1, x_2, \dots, x_s) \Big|_{\substack{x_r=1-r \\ x_i=x_i^* (i \neq r)}} &= 1. \end{aligned} \right\} \quad (6.9)$$

Распределение вероятностей x_α , задаваемое

$$\left. \begin{aligned} \text{Pr}(x_i = x_i^*) &= 1 \quad (i \neq r), \\ \text{Pr}(x_r = r) &= \text{Pr}(y = 0) \end{aligned} \right\} \quad (6.10)$$

вызывает требуемое распределение по y , откуда $C^* = C$, что и требовалось доказать.

6.4. Предельная теорема для вычислительного канала с шумом

Теорема 6.1, которую можно сформулировать следующим образом:

$$\max_x I[X; Z] \leq \max_Y I[Y; Z], \quad (6.11)$$

указывает, что максимальное количество информации относительно входов x_α , которое может быть извлечено на

выходе z_γ ненадежной модульной сети, вычисляющей функцию $f(x_\alpha(t))$, равно, самое большее, максимальному количеству информации, которое может быть извлечено из z_γ относительно выходов y_α гипотетической идеальной модульной сети, вычисляющей данную функцию. Это вполне согласуется с тем, что уже было сказано выше относительно вычисления.

Из глав 3 и 4 ясно, что такое извлечение информации требует использования кодов с исправлением ошибок и, следовательно, избыточности либо сигнальной последовательности, либо канала, либо схемы. Теорема 3.2 указала связь C с этой избыточностью, и при этом обеспечила методы ее использования. В последующем обсуждении мы покажем, как для C^* можно дать подобные методы.

В качестве первого шага докажем «теорему случайного кодирования» (подобную теореме 3.2), которая связывает избыточность сигнальной последовательности с C^* . Для этого мы используем модель, изображенную на рис. 6.2 (см. также рис. 5.1), состоящую из модуля (или модульной сети), обслуживаемого идеальными кодирующими и декодирующими устройствами. Предполагается, что сеть не имеет памяти и что каждый вход x_1, x_2, \dots, x_s кодируется независимо (рис. 6.2). Тогда имеем следующую теорему.

Теорема 6.3. *Рассмотрим дискретный источник S , выбирающий сообщения из ансамбля X , содержащего количе-*

ство информации $H(X) = \sum_{r=1}^s H(X_r)$, и ненадежную модуль-

ную вычислительную сеть с работоспособностью C^ . Если существует источник S' , выбирающий последовательности*

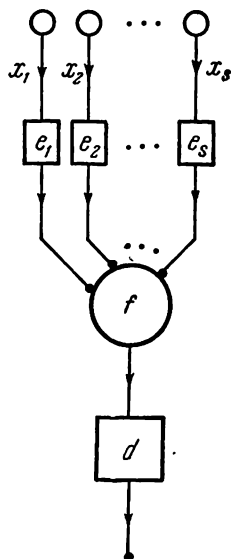


Рис. 6.2. Модульная вычислительная система: x_1, x_2, \dots, x_s — входные переменные; e_1, e_2, \dots, e_s — кодирующие устройства; f — логическая функция; d — декодирующее устройство.

из ансамбля X' , такой, что

$$(1) H(X_r) < H(X'_r) \quad (r = 1, 2, \dots, s)$$

и

$$(2) H(X') = \sum_{r=1}^s H(X'_r) = C^* + H(X'|Z'),$$

где Z' — ансамбль выходных последовательностей, то существует по крайней мере один код с тем свойством, что надлежащим образом закодированные сообщения из S могут быть обработаны сетью и восстановлены с произвольно малой частотой ошибки.

Доказательство этой теоремы, приведенное в приложении, строго следует доказательству теоремы Шеннона (глава 3).

Теорема показывает, что ненадежная модульная вычислительная сеть, снабженная соответствующими кодирующими и декодирующими устройствами, может быть использована таким образом, что декодированные выходные последовательности однозначно определяют входные последовательности, т. е. так, что $H(Z) = H(X)$. Из уравнений (5.5) и (5.7) следует, что сеть используется только для связи, а не для вычислений. Положив $H(X) \leq C^*$, мы тотчас же исключаем любое использование сети для вычисления, так как это означает, что $H(X) < H(Y_0)$ (см. обозначения в приложении). Выбор такого $H(X)$, для которого $H(X) = H(X_0 | Y_0)$ меньше или равно C^* , позволил бы, тем не менее, выполнить вычисления в этой сети. Само по себе это не значит, что вычисление, выполняемое этой сетью, соответствует заданному. Таким образом, сеть, представленную на рис. 6.2, можно использовать в качестве системы связи, в которой последовательности сообщений длины k кодируются в сигнальные последовательности длины n таким образом, что k/n численно равно $C^* + H(X_0 | Y_0)$. Согласно теореме 3.2 после передачи информации останется неопределенность не менее, чем $H(X_0 | Y_0)$.

Чтобы обеспечить в сети действительное выполнение заданных вычислений, эта неопределенность должна точно соответствовать информации, обязательно разрушаемой в сети при заданном вычислении. То есть уравнение (5.5)

должно быть удовлетворено почленно в соответствии с обозначением $y_\beta = f(x_\alpha(t))$, а именно,

$$\begin{aligned} -\Pr(y_\beta) \log_2 \Pr(y_\beta) + \sum_{\substack{\gamma_{\alpha\beta} \\ k_1^s}} \Pr(x_\alpha) \log_2 \Pr(x_\alpha) = \\ = \Pr(y_\beta) \sum_{\alpha=1} \Pr(x_\alpha | y_\beta) \log_2 \Pr(x_\alpha | y_\beta) \quad (6.12) \end{aligned}$$

или

$$-H(Y_\beta) + H(\chi_{\alpha\beta}) = \Pr(y_\beta) H(X | Y_\beta) \quad (\beta = 1, \dots, k_2), \quad (6.13)$$

где $\chi_{\alpha\beta}$ — ансамбль, соответствующий $\gamma_{\alpha\beta}$.

Это условие в некотором смысле требует, чтобы в кодирующее устройство были заложены сведения о функции, которая вычисляется сетью, и вероятно, что если не приняты специальные меры предосторожности, то часть требуемых вычислений будет фактически выполняться в кодирующем устройстве. Если к тому же кодирующее устройство работает, как это предполагалось, без сбоев, то работоспособность модульной сети, обслуживаемой таким кодирующим устройством, не имеет смысла. Можно принять меры, чтобы обеспечить полное отсутствие операции по реализации функции f в кодирующих сетях за счет независимого кодирования каждой переменной. Однако из теоремы 6.3 кажется вероятным, что любое вычисление не может быть выполнено и в идеальном модуле; все вычислительные операции переносятся в декодирующее устройство. С другой стороны, если и декодирующее устройство работает без сбоев, то эта процедура также неудовлетворительна по тем же соображениям. Мы можем добиться того, что никаких вычислений не будет в декодирующем устройстве, так же как и в кодирующем устройстве, при условии, что декодирующее устройство осуществляет однозначное преобразование в отсутствие шума. Тогда последует результат Элайса [10] и Питерсона и Рэбина [29], а именно, при вычислении в модульных сетях поразрядно событий, представляемых либо булевыми функциями двух переменных, либо функциями, составленными из них, в общем, можно использовать только $(n, 1)$ -коды, и поэтому $C^* = 0$. Эти результаты недавно были распространены одним из авторов (Виноград [46]) на

сети, которые поразрядно вычисляют булевы функции любого конечного числа переменных. Было показано, что только $2^n + 1$ из 2^{2^n} возможных булевых функций двоичных переменных допускают положительные значения работоспособности канала вычислений при наличии ранее отмеченных ограничений, и что эти функции образуют особый класс, куда не входят функции, необходимые для реализации остальных $2^{2^n} - 2^n - 1$ функций.

Итак, мы видим, что теорема 6.3 привела нас в конце концов к результатам, которые были получены Элайсом и др., а именно, если мы строго следуем шенноновской модели системы связи и просто пытаемся получить надежные вычисления при положительных скоростях передачи информации такими схемами, как мы описали, то это приводит к отрицательным ответам. Далее мы покажем, как можно получить положительные ответы на вопросы, касающиеся работоспособности канала вычислений, задаваясь новыми путями введения избыточности и системами, менее строго следующими традиционной схеме кодирующего устройства, канала и декодирующего устройства.

ГЛАВА СЕДЬМАЯ

СИГНАЛЬНАЯ И МОДУЛЬНАЯ ИЗБЫТОЧНОСТИ

Сущность ограничений, возникающих при попытке рассмотрения вычислительной системы как строгой модели системы связи, данной Шенноном ([35], см. рис. 3.1), заслуживает дальнейшего изучения. Главные черты этой модели системы связи следующие:

1. Система состоит из трех функционально различных частей: кодирующего устройства, канала связи и декодирующего устройства.

2. Предполагается, что в кодирующем и декодирующем устройствах нет шума и шум воздействует только на канал связи.

3. Коды, применяемые для получения надежной связи при помощи такой системы, используют избыточность сигнальной последовательности, т. е. сообщения длины k кодируются в сигнальные последовательности длины n , пе-

редаются по каналу с шумом и затем декодируются в последовательности сообщений.

4. Канал связывает обрабатывает сигнальные последовательности поразрядно (по крайней мере, в тех простых случаях, которые мы рассмотрели), так что никакие операции не выполняются внутри последовательности.

Эта совокупность ограничений приводит к теореме 3.2, которая гласит, что для получения надежной связи можно применять коды, использующие избыточность сигнальной последовательности (которая измеряется как n/k) таким образом, что структура канала связи не меняется в результате работы кода; то есть канал остается фиксированным и только принимает кодированные входные последовательности и поставляет кодированные выходные последовательности для декодирования. Более того, с этой фиксированной структурой связана определенная величина, а именно, пропускная способность канала C , которая устанавливает максимальную среднюю скорость передачи информации без потерь по такому каналу. Пропускная способность не зависит от длины n сигнальной последовательности, а зависит только от статистической структуры шума в канале. Теорема 3.2 указывает на существование по меньшей мере одного кода с тем свойством, что если он обеспечивает избыточность сигнальной последовательности больше некоторой минимальной величины, определяемой пропускной способностью канала, то количество информации, равное C бит, может быть передано по каналу за единицу времени и восстановлено декодирующим устройством с произвольно малой частотой ошибки.

Как было показано, слишком тесная привязка к этой модели приводит в случае вычислительных систем к теореме 6.3 и к результатам Элайса и др. (цит. работы), которые показывают, что эквивалентная работоспособность канала вычислений C^* в общем случае равна нулю, т. е. не существует минимальной избыточности сигнальной последовательности, обеспечивающей надежное проведение вычислительного процесса, и что надежность вычислений пропорциональна длине используемой сигнальной последовательности. В главе 4 мы описали исследования фон Неймана и др., в которых были получены, по существу, такие же результаты, когда использовалась модульная избыточность

вместо избыточности сигнальной последовательности. В следующем разделе мы покажем возможность создания вычислительных систем (несколько отличающихся от вышеупомянутых), которые обеспечивают надежность вычислений и положительные значения работоспособности канала вычислений.

7.1. Функциональное кодирование

Модель системы связи, если ей строго следовать, хотя и приводит к нулевой работоспособности канала вычислений, дает нам ценное понимание природы кодирования для вычислительных каналов с шумом. Должно быть ясно, что процесс кодирования, замена последовательностей сообщений последовательностями сигналов является, по существу, процессом, посредством которого последовательности сообщений «подбираются соответственно» каналу с шумом. То есть эти последовательности заменяются другими, выбранными так, чтобы канал с шумом не вносил неопределенности больше некоторого неизбежного количества. Это осуществляется заменой «простых» последовательностей сообщений, состоящих из k независимых символов (полученных, если это нужно, в результате предварительного кодирования), «сложными» сигнальными последовательностями, состоящими из n взаимосвязанных символов. Таким образом, идеальное вычисление достигается путем проведения операции непосредственно над входными по отношению к каналу с шумом последовательностями с целью получения сложных последовательностей взаимосвязанных символов, которые эффективно «защищают» сообщения от воздействий канала с шумом.

