

Р. Ландауэр

IEEE fellow and member of the IBM Thomas J. Watson
Research Center in Yorktown Heights, N.Y., USA.

НЕОБРАТИМОСТЬ И ВЫДЕЛЕНИЕ ТЕПЛА В ПРОЦЕССЕ ВЫЧИСЛЕНИЙ

Приводятся аргументы в пользу того, что любые счетные машины неизбежно вычисляют логические функции, которые не имеют обратных. Такая логическая необратимость соответствует физической необратимости и требует выделения конечного количества тепла за машинный цикл, обычно порядка kT на каждую необратимую функцию. Эта диссипация способствует стандартизации сигналов и делает их независимыми от точной логической истории. Подробный обзор двух простых, но полезных моделей бистабильных устройств позволяет подробно изучить кинетику переключений и получить соотношение между скоростью вычислений и диссипацией энергии, а также оценить влияние ошибок, индуцируемых тепловыми флуктуациями.

1. Введение

Поиски более быстрых и компактных вычислительных схем непосредственно приводят к вопросу о том, каковы принципиальные физические ограничения на прогресс в этом направлении. На практике ограничения, вероятно, возникают из-за необходимости обеспечения доступа к каждому логическому элементу. Пока, однако, трудно понять, какие именно физические условия это накладывает на те степени свободы, которые переносят информацию. Существование такой компактной среды хранения, как генетическая, позволяет надеяться, что можно весьма далеко продвинуться в направлении все большей компактификации устройств, по крайней мере, если мы готовы пожертвовать на этом пути скоростью и произвольным доступом.

Однако, отвлекаясь от проблемы доступа, можно показать или по меньшей мере предположить с большой степенью уверенности, что обработка информации неизбежно сопровождается определенным минимальным выделением тепла. Вообще говоря, это и не удивительно. Вычисления, как все процессы, протекающие с конечной скоростью, должны приводить к некоторой диссипации. Тем не менее наши аргументы зна-

чительно более фундаментальны и приводят к выводу о минимальном выделении тепла вне зависимости от скорости процесса. Естественно, количество выделенного тепла на много порядков меньше, чем диссипация энергии в любом практически осуществимом устройстве. Принципиальным моментом тем не менее оказывается тот факт, что диссипация приводит к реальным следствиям, а не является просто досадной и устранимой неприятностью. На практике к тем же следствиям могут привести несравненно большие количества рассеянной энергии.

Результат нашего исследования о диссипации можно предсказать несколькими способами, и нашей основной целью будет сжатое изложение главных идей, которые помогут достичь более глубокого понимания физических требований к логическим устройствам. Простейший путь предсказания нашего вывода состоит в учете того, что бинарное устройство имеет по крайней мере одну информационную степень свободы. На классическом уровне одна степень свободы соответствует kT тепловой энергии. Любой переключающий сигнал, проходящий через данное устройство, должен, следовательно, обладать большей энергией для подавления шума. Этот довод не означает, что энергия сигнала действительно должна рассеиваться. В качестве альтернативного способа предсказания наших выводов может быть использована ссылка на мнения Бриллюэна и более ранних авторов, собранные Бриллюэном в его книге «Наука и теория информации» [1], о том, что процесс измерения требует рассеяния энергии порядка kT . Процесс вычислений, в котором состояния различных элементов зависят от состояний других элементов в предшествующие моменты времени, тесно связан с измерением. Сложно, однако, дать более точную характеристику такой связи. Кроме того, аргументы, касающиеся процесса измерения, основаны на анализе специфических моделей, которые весьма далеки от устройств, осуществляющих обработку данных. На самом деле, при изучении процесса измерений само понятие *измерения* не определяется достаточно четко, и, кроме того, обходится следующий чрезвычайно существенный вопрос: когда система A , взаимодействующая с системой B , осуществляет измерение? Сам по себе тот факт, что две физические системы взаимодействуют, не обязательно приводит к диссипации.

Наш главный аргумент возникает в результате следующей цепочки рассуждений. Простейшее бинарное устройство представляет собой частицу в бистабильной потенциальной яме, показанной на рис. 1. Назовем частицу, находящуюся в левой яме, состоянием **нуль**, а частицу в правой яме — состоянием **единица**. Рассмотрим теперь операцию **установки в**