Если возвратиться теперь к модели вычислительной системы (глава 5), состоящей из последовательного соединения вычислительного канала без шума и канала связи с шумом (см. рис. 5.2), то из предшествующих рассуждений следует, что вместо x_1, x_2, \dots, x_s требуемый выход y должен быть согласован с каналом связи с шумом. Факт, что мера информации в этой модели $I[X; Z]$ была сведена к $I[Y; Z]$ (см. главу 5), также указывает на это, т. е. *вместо аргументов $x_\alpha(t)$ кодироваться должны функциональные значения y_β .*

Предположим при этом, что выходные последовательности длины k , т. е. $y_{\beta k}$, заменены последовательностями длины n , $y_{\beta n}$, таким образом, что

$$y'_{\beta\mu} = e_{\mu}(y_{\beta\lambda}) \quad (\lambda = 1, \dots, k; \mu = 1, \dots, n), \quad (7.1)$$

где e_{μ} — кодирующие функции, предписанные некоторым данным кодом связи (см. раздел 3.6). Из теоремы 3.2 сразу же следует, что если k/n меньше или равно C^* , где C^* — работоспособность канала вычислений данной ненадежной сети (или совпадает с C — пропускной способностью эквивалентного канала связи с шумом), то $y_{\beta k}$ могут быть восстановлены с произвольно малой частотой ошибки по окончательным, претерпевшим воздействие шума выходным последовательностям $z_{\gamma n}$. То есть можно достигнуть положительных значений работоспособности канала вычислений, если использовать функциональное кодирование (уравнение (7.1)). Для пояснения этого метода приведем следующий пример.

Рассмотрим код (5,2), упомянутый в главе 3. Его использование показано в табл. 7.1. На рис. 7.1 показан автомат, в котором использован такой код. Соответствующий неизбыточный автомат показан на рис. 4.1. Каждая модульная сеть вычисляет некоторую функцию f'_{λ} (см. главу 8). Каждая декодирующая сеть вычисляет функцию $d_{\lambda}(z_{\gamma\mu}) = y_{\beta\lambda}$, т. е. восстанавливает требуемый выход.

Очевидно, что функционально закодированная система заметно отличается от ранее рассмотренных. В частности:

1. Объединены функции вычисления и кодирования.

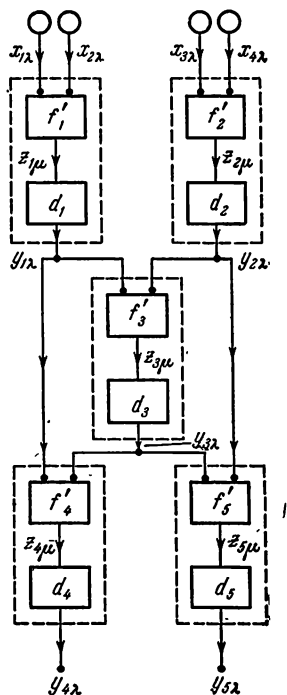


Рис. 7.1. Модульная сеть с избыточностью сигнальной последовательности.

2. Некоторые операции производятся над целыми сигнальными последовательностями.

Эти особенности означают, что модульные сети, реализующие события E , должны иметь какую-то память,

Таблица 7.1

Некоторые кодирующие и декодирующие функции для уравнения (7.1)

Кодирующая функция e_μ	Декодирующая функция d_λ
$y'_{\beta 1} = y_{\beta 1}$	$y_{\beta 1} = z_{\gamma 1} \oplus (z_{\gamma 1} \oplus z_{\gamma 2})$
$y'_{\beta 2} = y_{\beta 1}$	$\&(z_{\gamma 1} \oplus z_{\gamma 3} \oplus z_{\gamma 5})$
$y'_{\beta 3} = y_{\beta 1} \oplus y_{\beta 2}$	$y_{\beta 2} = z_{\gamma 5} \oplus (z_{\gamma 5} \oplus z_{\gamma 1})$
$y'_{\beta 4} = y_{\beta 2}$	$\&(z_{\gamma 1} \oplus z_{\gamma 3} \oplus z_{\gamma 5})$
$y'_{\beta 5} = y_{\beta 2}$	

включая элементы задержки, регистры сдвига и тому подобное (см. Фано [12] и Питерсон [28]). Таким образом, «простые» модули заменены «сложными» модульными вычислительными сетями, которые в одной части выполняют кодирование и вычисление, а в другой — декодирование. От декодирующего устройства не требуется выполнения каких-либо вычислений, но требуется однозначное преобразование информации при отсутствии шума.

С другой стороны, получающиеся в результате функционально закодированные сети чрезвычайно сложны по сравнению с первоначальными сетями или прототипами и, по мере того, как возрастают k и n (что и следует ожидать, если необходимо получить произвольно малую частоту ошибки в вычислениях при некотором заданном ненулевом отношении k/n), их сложность быстро возрастает. Желательно, чтобы сложность таких закодированных сетей оставалась возможно меньшей. Поэтому мы рассмотрим теперь другие средства функционального кодирования, вводя модульную избыточность вместо избыточности сигнальной последовательности.

7.2. Функциональное кодирование с использованием модульной избыточности

В предыдущей схеме мы закодировали требуемые выходы $y_{\text{вл}}$, т. е. мы отдельно закодировали требуемые выходы вместо кодирования общих требуемых выходных конфигураций. Это было сделано для того, чтобы иметь возможность использовать избыточность сигнальной последовательности. Теперь для кодирования $y_{\text{в}}$ воспользуемся той избыточностью модулей и соединений, которую ранее использовали фон Нейман и др. (см. главу 4), т. е. заменим множество требуемых выходов y_k другим множеством y'_n , так что

$$y'_\mu = e_\mu(y_\lambda) \quad (\lambda = 1, \dots, k; \mu = 1, \dots, n). \quad (7.2)$$

Уравнение (7.2) можно интерпретировать как уравнение, определяющее замену сети, вычисляющей функции f_1, f_2, \dots, f_k , другой сетью, вычисляющей новые функции f'_1, f'_2, \dots, f'_n . Следуя разделу 4.3, определим модульную избыточность N новой сети как отношение числа модулей в этой сети к числу модулей прототипа, т. е. n/k . Это аналогично избыточности сигнальной последовательности.

Подобным образом, по аналогии с ранее рассмотренной скоростью передачи информации на символ k/n , определим удельную нагрузку на модуль — R . Так,

$$N = \frac{n}{k}, \quad R = \frac{k}{n}. \quad (7.3)$$

Рис. 7.2 наглядно иллюстрирует это преобразование (здесь e_μ — кодирующие функции для простого итеративного $(n, 1)$ кода).

В более сложном случае, когда e_μ задается общими (n, k) кодами, k требуемых функций (или исходных функций) заменяются n кодированными функциями. Для иллюстрации этого на рис. 7.3 приводится пример, в котором использован (5,2) код Хэмминга.

Рис. 7.3, а и б показывают исходную и избыточную модульные сети, реализующие определенные события E_1 и

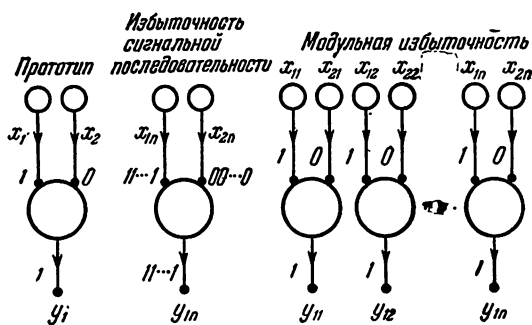


Рис. 7.2. Различные типы избыточности.

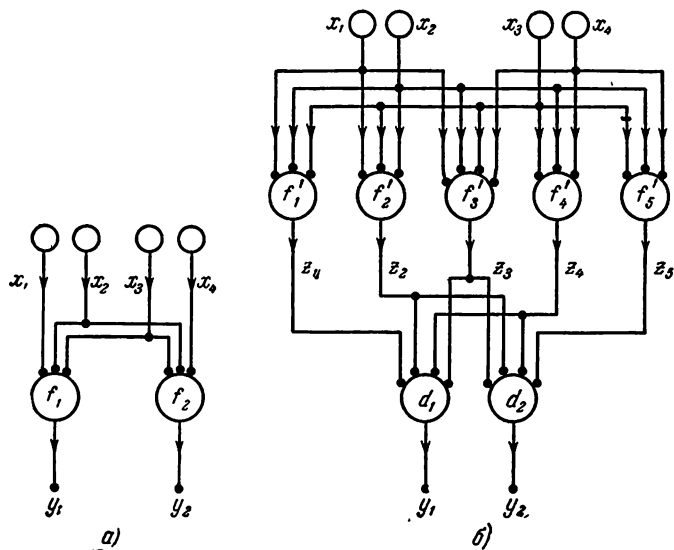


Рис. 7.3. Модульная сеть, иллюстрирующая использование модульной избыточности.

E_2 . Предполагается, что имеются модули, которые могут реализовать любую из требуемых функций f_λ и f_μ с вероятностью ошибки ε . Предполагается также, что в модулях, вычисляющих функции d_1 и d_2 , нет шума. Отметим, что модули, реализующие эти декодирующие функции, не считаются избыточными. Это не очень удовлетворительно по ряду причин. Однако в главе 8 мы приведем другую схему, которая устраняет эти недостатки.

В качестве дополнительного примера на

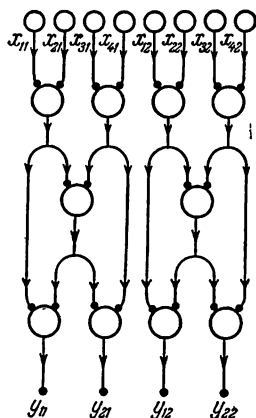


Рис. 7.4. Неизбыточная модульная сеть глубины 3, являющаяся исходной для введения модульной избыточности.

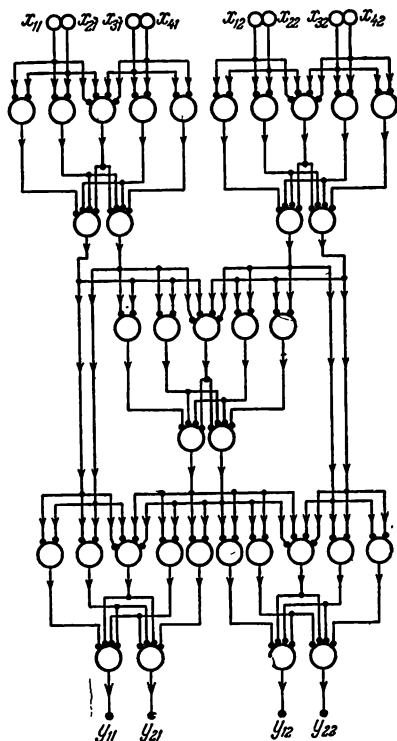


Рис. 7.5. Результирующая избыточная модульная сеть, в которой модульная избыточность вводится с использованием (5,2) кода Хэмминга.

рис. 7.4 и 7.5 показаны исходная и избыточная модульные сети, использующие код (5,2) и реализующие событие,

аналогичное событию на рис. 4.1. Читатель должен сравнить эту схему со схемой фон Неймана (см. рис. 4.2).

Нечто подобное было предложено независимо от нашей работы Армстронгом [2] и Рой-Чаудхури [32], которые модульную сеть, реализующую событие E , заменили другой сетью, реализующей событие E' . Однако они не использовали возможность увеличения исходного множества так, чтобы k/n оставалось постоянным. Работа Идена [9] также напоминает нашу потому, что он предложил схему, которая заменила одиночное событие более сложным событием, позволяя при положительных скоростях передачи использовать модульную избыточность и избыточность сигнальной последовательности для целей обнаружения одиночной ошибки. Эта схема примечательна тем, что часть вычисления происходит в декодирующем устройстве.

Из рис. 7.4 и 7.5 ясно, что хотя трудности, возросшие с увеличением сложности f'_μ , несколько уменьшились за счет использования модульной избыточности, не было найдено удовлетворительного ответа для решения проблемы сложности декодирующих устройств. Иными словами, даже если в декодерах d_λ никаких вычислений не происходит, они по крайней мере так же сложны, как и сети, реализующие f'_μ .

Таким образом, предположение о том, что в этих декодирующих сетях по сравнению с модульными сетями нет шума, неудачно. В самом деле, из примеров, приведенных в этом разделе, должно быть ясно, что любая попытка анализировать модульные сети независимо от кодирующих и декодирующих сетей, неудовлетворительна уже потому, что практически нельзя провести границу между такими сетями, как это сделано в системах связи. Это значит, что различие, проводимое между кодирующими, вычислительными и декодирующими сетями, довольно искусственное, и что было бы лучше не различать сети на такой основе.

В следующей главе мы приведем методы реализации (при условии некоторых остающихся ограничений) положительных значений работоспособности канала вычислений в избыточных модульных сетях, где вычисление и кодирование не разграничиваются, а фактически полностью объединяются.

ГЛАВА ВОСЬМАЯ

АНАСТОМОТИЧЕСКИЕ МОДУЛЬНЫЕ СЕТИ

Как мы показали ранее, расчленение модульной сети на кодирующую, вычислительную и декодирующую сети (или некоторую комбинацию из них) связано с рядом трудных проблем. В частности, понятие работоспособности канала вычислений становится несколько двусмысленным. Единственный способ исключить эту двусмысленность заключается, очевидно, в том, чтобы определить работоспособность для всех блоков сети, какие бы функции они ни выполняли. Кроме того, допущение, что в кодирующих и декодирующих сетях нет шума, не обосновано, так как такие сети, в сущности, не отличаются по своему характеру от сетей, которые реализуют определенные события. Далее мы будем рассматривать сети, предполагая, что шум имеется в каждом элементе сети независимо от его функционального назначения. Примем допущение, что любое неправильное функционирование отдельного модуля в любой исходной сети будет давать ошибку в функции, вычисляемой сетью, т. е. выход сети будет отличаться от значения заданной функции. Это, несомненно, переоценка воздействия модульных сбоев, но мы делаем такое допущение для того, чтобы обеспечить упрощение наших расчетов. Мы говорим, что сеть *произвольно надежна*, если вероятность неправильного функционирования сети произвольно близка к вероятности неправильного функционирования ее выходных элементов. Иными словами, так как ошибки могут иметь место где угодно внутри сети, некоторые из них могут возникать в выходных элементах сложной сети, и они не могут быть исправлены внутри сети. Эти ошибки можно исправить внешними способами, но тогда возникает проблема: «Кто исправляет корректора?» — «Sed quis custodiet ipsos custodies?» (фон Нейман [40]). Мы будем удовлетворены синтезом сетей, которые исправляют внутренние ошибки и дают выход, искажаемый только шумом в выходных модулях.

8.1. Расширение ансамбля. Основная теорема

В разделе 7.2 отмечалось, что при проектировании сетей для получения произвольно малых частот ошибки и положительных значений работоспособности канала вычисле-

ний можно использовать либо избыточность сигнальной последовательности, либо модульную избыточность. Это было возможно при условии, что имелись декодирующие устройства без шума и что удельную нагрузку k/n можно было зафиксировать, одновременно увеличивая k и n . Пока мы не введем идеальные модули любого типа, основным средством достижения цели останется расширение ансамблей как избыточных прототипов, так и избыточных сетей. То есть для сохранения фиксированной удельной нагрузки все большие и большие ансамбли прототипов должны заменяться все большими и большими избыточными ансамблями. Эти избыточные ансамбли должны быть составлены так, чтобы контролировать сбои, и должны состоять, как мы покажем, из более сложных модулей, чем модули прототипов.

Теперь сформулируем и докажем основную теорему, относящуюся к синтезу надежных автоматов.

Т е о р е м а 8.1. *Даны модули, которые могут вычислять с одинаковой работоспособностью S^* любую булеву функцию. Для любого автомата A , вычисляющего определенное событие, и любой удельной нагрузки на модуль $R < S^*$ (ср. главу 3) можно синтезировать произвольно надежный автомат A , который вычисляет определенное событие с удельной нагрузкой на модуль $\geq R$.*

Д о к а з а т е л ь с т в о. Пусть M — число модулей, составляющих автомат A , и пусть a_v ($v = 1, \dots, M$) — число воздействий выхода v -го модуля на отдельный выход автомата A . (Заметим, что мы не делаем допущения, что автомат A ациклический, а лишь предполагаем, что событие, реализуемое A , является определенным.) Пусть ε — вероятность любого отдельного модульного сбоя и P_A — вероятность неправильного функционирования A . Тогда

$$1 - P_A \geq \prod_{v=1}^M (1 - \varepsilon)^{a_v} = (1 - \varepsilon)^{\sum_{v=1}^M a_v} = (1 - \varepsilon)^b, \quad (8.1)$$

где

$$b = \sum_{v=1}^M a_v. \quad (8.2)$$

(Заметим, что если автомат A ациклический, то $a_v = 1$ и $b = M$.)

Таким образом, чтобы синтезировать автомат A с вероятностью неправильного функционирования, не превышающей $\delta (\geq 0)$, нам требуются модули с вероятностью неправильного функционирования, удовлетворяющей уравнению

$$\varepsilon \leq 1 - (1 - \delta)^{b^{-1}}. \quad (8.3)$$

Это значит, что если требуется произвольно высокая надежность, то ε должно быть произвольно мало, т. е. только при наличии идеальных модулей можно построить произвольно надежный автомат A . Для того чтобы получить произвольно надежные автоматы, состоящие из модулей с заданными вероятностями неправильного функционирования, мы рассмотрим сначала вместо автомата A другой автомат, $A(=A^k)$, который состоит из k копий автомата A , каждая из которых вычисляет одно и то же определенное событие, но не обязательно имеет те же самые входы. Таким образом, автомат A можно рассматривать как блок из k идентичных автоматов, работающих одновременно, но не обязательно по одной и той же программе или при одних и тех же исходных данных.

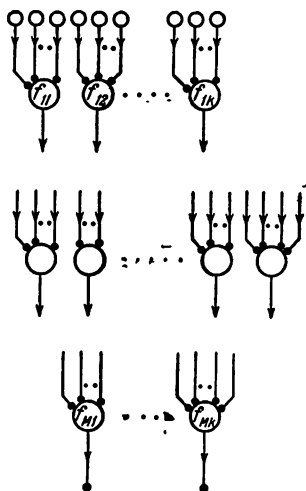


Рис. 8.1. Избыточный автомат A^k .

Обозначим булевы функции, вычисляемые в A , через $f_{\lambda\lambda}$. Рис. 8.1 дает представление о произвольном автомате A .

Очевидно, что если ε означает вероятность того, что по крайней мере один из k модулей $m_{\lambda\lambda}$ функционирует неправильно, то (если уравнение (8.3) удовлетворено) вероятность неправильного функционирования любого из k автоматов будет меньше δ . В том случае, когда ε фиксировано, мы рассматриваем вместо автомата A другой автомат A , полученный из A следующим образом:

1. Пусть x_r ($r = 1, \dots, s_\lambda$) входы модуля $m_{\lambda\lambda}$ исходного автомата A . Обозначим k входов $x_{r\lambda}$ ($\lambda = 1, \dots, k$) через

Уравнение (8.5) и рис. 8.2 показывают, что каждый модуль A вычисляет функцию $f'_{v\mu}$, которая реализует операции декодирования закодированных входов, вычисления функции $f_{v\lambda}$ и кодирования результирующего выхода для передачи к следующим модулям. В общем, любой ряд из

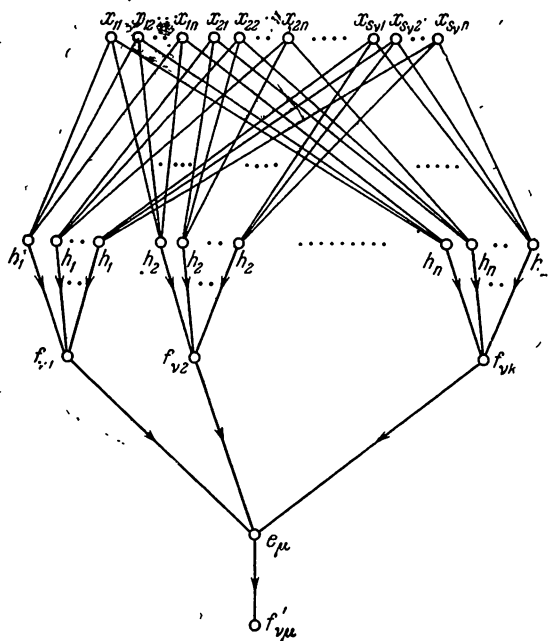


Рис. 8.2. Граф, представляющий структуру типичного модуля, вычисляющего функцию $f'_{v\mu}$.

n модулей A декодирует выходы предыдущего ряда и восстанавливает требуемое сообщение с некоторой вероятностью ошибки P , меньшей или равной δ . Эта вероятность определяется только модульным шумом (e) и (n, k) кодами, которые определяют операции декодирования и кодирования d_λ и e_μ соответственно. Другими словами, считается, что шум имеет место при передаче сообщений от ряда к ряду, и, следовательно, по теореме Шеннона (теорема 3.2) для достаточно большого k существует по крайней мере

один (n, k) код такой, что если k/n — удельная нагрузка на модуль и C^* — работоспособность модулей, вычисляющих $f'_{\nu\mu}$ (функцию e), то

$$P \cong 2^{k-nC^*}. \quad (8.6)$$

Так как между модулями в A и модульными рядами в A существует однозначное соответствие, из этого следует,

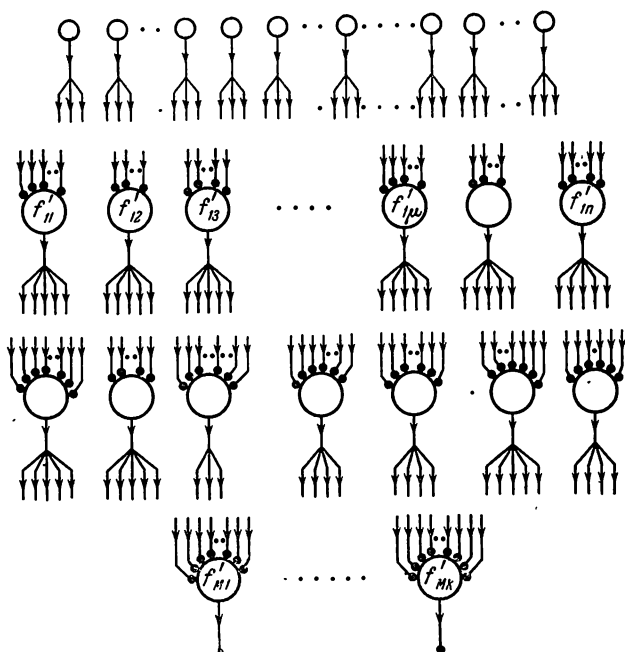


Рис. 8.3. Избыточно закодированный автомат A .

что P_A — вероятность неправильного функционирования автомата A , задается выражением

$$1 - P_A \geq (1 - P)^b$$

или

$$P \leq 1 - (1 - P_A)^{b^{-1}}. \quad (8.7)$$

Из уравнений (8.6) и (8.7) имеем

$$2^{k-nC^*} \leq 1 - (1 - P_A)^{b^{-1}} \quad (8.8)$$

Таким образом, для любого конечного b и любого P_A уравнение (8.8) может быть удовлетворено за счет увеличения k и n и фиксирования k/n при некотором значении, меньшем C^* . В частности, таким путем может быть получена произвольно высокая надежность ($P_A \rightarrow 0$).

Количество модулей в неизбыточном прототипе A равно $M \cdot k$. Аналогично, количество модулей в A равно $(M - u)n + uk$, где u — число выходных модулей в A . Модульная избыточность A , следовательно, равна $[(M - u)n + uk]/Mk$, а удельная нагрузка на модуль R' равна

$$R' = \frac{Mk}{(M - u)n + uk} > \frac{k}{n} = R. \quad (8.9)$$

Таким образом, для удельной нагрузки $R' \cong R$ при условии, что $R \leq C^*$, можно синтезировать автомат A , который вычисляет с произвольно высокой надежностью (если заданы модули с работоспособностью C^*) определенное событие, вычисляемое прототипом A , что и требовалось доказать.

8.2. Некоторые примеры надежных модульных сетей

Для пояснения смысла этой теоремы дадим несколько примеров надежных модульных сетей, работающих при различных удельных нагрузках и основанных на кодах, приведенных в разделе 3.6. Для наглядности приведем примеры, в которых автомат A — ациклический.

Рассмотрим автомат A , показанный на рис. 8.4. В общем случае f_i может быть произвольной, но мы ограничимся

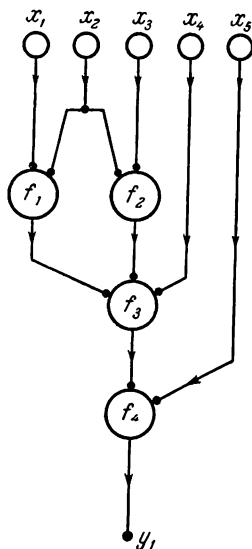


Рис. 8.4. Неизбыточный автомат A .