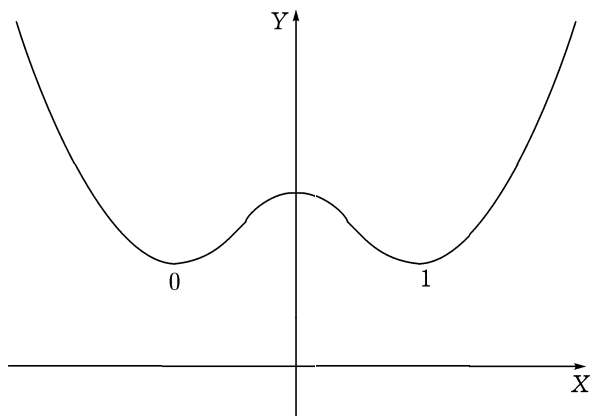


Рис. 1. Бистабильная потенциальная яма. X — обобщенная координата, по которой происходит переключение

единицу, которая переводит частицу в состояние **единицы**, вне зависимости от ее начального положения. Если мы говорим, что частица находится в состоянии **единицы**, то ее легко сохранить в этом состоянии без каких-либо затрат энергии. Если же частица находится в состоянии **нуль**, то необходимо приложить силу, чтобы она смогла перескочить барьер, а когда она пройдет максимум потенциала — обратную силу, чтобы частица, попав в состояние **единицы**, не имела избытка кинетической энергии. Таким образом, полный процесс не требует затрат энергии, которая восстанавливается благодаря увеличению энергии скатывающейся вниз частицы. Таким образом, на первый взгляд, возможно осуществить **установку в единицу** без потери энергии. Отметим, однако, что для предотвращения потерь энергии потребовалось использовать две различные программы в зависимости от начального состояния устройства. Компьютер работает совсем не так. В большинстве случаев способ перемещения информации компьютером не зависит от введенных данных и является функцией только физической реализации вычислительной схемы.

Можем ли мы найти меняющуюся со временем силу $F(t)$, которая, будучи приложенной к консервативной системе на рис. 1, переведет частицу в состояние **единицы**, если в начальный момент времени она находилась либо в **единице**, либо в **нуле**? Поскольку система консервативна, полное обращение времени приводит к другой системе, удовлетворяющей уравнениям движения. В этой системе имеется возможность того, что для определенного начального условия (частица находится в **единице**, ско-

рость равна нулю) результатом будут по меньшей мере два состояния: **единица** и **нуль**. Это, однако, невозможно. Законы механики полностью детерминированы, и траектория определяется начальными положением и скоростью. (Неустойчивое начальное положение является в некотором смысле исключением, т. к. из неустойчивой точки можно выйти по одному из по крайней мере двух направлений. Наша начальная точка **единица**, однако, является точкой устойчивого равновесия.) Возвращаясь к исходному направлению времени, можно сделать вывод, что невозможно однозначно построить силу $F(t)$, которая обеспечивает переход частицы в состояние **единицы** вне зависимости от ее начального положения.

Если, однако, допустить возможность потенциальной ямы с потерями, такое построение упрощается. Большая направленная вправо начальная сила, приложенная достаточно медленно для того, чтобы подавить колебания, вытолкнет частицу вправо, в состояние **единицы**, вне зависимости от ее начального состояния. Тогда, если сила убирается достаточно медленно, так что колебания опять подавлены, частица оказывается связанной в состоянии **единицы**. Этот пример иллюстрирует общее положение, обоснованное в [2] более подробно: хотя система с большим затуханием по очевидным причинам неприемлема, т. к. она чрезвычайно медлительна, система со слишком малым затуханием также нежелательна для реализации переключения, поскольку система может перескочить обратно в неправильное состояние, когда переключающая сила прикладывается или убирается слишком быстро.

2. Классификация

Прежде чем начать более подробное обсуждение, необходимо классифицировать устройства для обработки данных по способам хранения информации, когда она ни с чем не взаимодействует и не обрабатывается. Простейший класс, который будет рассматриваться в дальнейшем, включает устройства, имеющие возможность хранить информацию без рассеяния энергии. Система, показанная на рис. 1, относится к этому классу. Близкое отношение к механическому примеру на рис. 1 имеют ферриты, ферроэлектрики и тонкие магнитные пленки. Последние могут переключаться без передвижения доменной стенки и весьма тесно связаны с устройством, изображенным на рис. 1. Криотроны также являются устройствами, диссипация в которых имеет место только при переключении. Они отличаются, однако, от устройств типа, изображенных на рис. 1, поскольку в них **единица** и **нуль** не являются энергетически

предпочтительными. Криотрон похож скорее на механическое устройство с рис. 2, изображающее частицу в ящике. Два конкретных положения в ящике выбираются в качестве состояний **нуль** и **единица**, и сохранение информации зависит от того факта, что броуновское движение в ящике очень медленное. Упор на медленность броуновского движения, а не на сохраняющие силы, характеризует не только криотроны, но и большинство наиболее известных способов хранения информации. Из литературы известно, что все существенные логические функции могут быть реализованы устройствами первого класса. Это означает, что можно построить компьютеры, содержащие только криотроны или только магнитные ядра [3, 4].