тем, что положим

$$\left. \begin{aligned} f_1 &= x_1 \& x_2, \\ f_2 &= \bar{x}_2 \vee \bar{x}_3, \\ f_3 &= x_6 \oplus (\bar{x}_7 \& x_4), \\ f_4 &= \bar{x}_8 \oplus \bar{x}_5. \end{aligned} \right\} \quad (8.10)$$

В соответствии с уравнениями (8.4) и (8.5), применим к этому автомату (3,1), (5,2) и (7,4) коды из раздела 3.6, чтобы получить избыточно закодированные автоматы A_1 , A_2 , A_3 соответственно:

1. (3,1) кодированный автомат A_1 .

Так как используется (3,1) код, то $k=1$ и $A=A$. Следовательно, мы получаем для искомым $f'_{\nu\mu}$ ($\mu=1, 2, 3$) следующие уравнения:

$$\left. \begin{aligned} f'_{1\mu} &= x_1 \& x_2, \\ f'_{2\mu} &= \bar{x}_2 \vee \bar{x}_3, \\ f'_{3\mu} &= ((x_{61} \& x_{62}) \oplus (x_{62} \& x_{63}) \oplus (x_{63} \& x_{61})) \oplus \\ &\quad \oplus (-(x_{71} \& x_{72}) \oplus (x_{72} \& x_{73}) \oplus (x_{73} \& x_{71})) \& x_4), \\ f'_{4\mu} &= -((x_{81} \& x_{82}) \oplus (x_{82} \& x_{83}) \oplus (x_{83} \& x_{81})) \& \bar{x}_5. \end{aligned} \right\} \quad (8.11)$$

2. (5,2) кодированный автомат A_2 .

В этом случае $k=2$ и $A=A^2$, т. е. исходный автомат A состоит из двух копий автомата A и исходные функции $f_{\nu\lambda}$ ($\lambda=1, 2$) равны

$$\left. \begin{aligned} f_{1\lambda} &= x_{1\lambda} \& x_{2\lambda}, \\ f_{2\lambda} &= \bar{x}_{2\lambda} \vee \bar{x}_{3\lambda}, \\ f_{3\lambda} &= x_{6\lambda} \oplus (\bar{x}_{7\lambda} \& x_{4\lambda}), \\ f_{4\lambda} &= \bar{x}_{8\lambda} \& \bar{x}_{5\lambda}. \end{aligned} \right\} \quad (8.12)$$

По уравнениям (8.4) и (8.5) и таблицам 3.2 и 3.3 получаем

следующие функции $f'_{\nu\mu}$ ($\mu = 1, \dots, 5$):

$$\left. \begin{aligned} f'_{11} &= f'_{15} = x_{11} \& x_{21}, \\ f'_{12} &= f'_{14} = x_{21} \& x_{22}, \\ f'_{21} &= f'_{25} = \bar{x}_{21} \vee \bar{x}_{31}, \\ f'_{22} &= f'_{24} = \bar{x}_{22} \vee \bar{x}_{32}, \\ f'_{\sigma 3} &= f'_{\sigma 1} \oplus f'_{\sigma 2} \quad (\sigma = 1, 2), \\ f'_{31} &= f'_{35} = (x_{61} \oplus (x_{61} \oplus x_{62}) \& (x_{61} \oplus x_{63} \oplus x_{65})) \oplus \\ &\quad \oplus (-(x_{71} \oplus (x_{71} \oplus x_{72}) \& (x_{71} \oplus x_{73} \oplus x_{75})) \& x_{41}), \\ f'_{32} &= f'_{34} = (x_{65} \oplus (x_{65} \oplus x_{64}) \& (x_{61} \oplus x_{63} \oplus x_{65})) \oplus \\ &\quad \oplus (-(x_{75} \oplus (x_{75} \oplus x_{74}) \& (x_{71} \oplus x_{73} \oplus x_{75})) \& x_{42}), \\ f'_{33} &= f'_{31} \oplus f'_{32}, \\ f'_{41} &= -(x_{81} \oplus (x_{81} \oplus x_{82}) \& (x_{81} \oplus x_{83} \oplus x_{85})) \& x_{51}, \\ f'_{42} &= -(x_{85} \oplus (x_{85} \oplus x_{84}) \& (x_{81} \oplus x_{83} \oplus x_{85})) \& x_{52}. \end{aligned} \right\} \quad (8.13)$$

3.(7,4) кодированный автомат A_3 .

В этом последнем случае $A = A^4$, т. е. автомат A теперь состоит из четырех копий автомата A , а исходные функции $f_{\nu\lambda}$ аналогичны функциям, заданным уравнениями (8.12), за исключением того, что $\lambda = 1, \dots, 4$. Функции $f'_{\nu\mu}$ ($\mu = 1, \dots, 7$) задаются, как и во втором примере, следующими уравнениями:

$$\left. \begin{aligned} f'_{1\varphi} &= x_{1\varphi} \& x_{2\varphi}, \\ f'_{2\varphi} &= \bar{x}_{2\varphi} \vee \bar{x}_{3\varphi}, \\ f'_{3\varphi} &= (x_{6\varphi} \oplus \pi_6) \oplus (-(x_{7\varphi} \oplus \pi_7) \& x_{4\varphi}), \\ f'_{x5} &= f'_{x1} \oplus f'_{x3} \oplus f'_{x5}, \\ f'_{x8} &= f'_{x1} \oplus f'_{x2} \oplus f'_{x4}, \\ f'_{x7} &= f'_{x1} \oplus f'_{x2} \oplus f'_{x3}, \\ f'_{4\varphi} &= -(x_{8\varphi} \oplus \pi_7) \& x_5 \quad (\varphi = 1, \dots, 4; \chi = 1, 2, 3), \end{aligned} \right\} \quad (8.14)$$

где

$$\begin{aligned} \pi_\psi &= (x_{\psi 1} \oplus x_{\psi 2} \oplus x_{\psi 4} \oplus x_{\psi 7}) \& (x_{\psi 1} \oplus x_{\psi 3} \oplus x_{\psi 6} \oplus x_{\psi 7}) \& \\ &\quad \& (x_{\psi 2} \oplus x_{\psi 5} \oplus x_{\psi 6} \oplus x_{\psi 7}) \quad (\psi = 5, 6, 7). \end{aligned}$$

На рисунках 8.5 — 8.7 показаны получающиеся в результате избыточно закодированные автоматы.

Для этих автоматов можно вычислить вероятности ошибки P_A по следующей формуле:

$$P_A = 1 - Q^p, \quad (8.15)$$

где Q — вероятность того, что данные коды не дают ошибки, и p — число различных «блоков», использующих такой

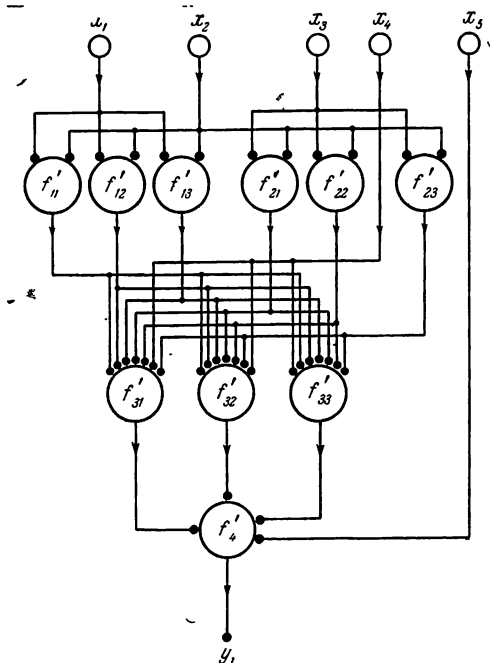


Рис. 8.5. Автомат A_1 . Избыточно закодированный вариант $A_1 (=A)$, использующий (3,1) код Хэмминга.

код. Вместо средней вероятности отсутствия ошибок $Q \cong \cong 1 - 2^{-n(C-R)}$ мы используем точную формулу, приведенную для таких кодов в разделе 3.6.

В табл. 8.1 даны вероятности ошибок, полученные для вероятности неправильного функционирования модуля $\varepsilon = 0,005$. Ошибками в выходных рядах модулей автоматов мы, конечно, пренебрегаем.

Таблица 8.1
Вероятность ошибки для A , A_1 , A_2 , A_3

Автомат	Q	p	P_A
A	0,99500	3	0,0149
A_1	0,99985	2	0,0003
A_2	0,99955	2	0,0009
A_3	0,99795	2	0,0041

Отметим, что все кодированные автоматы имеют вероятность ошибки P_A , несколько меньшую ε . Для получения в этом случае вероятностей ошибки, значительно меньших

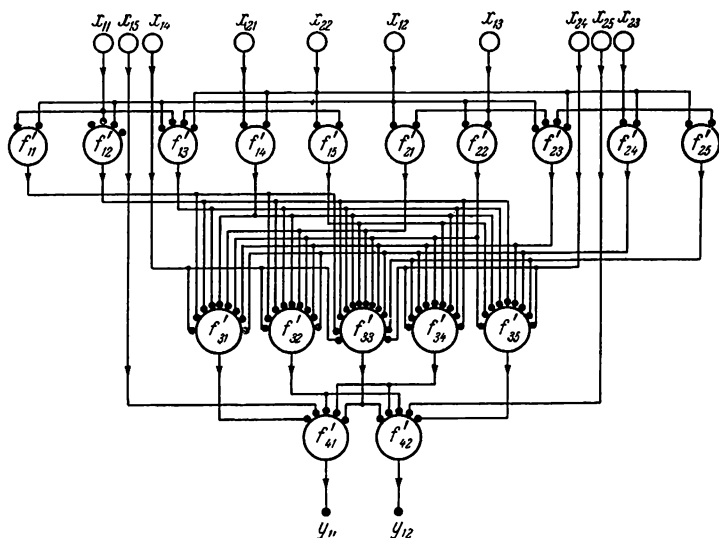


Рис. 8.6. Автомат A_2 . Избыточно закодированный вариант $A_2 (=A^2)$, использующий (5,2) код Хэмминга.

ε , требуются большие значения k и n . С другой стороны, при ε порядка 10^{-8} (7,4) закодированный автомат имел бы вероятность ошибки P_{A_3} , ощутимо меньшую, чем 10^{-6} , и меньшую, чем P_{A_1} и P_{A_2} .

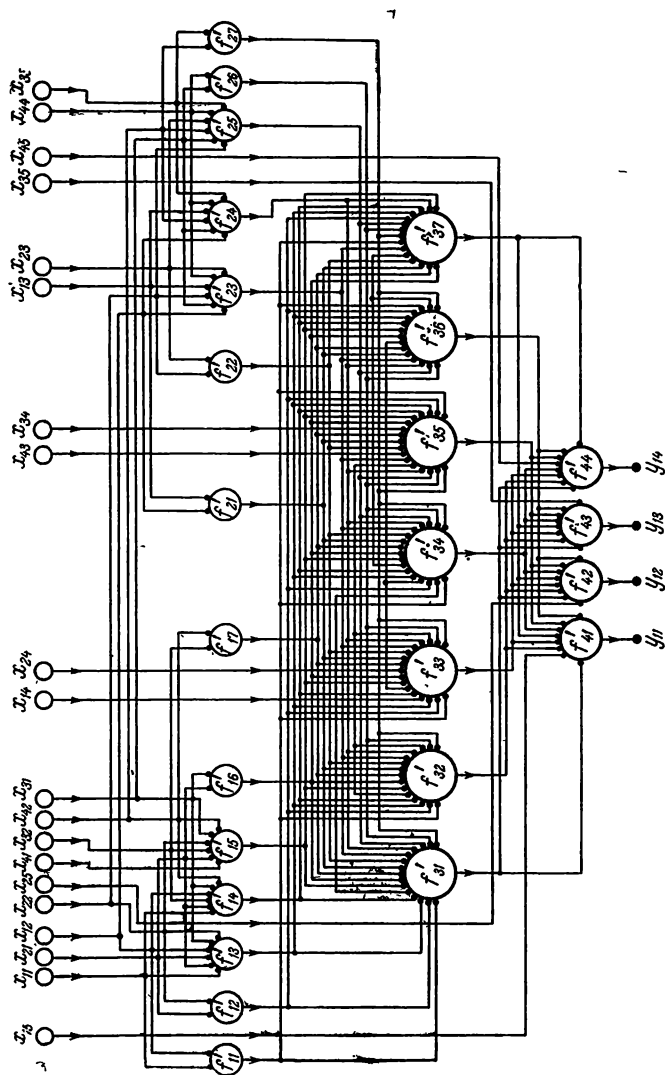


Рис. 8.7. Автомат A_8 . Избыточно закодированный вариант $A_8 (=A^4)$, использующий (7,4) код Хэмминга.