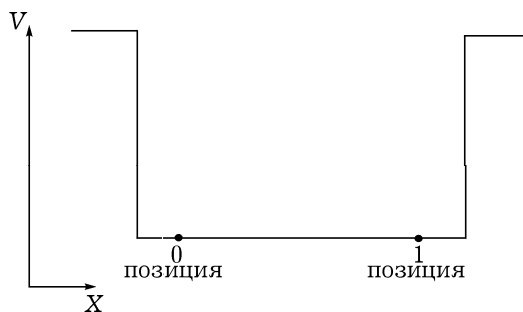


Рис. 2. Потенциальная яма, в которой состояния 0 и 1 не разделены барьером. Информация сохраняется, так как броуновское движение медленно

Второй класс устройств включает структуры, находящиеся в стабильном (инвариантном по времени) состоянии, которое оказывается диссипативным, когда дело идет о хранении информации. К этому классу принадлежат триггерные электронные схемы, реле и туннельные диоды. Последние обнаруживают типичное поведение, проиллюстрированное на рис. 3. Две стабильные точки разделены нестабильной областью так же, как и для устройства на рис. 1. Интересно, что этот класс не имеет известных представителей, аналогичных рис. 2. Все активные бистабильные устройства обладают встроенными средствами для восстановления нужного состояния. Сходство между рис. 3 и устройством на рис. 1 станет более очевидным, если представить бистабильную яму на рис. 1 графиком зависимости силы от расстояния (см. рис. 4). Линия $F = 0$ пересекает кривую в трех точках, аналогично сплошной линии (линии постоянного тока) на рис. 3. Такая аналогия наводит на мысль о том, что в случае диссипативного устройства в результате теплового возбуждения или квантово-

механического туннелирования возможны переходы между различными стабильными состояниями, так же, как и в недиссипативном случае, что

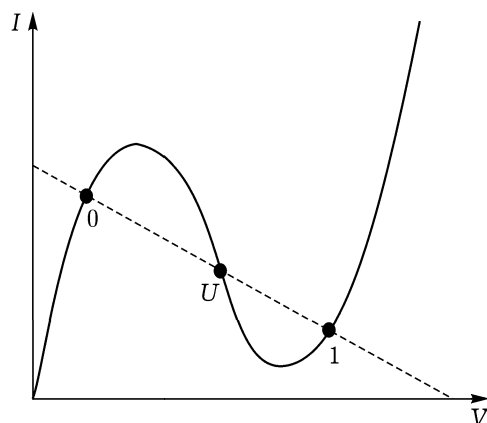


Рис. 3. Характеристика отрицательного сопротивления (сплошная линия) и линия нагрузки. 0 и 1 — стабильные состояния, U — нестабильное

подробно обсуждается Свенсоном [5]. Диссипативное устройство, такое, как туннельный диод, строго говоря, является аналогом несимметричной потенциальной ямы, а не симметричной ямы с рис. 1. Следовательно, можно ожидать, что из двух возможных состояний устройства с отрицательным сопротивлением только одно действительно стабильно, другое же — метастабильно. Ансамбль бистабильных туннельных диодов, представленный самому себе на достаточно длительное время, будет эволюционировать к одному и тому же абсолютно стабильному состоянию.

В общем случае, когда такие запирающие устройства используются в вычислительных схемах, понижение диссипации в двух допустимых состояниях может быть достигнуто перемещением этих состояний как можно ближе к оси токов или напряжений. Если бы удалось почти полностью исключить диссипацию в стационарном состоянии, такое устройство можно было бы отнести к первому классу. Следовательно, интуитивно можно ожидать, что в стационарном состоянии диссипативного устройства диссипация на одно переключение по крайней мере так же велика, как и в устройствах первого класса, и что эта диссипация, связанная с переключением, дополняется диссипацией в стационарном состоянии.

Последним, третьим классом являются устройства типа «catch-all», а именно, такие, в которых для распознавания информации существенно изменение во времени. Этот класс включает линии задержки, а также несущие схемы, например, фазовые бистабильные системы фон Неймана [6].

Последние представляют собой изящный пример использования эффектов диссипации; большинство других представителей третьего класса более сложны для обсуждения в простых физических терминах.