8.3. Обсуждение теоремы.

Связь с другими результатами

Справедливость теоремы 8.1 основывается на некоторых требованиях и допущениях, смысл которых мы теперь рассмотрим.

Увеличение ансамбля. Первое основное требование, содержащееся в теореме, заключается в том, что ансамбль k прототипов должен быть сделан произвольно большим. Это требование следует из того, что теорема, по существу, берет начало от теоремы Шеннона о случайном кодировании [35], т. е. теоремы 3.2, которая гласит, что вероятность неправильного функционирования P_A стремится к нулю только в случае, если k и n стремятся к бесконечности. Примеры, изображенные на рис. 8.4—8.7, показывают, как используются такие ансамбли. Изучение этого требования несколько проясняет смысл результатов, ранее рассмотренных в главе 4. Если положить $k = 1$, как у фон Неймана, Вербеека и др. и Мурогги, то ясно, что для получения произвольно малой вероятности неправильного функционирования можно использовать только $(n, 1)$ коды. В самом деле, автоматы, синтезированные Вербееком и др. и Мурогой, являются частными случаями наших автоматов, в которых каждый ряд, в общем случае, декодирует, вычисляет и кодирует, за исключением того, что k ограничено единицей и, следовательно, должны использоваться $(n, 1)$ коды.

Топология прототипа. Второе требование касается топологии каждого из k прототипов. Рассмотрим прототип A ,

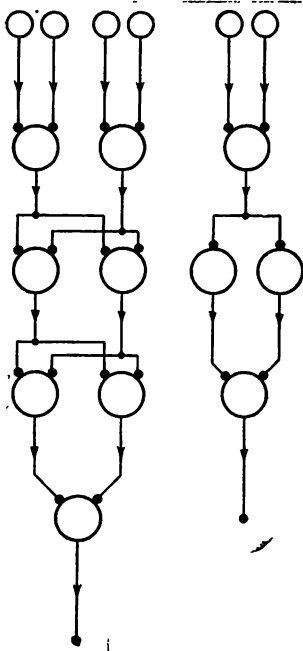


Рис. 8.8. Прототип A , состоящий из двух определенных событий различной длины.

показанный на рис. 8.8. Он включает в себя два определенных события различной глубины и различной топологии. Если на различных уровнях используются разные коды, то можно синтезировать автомат, вычисляющий эти события с

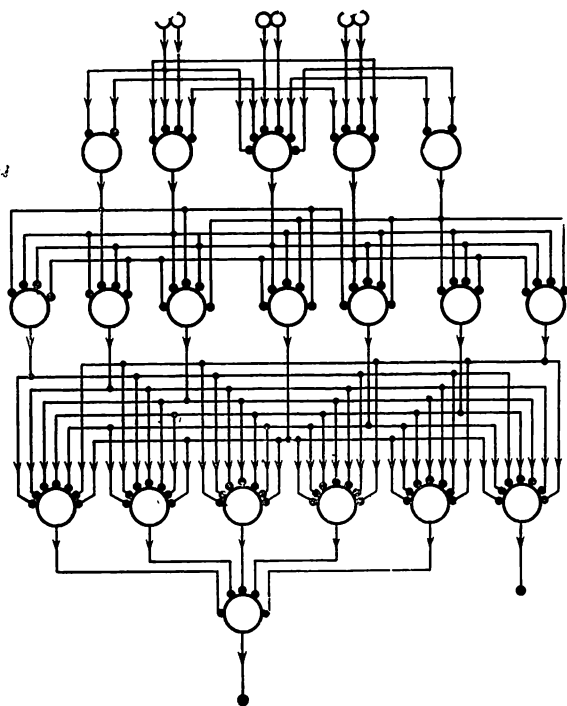


Рис. 8.9. Избыточный вариант А.

большей надежностью, чем прототип, при некотором разумном значении удельной нагрузки. На рис. 8.9 показан такой автомат, в котором для трех рядов используются соответственно (5,3), (7,4) и (5,2) коды. Возникает вопрос, можно ли применять теорему 8.1, если k определенных событий различной топологии заменены избыточным автоматом. Ответ таков, что если каждый ряд исходного ансамбля не увеличивается в ширину с возрастанием k , то теорема неприменима. Единственный случай, когда для всех рядов ширина

ряда возрастает вместе с k , имеет место, когда все k событий вычисляются сетями с одной и той же топологией. Таким образом, мы видим, что хотя k прототипов могут вычислять различные определенные события, все они должны иметь одну и ту же топологию.

Модульная сложность. Из рисунков 8.5—8.7 очевидно, что по мере возрастания k и n требуются все более сложные модули, т. е., число входов на модуль возрастает вместе с k и n . Причиной является то, что эффективность (n, k) кодов в исправлении ошибок зависит от способности таких кодов распределять информацию во многие места. Иными словами, эффективность достигается за счет одновременной посылки информации до декодирования по многим каналам, а это в свою очередь требует использования модулей, которые вычисляют функции многих переменных (см. рис. 8.2 и уравнение (8.5)). Если существует ограничение на сложность модулей, то такие коды использовать нельзя.

В самом деле, недавно было доказано (Виноград [47]), что если где-либо в автомате следует исправить t ошибок, то среднее число входов на модуль \bar{s} ограничивается выражением

$$\bar{s} \geq \frac{(2t+1)k}{n} \quad (8.16)$$

или, эквивалентно,

$$\frac{k}{n} \leq \frac{\bar{s}}{2t+1}. \quad (8.17)$$

Таким образом, если удельная нагрузка на модуль $R \cong \cong k/n$ должна оставаться постоянной, то по мере того, как исправляется все больше ошибок, число входов \bar{s} должно также возрастать. Иначе говоря, возрастание сложности модулей не только достаточно для исправления ошибок (как мы здесь показали), но также и необходимо в той или иной форме (см. Лофгрэн [19]).

Если даны модули, способные вычислять функции $n^{\bar{s}}$ переменных, то необходимо, чтобы такие функции были любыми из $2^{2^{n^{\bar{s}}}}$ возможных булевых функций. Иначе говоря, мы не ограничиваемся модулями, вычисляющими только пороговые функции (раздел 2.2), но делаем допущение, что наши модули достаточно сложны для того, чтобы вычислить любую требуемую функцию (раздел 2.7).

Ненадежность модулей в сравнении со сложностью. Очевидно также, что теорема 8.1 справедлива только тогда, когда вероятность неправильного функционирования ε модулей, составляющих избыточные сети, не зависит от k и n , т. е. работоспособность C^* не зависит от k и n . Если бы модульные ошибки возрастали с k и n (с возрастанием модульной сложности), то C^* была бы убывающей функцией k и n , и R стремилась бы к нулю вместе с P_A . Допущение, что неправильное функционирование модуля не зависит от сложности, фактически эквивалентно допущению, сделанному в шенноновской модели связи [35], о том, что по сравнению с каналами связи в кодирующих и декодирующих устройствах нет шума. Теорема 8.1 справедлива лишь в том случае, когда сделано это допущение.

В случае, когда рассмотренные выше требования не удовлетворены, т. е., если модульные сбои возрастают с модульной сложностью, или заданы только модули ограниченной сложности, то для эффективного синтеза надежных автоматов требуются несколько более сложные расчеты, чем описанные в этой монографии.

Определенные события в сравнении с неопределенными. Существует еще одно требование, необходимое для справедливости теоремы 8.1 и относящееся к характеру событий, вычисляемых модульными сетями. Напомним, что уравнение (8.1) означает следующее: для автоматов, вычисляющих события фиксированной длины (конечное b), вероятность неправильного функционирования P_A может быть сделана сколь угодно малой за счет увеличения k и n . Если, однако, b не фиксировано, то такого результата не последует. Это указывает на то, что события, вычисляемые автоматами, удовлетворяющими теореме 8.1, должны иметь конечную длину; т. е. они должны быть либо определенными событиями, либо фиксированными частями неопределенных событий. Недавно было показано (Рэбин [33]), что никакой автомат с конечными вероятностями (отличными от нуля и единицы) перехода от любого внутреннего состояния к другому (см. разделы 2.3—2.6) не может реализовать с произвольно высокой надежностью неопределенное событие. Смысл этого результата ясен. Напомним из главы 2: неопределенное событие таково, что его таблица истинности имеет неограниченную длину, и для вычисления неопреде-

ленного события, в общем, требуются циклические сети (сети с обратной связью и памятью). В таких сетях неисправленная ошибка может повлечь за собой ошибку на всех последующих выходах сети. Такая ситуация отсутствует в ациклических сетях конечной глубины, где, так сказать, для каждого вычисления производят (не принимая во внимание воздействия постоянных ошибок от повреждений) «уборку начисто». Эти результаты, вместе взятые, означают, что можно вычислить с произвольно высокой надежностью определенные события или конечные части неопределенных событий, но нельзя таким образом вычислить неопределенные события. Итак, для теоремы 8.1 мы требуем, чтобы события, вычисляемые данными автоматами, имели конечную глубину, т. е. были определенными событиями. Это ограничение определенными событиями в некотором смысле более математическое, чем физическое, так как на практике нас интересуют только таблицы некоторой фиксированной длины; можно сказать, что любой автомат вычисляет в таком случае определенные события. Таким образом, для всех практических целей у нас остается результат, что можно так синтезировать автоматы, что они будут производить вычисления с произвольно высокой надежностью и ненулевыми удельными нагрузками на модуль для некоторого фиксированного отрезка времени, но после этого времени с вероятностью, произвольно близкой к единице, они деградируют.

8.4. Эффект функционального кодирования

Заканчивая главу, подчеркиваем основной принцип, введенный в разделе 8.2, а именно: *каждый модуль кодируемого автомата декодирует закодированный вход, вычисляет заданную функцию и затем кодирует значение выхода для передачи дальше*. Эффект этой процедуры заключается в том, что ошибки исправляются в тот момент, когда они возникают, и поэтому не распространяются по автомату. Следующий

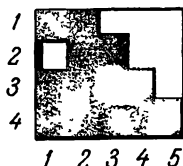


Рис. 8.10. Распределительная матрица, соответствующая прототипу A на рис. 8.4; $1, \dots, 5$ — входные переменные x_1, \dots, x_5 соответственно; $1, \dots, 4$ — модули m_1, \dots, m_4 соответственно.

эффект такой процедуры заключается в том, что с точки зрения понимания разделение системы на вычислительную и кодирующую части становится несущественным, так как эти функции выполняются повсюду внутри автомата. Последняя причина, по которой избыточно закодированные автоматы этого типа эффективны для исправления сбоев, такова, что кодирование функций описанным ранее способом (см. уравнения (8.4) и (8.5)) приводит к *распределению* данных функций повсюду в таком автомате и к *разновременности* функции каждой составной части модели. Иными словами, любая функция вычисляется во многих местах, и любой модуль участвует в вычислении многих функций. Именно *многократное распределение функции* дает возможность получить надежное функционирование, а *многократная разновременность функции* приводит к положительной удельной нагрузке.

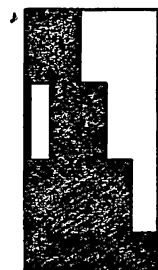


Рис. 8.11. Распределительная матрица, соответствующая автомату A_1 на рис. 8.5.

Рис. 8.12. Распределительная матрица, соответствующая автомату A_2 на рис. 8.6.

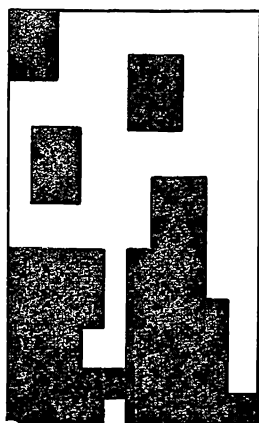


Рис. 8.12. Распределительная матрица, соответствующая автомату A_2 на рис. 8.6.

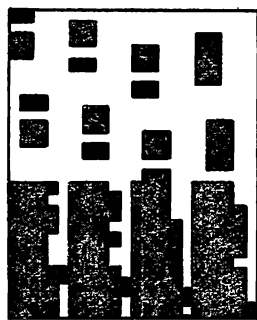


Рис. 8.13. Распределительная матрица, соответствующая автомату A_3 на рис. 8.7.

Мы можем продемонстрировать эту особенность функционального кодирования с помощью распределительных

матриц, то есть массивов, в которых входы автоматов являются ординатами, а модульные координаты — абсциссами. Например, прототип, изображенный на рис. 8.4, представлен матрицей, приведенной на рис. 8.10. Рисунки 8.11—8.13 изображают матрицы для автоматов A_1 , A_2 и A_3 , показанных на рисунках 8.5—8.7. Возросшую разновременность в A_2 и A_3 легко заметить по увеличению количества входов, соединенных со все большим количеством модулей. По мере возрастания k и n эффект становится все более заметным, до тех пор, пока структурно и функционально любой модуль не становится связанным с большинством других. Мы назвали получающиеся в результате сети *анастомотическими* из-за их структурного сходства с нервными сетями коры головного мозга.

ГЛАВА ДЕВЯТАЯ

ОШИБКИ ПЕРЕДАЧ И СТРУКТУР

При анализе модульных ошибок в анастомотических сетях было принято, что ошибки ни в *соединениях*, ни в *структуре* *) не имели места. Для сетей со сколь угодно большой степенью анастомоза это допущение несправедливо. При последующем обсуждении будем предполагать, что наряду с модульными ошибками могут иметь место следующие ошибки:

1. Каждое соединение функционирует, как двоичный передающий канал с шумом с $\text{Pr}(1|0) = p_4$, $\text{Pr}(0|1) = p_3$.

2. Каждое соединение может быть выполнено правильно с вероятностью $1 - p$ и неправильно с вероятностью p .

9.1. Синаптический шум

Вообще говоря, для ошибок первого типа в анастомотических сетях могут возникнуть два случая. Рис. 9.1 пояснит эту ситуацию. Каждый модуль анастомотической сети

*) Под структурой сети в данном случае понимается множество всех взаимных соединений модулей. (Прим. ред.)

связан с другими модулями соединительной структурой, состоящей из одного соединения, которое разветвляется или «карборизируется» во множество окончаний. Ошибки могут иметь место либо до точки разветвления v_0 либо после нее.

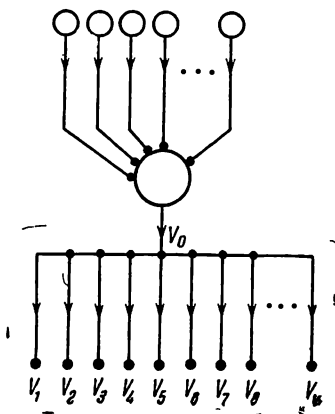


Рис. 9.1. Анастомозический модуль.

Ошибки, имеющие место до точки разветвления v_0 , очевидно, могут быть объединены с модульными ошибками так, что этот вид может быть отнесен к общему случаю модульной ошибки, который рассматривался в главе 8. Итак, предположим, что модульные ошибки характеризуются двоичным каналом с переходными вероятностями $\text{Pr}(0|1) = \varepsilon_1$, $\text{Pr}(1|0) = \varepsilon_2$. Этот случай изображен на рис. 9.2, где y_β представляет требуемый выход модуля, z_γ — выход модуля с шумом и x_r — выход соединения с шумом.

Мы можем объединить два множества ошибок, используя формулу

$$\text{Pr}(x_r | y_\beta) = \sum_{\gamma=1}^{k_s} \text{Pr}(x_r | z_\gamma) \text{Pr}(z_\gamma | y_\beta). \quad (9.1)$$

Таким образом,

$$\left. \begin{aligned} \text{Pr}(x_1 | y_1) &= (1 - \varepsilon_1)(1 - p_3) + \varepsilon_1 p_4 = 1 - r_1, \\ \text{Pr}(x_0 | y_0) &= (1 - \varepsilon_2)(1 - p_4) + \varepsilon_2 p_3 = 1 - r_2, \\ \text{Pr}(x_0 | y_1) &= 1 - \text{Pr}(x_1 | y_1) = r_1, \\ \text{Pr}(x_1 | y_0) &= 1 - \text{Pr}(x_0 | y_0) = r_2. \end{aligned} \right\} \quad (9.2)$$

Получающийся в результате двоичный канал показан на рис. 9.3.

Ошибки, которые имеют место после точки разветвления v_0 , можно объединить с модульными ошибками таким же образом. Рассмотрим модуль, показанный на рис. 9.1. Предположим, что в добавление к модульным ошибкам

ε_1 и ε_2 и доузловым ошибкам передачи p_3 и p_4 , независимо имеют место послеузловые ошибки передачи p_5 и p_6 в каждом из соединений от v_0 до v_1, v_2, \dots, v_w соответственно. Очевидно, мы можем объединить эти ошибки с другими, как и

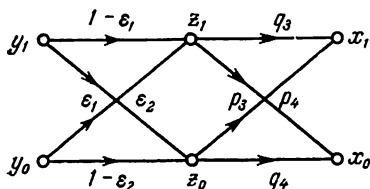


Рис. 9.2. Двойной канал с шумом, представляющий модульную ошибку и ошибку соединения.

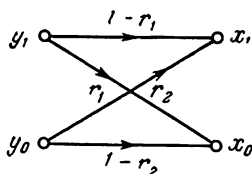


Рис. 9.3. Приведенный двойной канал.

раньше. Итак, рассмотрим путь из v_0 к любому v_w , изображенный на рис. 9.4.

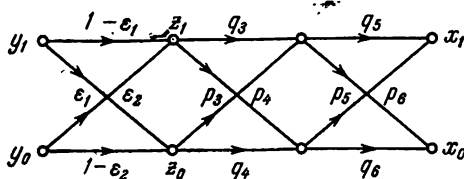


Рис. 9.4. Тройной канал с шумом, представляющий ошибки: модульную, передачи и синаптическую.

Последовательное применение уравнения (9.2) приводит к каналу с переходными вероятностями, равными

$$\left. \begin{aligned} \Pr(v_w = 1 | y_1) &= (1 - r_1)(1 - p_5) + r_1 p_6 = 1 - t_1, \\ \Pr(v_w = 0 | y_1) &= t_1, \\ \Pr(v_w = 1 | y_0) &= t_2, \\ \Pr(v_w = 0 | y_0) &= (1 - r_1)(1 - p_6) + r_2 p_5 = 1 - t_2. \end{aligned} \right\} \quad (9.3)$$

Таким образом, мы заменили тройной канал, показанный на рис. 9.4, тройным каналом, приведенным на рис. 9.5. В данном случае весь шум «взят вместе» как послеузловой

шум передачи. Это значит, что мы фактически можем рассматривать шум в сети как межмодульный или

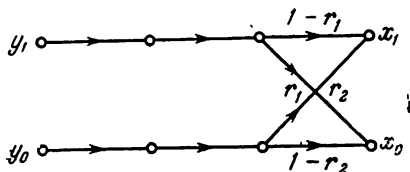


Рис. 9.5. Приведенный тройной канал.

синаптический шум (см. Аллансон [1]), т. е. мы можем заменить систему, показанную на рис. 9.1, схемой, приведенной на рис. 9.6, в которой послеузловой шум передачи снова можно объединить с существующим синаптическим шумом (p_1, p_2), применяя уравнение (9.2). Это прекрасно согласуется с тем фактом, что модули следующего ряда декодируют каждый входной сигнал, исправляя тем самым эти ошибки.

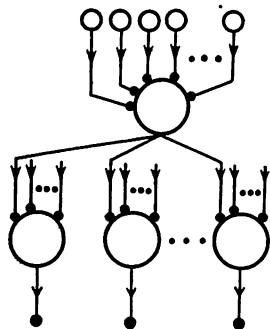


Рис. 9.6. Анастомотическая модульная сеть, в которой весь шум рассматривается с точки зрения синаптического происхождения.

Ясно также, что фактическое воздействие послеузлового шума заключается в декорреляции входных сигналов в следующий ряд. Это явно уменьшает работоспособность, так как декорреляция означает, что $H(Y | Z)$ возросла, а следовательно, $H(Y) - H(Y | Z)$ уменьшилась.

Если бы коррелированные ошибки имели место после узлов, то ситуация была бы несколько лучше. Однако этот вопрос требует дальнейшего изучения.

9.2. Ошибки в структурах

Полученные результаты предполагают, что все ошибки могут быть отнесены к синапсу и что можно использовать корректирующие коды, как это было показано ранее, для

исправления таких ошибок в следующем ряду. В частности, предполагается, что ошибки схем соединений, имеющие место с вероятностью p , могут рассматриваться как дополнительные ошибки в кодовом слове. Следующая теорема указывает на допустимость такой трактовки.

Теорема 9.1. Пусть B — структура надежного автомата A , работающего при удельной нагрузке R , и пусть C^* — работоспособность модулей, составляющих A . Если B нарушается с вероятностью $p < \beta(C^* - R)$ (т. е. каждое соединение автомата может быть выполнено неправильно с вероятностью p и правильно с вероятностью $1 - p$), то с вероятностью, произвольно близкой к единице, все же можно сделать так, что A будет функционировать с произвольно высокой надежностью. (β будет определено в доказательстве.)

Доказательство. Рассмотрим любой модуль этого автомата. Пусть его входными соединениями будут $c_{11}, c_{12}, \dots, c_{1n}; c_{21}, c_{22}, \dots, c_{2n}; \dots; c_{s_v1}, c_{s_v2}, \dots, c_{s_vn}$, где принимается, что все эти входы поступают от других модулей автомата. Пусть это множество представлено c_{rn} и пусть c'_{rn} — «шумовой» вариант c_{rn} . Разобьем множество всех c'_{rn} на два взаимно исключающих и исчерпывающих множества G и \bar{G} таких, что

$$c'_{rn} \in \begin{cases} G & \text{тогда и только тогда, когда } f'_{v\mu}(c'_{rn}) = f'_{v\mu}(c_{rn}), \\ \bar{G} & \text{тогда и только тогда, когда } f'_{v\mu}(c'_{rn}) \neq f'_{v\mu}(c_{rn}), \end{cases} \quad (9.4)$$

т. е. G есть множество c'_{rn} , вызывающее такой же модульный выход, как и его исходный вариант c_{rn} , а \bar{G} — дополнение G .

Очевидно, вероятность ошибки в этом случае равна

$$P_c = \sum_{c'_{rn} \in \bar{G}} \text{Pr}(c'_{rn} | c_{rn}). \quad (9.5)$$

В случае, когда необходимое соединение отсутствует, например, c_{11} , множество всех возможных c'_{rn} представляет собой множество всех таких c'_{rn} , что $c_{11} = 0$. Аналогично, если это соединение сделано, когда оно не должно быть,

множество всех c'_{rn} содержится в множестве всех таких c'_{rn} , что $c_{11} = 1$. Таким образом, наихудший случай ошибок, вызванных неправильными соединениями, имеет место тогда, когда значение c_{11} всегда меняется.

Пусть ${}_{11}c'_{rn}$ представляет множество $\bar{c}'_{11}, c'_{12}, \dots, c'_{s_v n}$. Оценим вероятность неправильного функционирования каждый раз, когда ${}_{11}c'_{rn}$ является входом вместо c'_{rn} . Сделаем это, рассматривая ${}_{11}c'_{rn}$ как результат модульного шума, а не как ошибку структуры.

Разобьем G на два взаимно исключающих и исчерпывающих множества G_1 и $G_2 = G - G_1$ таких, что

$$c'_{rn} \in \begin{cases} G_1 \text{ тогда и только тогда, когда } c'_{rn} \cup {}_{11}c'_{rn} \in G, \\ G_2 \text{ тогда и только тогда, когда } c'_{rn} \in G \text{ и } {}_{11}c'_{rn} \in \bar{G}, \end{cases} \quad (9.6)$$

т. е. c'_{rn} содержится в G_2 тогда и только тогда, когда перестановка соединения c_{11} переводит его из G в \bar{G} .

Таким образом, когда c_{11} неправильно соединено, имеем

$$P_c \leq \sum_{c'_{rn} \in \bar{G} \cup G_2} \Pr(c'_{rn} | c_{rn}), \quad (9.7)$$

$$P_c = \sum_{c'_{rn} \in \bar{G}} \Pr(c'_{rn} | c_{rn}) + \sum_{c'_{rn} \in G_2} \Pr(c'_{rn} | c_{rn}). \quad (9.8)$$

Допустим, что ошибки на входах $c_{11}, \dots, c_{s_v n}$ независимы, так что

$$\Pr(c'_{rn} | c_{rn}) = \prod_{\substack{r=1, \dots, s_v \\ \mu=1, \dots, n}} \Pr(c'_{r\mu} | c_{r\mu}). \quad (9.9)$$

Таким же образом

$$\Pr({}_{11}c'_{rn} | c_{rn}) = \Pr(\bar{c}'_{11} | c_{11}) \prod_{\substack{r=2, \dots, s_v \\ \mu=2, \dots, n}} \Pr(c'_{r\mu} | c_{r\mu}). \quad (9.10)$$

Тогда

$$\frac{\Pr({}_{11}c'_{rn} | c_{rn})}{\Pr(c'_{rn} | c_{rn})} = \frac{\Pr(\bar{c}'_{11} | c_{11})}{\Pr(c'_{11} | c_{11})}. \quad (9.11)$$

Предположим, что ошибка структуры представлена двоичным каналом, показанным на рис. 9.7. Тогда $\Pr(\bar{c}_{11} | c_{11}) : \Pr(c'_{11} | c_{11})$ равна либо $p_7/(1 - p_7)$, либо $p_8/(1 - p_8)$.

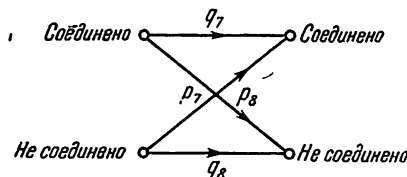


Рис. 9.7. Символическое представление ошибки в структуре.

Выберем большее из этих двух чисел (т. е. завысим ошибку). Назовем эту величину α , так что

$$\frac{\Pr(\bar{c}_{11} | c_{11})}{\Pr(c'_{11} | c_{11})} \leq \alpha. \quad (9.12)$$

Из уравнений (9.11) и (9.12) ясно, что

$$\Pr(c'_{11} | c_{rn}) \leq \alpha \Pr(c'_{rn} | c_{rn}). \quad (9.13)$$

Теперь каждому множеству c'_{rn} из G_2 соответствует множество c_{11} из \bar{G} . Таким образом, имеем

$$\begin{aligned} \sum_{c'_{rn} \in G_2} \Pr(c'_{rn} | c_{rn}) &= \\ &= \sum_{c'_{rn} \in \bar{G}} \Pr(c'_{rn} | c_{rn}) \leq \alpha \sum_{c'_{rn} \in G} \Pr(c'_{rn} | c_{rn}). \end{aligned} \quad (9.14)$$

Итак, из уравнений (9.8) и (9.14) мы имеем вероятность ошибки, когда c_{11} соединено неправильно:

$$\begin{aligned} P_c &\leq \sum_{c'_{rn} \in \bar{G}} \Pr(c'_{rn} | c_{rn}) + \alpha \sum_{c'_{rn} \in G} \Pr(c'_{rn} | c_{rn}) = \\ &= (1 + \alpha) \sum_{c'_{rn} \in G} \Pr(c'_{rn} | c_{rn}). \end{aligned} \quad (9.15)$$

Но

$$\sum_{c'_{rn} \in \bar{G}} \Pr(c'_{rn} | c_{rn}) = P, \quad (9.16)$$

где P — вероятность ошибки, связанная с данным (n, k) кодом, т. е. c'_{rn} рассматривается именно как «шумовой» вариант c_{rn} . Следовательно,

$$P_c \leq (1 + \alpha) P. \quad (9.17)$$

Аналогичным способом можно определить $11 \cdot 12 c'_{rn}$ и G_1 такие, что c'_{rn} содержится в G_1 тогда и только тогда, когда $c'_{rn} \cup 11 c'_{rn} \cup 12 c'_{rn} \cup 11 \cdot 12 c'_{rn}$ содержится в G . В этом случае коэффициент при P равен $(1 + \alpha + \alpha + \alpha^2)$, и

$$\begin{aligned} & \sum_{c'_{rn} \in G} \Pr(c'_{rn} | c_{rn}) + \\ & + \sum_{c'_{rn} \in \bar{G}_2} \Pr(c'_{rn} | c_{rn}) \leq (1 + 2\alpha + \alpha^2) P = (1 + \alpha)^2 P. \end{aligned} \quad (9.18)$$

Иными словами, два соединения могут быть выполнены неправильно в пучке из n соединений с вероятностью $P_c \leq (1 + \alpha)^2 P$. В общем случае, если не более l линий на пучок соединены неправильно, то

$$P_c \leq (1 + \alpha)^l P. \quad (9.19)$$

Таким образом, если не более l соединений на пучок в ряду из m модулей выполнены неправильно, то мы можем достигнуть невосприимчивости к такому воздействию, используя (n, k) код, для которого вероятность ошибки P' равна

$$P' \leq \frac{P}{(1 + \alpha)^l m}, \quad (9.20)$$

так что вероятность ошибки после декодирования будет меньше P .

Принимая, что P' есть вероятность ошибки соединения модулей, усредненная по ансамблю всех возможных кодов

(глава 3), т. е.

$$P' \cong 2^{-m(C^*-R)}, \quad (9.21)$$

получаем требование, чтобы

$$2^{-m(C^*-R)} \leq \frac{P}{(1+\alpha)^t m}. \quad (9.22)$$

Из него мы имеем следующее уравнение:

$$ms_v \leq \frac{s_v}{\log_2(1+\alpha)} (\log_2 P - \log_2 m + m(C^* - R)), \quad (9.23)$$

т. е. число неправильно выполненных соединений на модуль задается уравнением (9.23). Имеется, самое большее, ms_v соединений на модуль таких, что

$$\frac{ms_v}{ms_v} = f \leq \frac{1}{\log_2(1+\alpha)} \left(\frac{1}{m} \log_2 P - \frac{\log_2 m}{m} + (C^* - R) \right). \quad (9.24)$$

Для достаточно большого m

$$f \leq \frac{1}{\log_2(1+\alpha)} (C^* - R). \quad (9.25)$$

Ясно, что этот результат справедлив для всех модулей, так что если доля неправильно выполненных соединений в ряду из m модулей меньше f , то вероятность ошибки при декодировании P , которая имела бы место при отсутствии неправильных соединений, все же может быть достигнута за счет использования кода с вероятностью ошибки P' , задаваемой уравнением (9.20). По закону больших чисел (Феллер [13]), если неправильные соединения имеют место с вероятностью $p < f$, то с вероятностью, очень близкой к единице, доля неправильных соединений будет меньше f . Определим β как

$$\beta = \frac{1}{\log_2(1+\alpha)}. \quad (9.26)$$

Тогда с вероятностью, произвольно близкой к единице, ошибка после декодирования будет меньше P , если неправильные соединения имеют место с вероятностью $p < \beta(C^* - R)$, что и требовалось доказать.

9.3. Обсуждение результатов

Смысл этого результата очевиден. Задаваясь автоматом с минимальной избыточностью, таким, что $C^* - R$ очень мало, можно допустить очень малое число ошибок в структуре, если надежность автомата должна остаться ограниченной. Если использовать большую избыточность ($C^* - R$ больше), то можно дополнительно применить доступные «схемные способы», чтобы нейтрализовать воздействия неправильных соединений.

Следует, однако, отметить, что результат, по существу, получен при рассмотрении ансамбля всех возможных схем соединений и что в среднем такое поведение можно ожидать. Точнее, при доказательстве теоремы 9.1 мы рассматривали всевозможные схемы неправильных соединений как ансамбль, а затем к этому ансамблю применяли закон больших чисел. Объединяя теоремы 8.1 и 9.1, мы можем получить следующий результат: автомат, состоящий из модулей с положительной работоспособностью C^* , соединенных в вероятностном порядке, можно построить так, что с высокой степенью вероятности он будет произвольно надежным, если не учитывать ошибок в элементах последнего ряда и предусматривать, что автомат является ациклическим. Более широко, с большой вероятностью, можно построить автоматы, которые будут функционировать с произвольно высокой надежностью, за исключением ошибок на выходе, несмотря на неправильное функционирование их модулей, соединений и структур. Если такие автоматы должны иметь ненулевые удельные нагрузки R , то для этого необходимо только, чтобы неправильное функционирование модулей и соединений приводило к некоторой ненулевой работоспособности C^* такой, что $C^* > R$, и достаточно, чтобы автомат обладал многократной разновременностью функций и структур, отмеченной в главе 8.

ГЛАВА ДЕСЯТАЯ ЗАКЛЮЧЕНИЕ

В главах 5—9 мы кратко рассмотрели общую теорию вычислений при наличии шума и продемонстрировали применимость теории информации для синтеза надежных авто-

матов из менее надежных элементов. Основной результат заключается в том, что несмотря на модульные, синаптические ошибки и ошибки структур, можно построить автоматы, вычисляющие определенные события с произвольно малыми частотами ошибки, за исключением, конечно, окончательных ошибок на их выходе. В работе Рэбина [33] устанавливается тот факт, что без использования надежных элементов вообще нельзя построить автоматы для вычисления неопределенных событий с произвольно малыми частотами ошибок. Мы приняли это в том смысле, что практически из ненадежных элементов можно построить автоматы, вычисляющие с произвольно малыми частотами ошибки в течение некоторого конечного периода времени равномерные отрезки неопределенных событий. Было показано, что необходимые избыточности, требуемые в таком автомате, изменяются обратно пропорционально сложности заданных модулей. В частности, если модульные ошибки были независимы от модульной сложности, то для получения требуемой надежности было достаточно функционально избыточных анастомотических модульных сетей, основанных на кодах, корректирующих ошибки. Было также показано, что синаптические ошибки могут быть объединены в одно целое с основными ошибками так, что воздействия таких ошибок можно контролировать, используя еще более избыточные сети. Подобным образом показано, что точно так же можно контролировать ошибки в соединениях. Таким образом, по структурной организации такой автомат представляет собой взаимосвязанную сеть сложных модулей, имеющую некоторую долю случайности в соединениях, более вероятную в малом, чем в большом. Соответствующая функциональная организация представляет собой сеть, в которой исходные функции вычисляются во многих местах, и любой вычисляющий модуль представляет вообще отдельную сложную булеву функцию, являющуюся композицией многих исходных функций. Мы назвали эту организацию многократно разновременной функциональной организацией.

Связь этих результатов с другими работами. При сопоставлении этих результатов с другими работами следует отметить, что по типу избыточность, применяемую в этих автоматах, можно назвать фиксированной или статической

избыточностью. Иными словами, рассмотренные нами модульные автоматы имеют фиксированную структуру, которая не меняется во времени. Очень вероятно, что можно было бы создать более надежные автоматы с переменной структурой, которые могли бы контролировать повреждения и сбои за счет надлежащего структурного и функционального изменения, особенно, если бы имелись более сложные модули, чем те, которые мы до сих пор рассматривали. В самом деле, Пирс [30] уже показал эффективность адаптивных модульных сетей для исправления ошибок. Эти сети содержат модули, способные вычислять относительные надежности своих входов, и, соответственно, взвешивать такие входы. Использование обратной связи позволяет таким модулям стабилизировать собственную работу. При этом общая надежность (срок службы) таких модульных сетей значительно возрастает. Аналогично, Лофгрэн [19, 21] показал эффективность самовосстанавливающихся систем с точки зрения продления срока их службы. Такие системы автоматически заменяют неправильно функционирующие модули или используют самовоспроизводящиеся модульные сети (фон Нейман [40]; Мур [25]). Лофгрэн показал, что максимальный срок службы конечен только для такого самовоспроизводящегося автомата, структура которого хорошо локализована (т. е. чьи клеммы для входов и выходов определены в смысле их числа и их относительного положения). Однако срок службы не строго локализованных самовоспроизводящихся автоматов может быть неограниченным. Связи, если они вообще существуют, между такими жизнеспособными автоматами и нежизнеспособными автоматами, которые мы рассмотрели, требуют дальнейшего изучения.

Мы завершаем эту монографию, вновь подчеркивая ту мысль, что решающим фактором, фигурирующим во всех схемах, предназначенных для повышения надежности или срока службы автоматов, является *модульная сложность*. Мы показали формально, как модульную сложность можно заменить избыточностью и как такая замена влияет на надежность конечных автоматов.

ПРИЛОЖЕНИЕ

Доказательство теоремы 6.3

Мы будем строго следовать первоначальному доказательству теоремы 3.2, данному Шенноном [35]. Рассмотрим дискретный источник S , выбирающий сообщения из ансамбля X . Последовательности символов длины n используются в качестве входов и выходов в данной сети (см. рис. 6.2). С большой вероятностью (стремящейся к 1 по мере возрастания n) имеем, что:

1) входные последовательности принадлежат подансамблю X^n из X , состоящему примерно из $2^{nH(X)}$ последовательностей;

2) каждая из требуемых выходных последовательностей принадлежит подансамблю, содержащему $2^{nH(Y)}$ последовательностей, и может быть результатом любой из $2^{nH(X|Y)}$ последовательностей;

3) каждая действительная выходная последовательность принадлежит подансамблю Z^n из Z , содержащему $2^{nH(Z)}$ последовательностей, каждая из которых может быть результатом примерно $2^{nH(Y|Z)}$ последовательностей;

4) утверждения 1, 2 и 3 остаются справедливыми, если X , Y и Z заменены на X' , Y' и Z' соответственно.

Структура этой модели, которая соответствует в некотором смысле расчлененному модульному каналу на рис. 5.2, показана на рис. А.1. Мы устанавливаем случайное соответствие между последовательностями из X (последовательностями сообщений) и последовательностями из X' (последовательностями сигналов). То есть мы устанавливаем случайное соответствие между сигнальными последовательностями в высоковероятностном подансамбле ансамбля X' , содержащего около $2^{nH(X')}$ последовательностей,

и последовательностями сообщений высоковероятностного подансамбля ансамбля X_r , содержащего около $2^{nH(X_r)}$

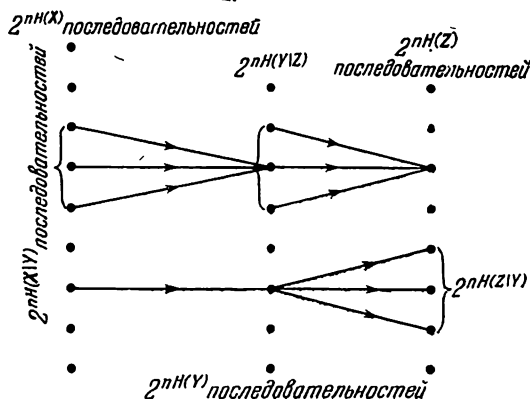


Рис. А.1. Структура модели рассчитанной вычислительной системы.

последовательностей. Если принять $H(X_r) < H(X'_r)$ ($r = 1, \dots, s$), то это соответствие можно всегда выполнить.

Тогда:

1. Вероятность того, что для заданной последовательности из X_r данная последовательность из X'_r не будет выбрана в качестве кодовой последовательности, равна $1 - 2^{-nH(X'_r)}$.

2. Вероятность, что для любой последовательности из X_r данная последовательность из X'_r не будет выбрана в качестве кодовой последовательности, равна

$$(1 - 2^{-nH(X'_r)})^{2^{nH(X_r)}} \cong 1 - 2^{n(H(X_r) - H(X'_r))}$$

3. Вероятность, что данная последовательность из X' является кодовой точкой, равна

$$2^{n \left(\sum_{r=1}^s H(X_r) - \sum_{r=1}^s H(X'_r) \right)} = 2^{n(H(X) - H(X'))}$$

4. Вероятность, что ни одна из $2^{nH(X'|Y')}$ последовательностей из X' , в результате которых может получиться данная последовательность из Y' , не выбрана в качестве кодовой последовательности, равна

$$(1 - 2^{n(H(X) - H(X'))})_{2^{nH(X'|Y')}}.$$

Следовательно, если данная последовательность из Z' принята декодирующим устройством, то она будет декодирована правильно только в том случае, если ни одна из $2^{nH(Y'|Z')}$ последовательностей из Y' (иных, чем та последовательность из Y' , которая действительно вызвала данную последовательность из Z') не является результатом тех последовательностей из X' , которые были выбраны в качестве кодовых последовательностей. Следовательно, вероятность Q правильного декодирования равна

$$Q \cong (1 - 2^{n(H(X) - H(X'))})_{2^{n(H(X'|Y') + H(Y'|Z'))}}, \quad (\text{A.1})$$

но по уравнению (5.4)

$$H(X') - H(X' | Y') = H(Y') \quad (\text{A.2})$$

и, следовательно (см. главу 5),

$$H(Y') - H(Y' | Z') = C^*, \quad (\text{A.3})$$

откуда следует

$$Q \cong (1 - 2^{n(H(X) - H(X'))})_{2^{n(H(X') - C^*)}} \cong 1 - 2^{n(H(X') - C^*)}. \quad (\text{A.4})$$

Таким образом, в среднем, вероятность неправильного декодирования сигнальной последовательности, выбранной случайно для последовательностей сообщений, равна $2^{-n(C^* - H(X'))}$ и, следовательно, существует по крайней мере один код, действие которого ничуть не хуже задаваемого этим средним значением вероятности, *что и требовалось доказать.*

Обсуждение. В том случае, когда $s = 1$, теорема 6.3 превращается в теорему 3.2. При этом первое условие теоремы 6.3 превращается в $H(X) < H(X')$ и является следствием факта, что в этом случае $H(X) < C^* = H(X') - H(X' | Z')$.

При формулировании и доказательстве теоремы 6.3 мы приняли допущение, что последовательности, выбираемые

любым источником S_r , кодируются независимо от последовательностей, выбираемых любыми другими источниками. Это допущение обеспечивает то, что требуемое вычисление f не выполняется в идеальном кодирующем устройстве. Такого же допущения, обеспечивающего отсутствие вычисления в декодирующем устройстве, не сделано. В самом деле, с большой вероятностью, любой случайно выбранный код является таким, что декодирующее устройство может опознать входные последовательности и последовательно выполнить вычисление f , то есть в данном случае ненадежный вычислительный модуль используется только для связи, а не для вычисления. Так, рассмотрим вероятность того, что данная последовательность из Y' может быть результатом некоторой кодовой последовательности из X' . Эта вероятность равна

$$2^{n(H(X) - H(X') + H(X'|Y'))}$$

и, следовательно, ожидаемое количество последовательностей из Y' , которые являются результатом случайного выбора кодовых последовательностей из X' , равно $2^{n(H(X) - H(X') + H(X'|Y) + H(Y'))}$, что сводится (см. главу 5) к $2^{nH(X)}$. То есть почти каждая закодированная последовательность из X' дает единственную последовательность из Y' и, таким образом, когда декодирующее устройство опознает последовательность из Y' , оно также опознает последовательность из X' , а не просто подмножество последовательностей.

Отметим, что этот аргумент не доказывает, что *каждый* код обладает этим свойством. Скорее он предполагает, что коды, для которых декодирующее устройство не выполняет требуемое вычисление, являются исключением, а не правилом. Например, если требуемым вычислением является сложение по модулю два ($x_1 \oplus x_2$), то можно так использовать группу кодов для достижения произвольно высокой надежности вычисления этой функции, что кодирующее и декодирующее устройства не будут играть роли в вычислении, а будут выполнять только функции кодирования и декодирования для целей обнаружения и исправления ошибок.

ЛИТЕРАТУРА

1. Allanson J. T., Proc. 1st Int. Cong. Cybernetics, 1956, Gauthier-Villars, Paris, 1959, стр. 687—694.
2. Armstrong D. B., Bell System Tech. J. 40, 577—594 (1961).
3. Blum M., 1960, Principles of Self-Organization, H. von Foerster and G. Zopf, Jr., Eds., Pergamon Press, New York (1962), стр. 95—119.
4. Boltzmann L., Wien. Ber. 63, 397 (1872).
5. Boole G., The Mathematical Analysis of Logic, Cambridge, 1847.
6. Бриллюен Л., Наука и теория информации, М., Физматгиз, 1960.
7. Cowan J. D., 1960a, Principles of Self-Organization, H. von Foerster and G. Zopf, Jr., Eds., Pergamon Press, 1962, стр. 135—179.
8. Cowan J. D., Proc. 1st Bionics Symposium, Technical Report WA60—600, Wright-Patterson Air Force Base, Ohio, 1960b, стр. 93—152.
9. Eden M., J. Information and Control 2, 310—313 (1959).
10. Elias P., IBM J. Res. Develop. 3, 346—353 (1958).
11. Фано Р., Nuovo cimento 13, X, Supplement 2, 353—372 (1959).
12. Фано Р., Передача информации. Статистическая теория связи, М., «Мир», 1965.
13. Feller W., An Introduction to Probability Theory and Its Applications, Wiley, New York, 1952.
14. Gabor D., J. Inst. Elec. Engrs. (London) 93, 429—457 (1946).
15. Гиббс Дж. В., Основные принципы статистической механики, излагаемые со специальным применением к рациональному обоснованию термодинамики, М.—Л., Гостехиздат, 1946.
16. Godel K., Monatsh. Math. u. Phys. 38, 173—189 (1931).
17. Hamming R. W., Bell System Tech. J. 29, 147—160 (1950).
18. Клини С. К., Автоматы, Сб. статей под ред. К. Э. Шеннона и Дж. Маккарти, М., ИЛ, 1956.
19. Lofgren L., 1960, Principles of Self-Organization, H. von Foerster and G. Zopf, Jr., Eds., Pergamon Press, New York, 1962, стр. 181—228.
20. Lofgren L., Report A510, Research Institute of National Defence, Stockholm, Sweden, 1962a.
21. Lofgren L., 1962b, Biological Prototypes and Synthetic Systems, E. Barnard and M. Kare, Plenum Press Inc., New York, 1962, стр. 342—369.

22. McCulloch W. S., Brookhaven Symposium in Biology 10, 207—215, Office of Technical Services, U. S. Department of Commerce, Washington, 1957.
23. McCulloch W. S., 1960, Principles of Self-Organization, H. von Foerster and G. Zopf, Jr., Eds., Pergamon Press, New York, 1962, стр. 91—94.
24. McCulloch W. S. and Pitts W. H., Bull. Math. Biophys. 5, 115—133 (1943).
25. Moore E. H., Proc. Symposium on Mathematical Problems in Biological Sciences, American Mathematical Society, New York, 1961.
26. Moore E. H. and Shannon C. E., J. Franklin Inst. 262, 191—208, 281—297 (1956).
27. Muroga S., Rome Air Development Center, New York, Technical Note 60-146 (1960).
28. Питерсон У. У., Коды, исправляющие ошибки, М., «Мир», 1964.
29. Peterson W. W., Rabin M. O., IBM J. Res. Develop. 2, 163—168 (1959).
30. Peirce W. H., Redundancy Techniques for Computing System, R. H. Wilcox and W. C. Mann, Eds., Spartan Books, Washington, D. C., 1962, стр. 229—251.
31. Pitts W. H. and McCulloch W. S., Bull. Math. Biophys. 9, 127—147 (1947).
32. Ray-Chaudhuri D. K., Bell System Tech. J. 40, 595—611 (1961).
33. Rabin M. O., Personal communication, 1962.
34. Шеннон К. Э., Работы по теории информации и кибернетике, Сб. статей, М., ИЛ, 1963.
36. Slepian D., Bell System Tech. J. 35, 203—234 (1956).
37. Szilard L., Z. Physik 53, 840—856 (1929).
38. Turing A. M., Proc. Lond. Math. Soc., Ser. 2 42, 230—265 (1937).
39. Verbeek L. A. M., 1960, Principles of Self-Organization, H. von Foerster and G. Zopf, Jr., Eds., Pergamon Press, New York, 1962, стр. 121—133.
40. von Neumann J., 1948, Proc. Hixon Symposia, L. Jeffrees, Ed., Wiley, New York, 1951, стр. 1—41.
41. von Neumann J., Lectures delivered at California Institute of Technology, 1952.
42. Фон Нейман Дж., Автоматы, Сб. статей под ред. К. Э. Шеннона и Дж. Маккарти, М., 1956.
43. Whitehead A. N. and Russell B. A. W., Principia Mathematica, Cambridge, 1910—1913.
44. Wiener N., Conference on Teleological Mechanisms, Annals N. Y. Acad. Sci. 50, 197—220 (1946).
45. Винер Н., Кибернетика, или Управление и связь в животном и машине, М., «Сов. радио», 1958.
46. Winograd S., IBM J. Res. Develop. 6, 430—436 (1962).
47. Winograd S., Information and Control, VI, 177—194 (1963).

Цена 37 коп.