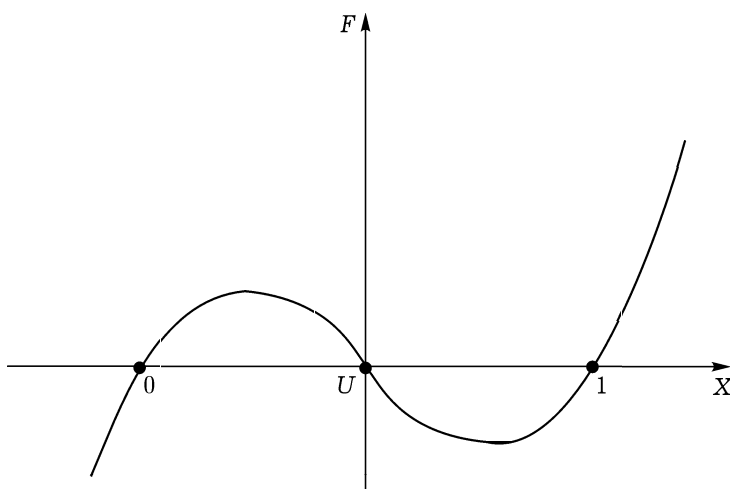


Рис. 4. Зависимость силы от расстояния для бистабильной ямы на рис. 1. 0 и 1 — стабильные состояния, U — нестабильное

В схеме фон Неймана, которую мы рассмотрим здесь не слишком подробно, используется сигнал накачки с частотой ω_0 , который, воздействуя на систему, настроенную на $\omega_0/2$ и обладающую нелинейной реактивностью, вызывает спонтанный переход сигнала на меньшую частоту. Низкочастотный сигнал может выбирать между двумя фазами (различающимися на 180° на низкой частоте), что и является причиной бистабильности. В схеме фон Неймана накачка прекращается после появления субгармоники, и субгармоника, соответственно, может пропадать в результате неизбежных потерь в схеме. Этот процесс представляет собой существенную часть схемы, контролируя направление прохождения информации, так что, на первый взгляд, потери здесь играют важную роль. Можно показать, однако, что редукция сигнала может быть осуществлена в нелинейной схеме без потерь для нужным образом фазированного сигнала накачки. Таким образом, представляется адекватным использование нелинейных схем без потерь, и вместо выключения накачки менять фазу накачки, что приведет к распаду сигнала вместо роста. При этом направленность потока информации не зависит от наличия потерь в схеме. Потери выполняют, однако, другую существенную функцию.

Система фон Неймана сильно зависит от схемы взаимодействия, на-

зываемой *majority logic*, в которой имеет место связь с тремя субгармоническими осцилляторами, а сумма их колебаний используется для синхронизации субгармонического осциллятора, накачка для которого включается несколько позже первых трех.

Каждый из трех складываемых сигналов может иметь одну из двух возможных фаз. Пропасть могут самое большее два из них, а один всегда останется, таким образом, всегда сохраняется фаза, предназначенная для возбуждения очередного колебания. Синхронизирующий сигнал может, следовательно, иметь две возможные амплитуды. Если все три входных сигнала согласованны, то синхронизирующий сигнал в три раза сильнее, чем в случае, когда только два сигнала имеют заданную фазу. Если в субгармонической схеме отсутствуют потери, последующая настройка приведет к двум различным амплитудам в зависимости от силы исходного синхронизирующего сигнала. Это, однако, будет интерферировать с основной операцией схемы на следующем этапе, когда потребуется вновь просуммировать выходные сигналы трех осцилляторов, которые должны будут иметь равные амплитуды. Станет понятно, что отсутствие потерь дает выходные амплитуды от каждого осциллятора, также зависящие от входных сигналов на предыдущем этапе. В то же время, хотя отклонения от ожидаемых амплитуд не имеют особого значения в одном цикле, они могут увеличиться за время, равное нескольким машинным циклам. Следовательно, потери необходимы для того, чтобы уничтожались несущественные детали предыстории сигнала. Потери имеют принципиальное значение для стандартизации сигналов, эта функция долгое время не получала адекватного признания в теоретических дискуссиях, но весьма четко обозначена в недавней статье А. У. Ло [7].

3. Логическая необратимость

Во введении при анализе команды **сохранение в один** был изучен рис. 1 и установлена необходимость диссипации энергии. Теперь мы предпримем попытку обобщения такого способа рассуждений. **Сохранение в один** является примером логической функции истинности, которую мы будем называть *необратимой*. Назовем устройство *логически необратимым*, если по сигналу на выходе нельзя определить однозначно сигнал на входе. Мы считаем, что логически необратимые устройства имеют принципиальное значение для вычислительной техники. Мы считаем также, что логическая необратимость предполагает в свою очередь физическую

необратимость, а последняя сопровождается диссипативными эффектами.

Будем рассматривать компьютер как строго конечный массив из N двоичных элементов, способных сохранять информацию без диссипации. Будем считать компьютер синхронным, что означает, что существует хорошо определенный машинный цикл, и в конце каждого цикла состояние N элементов является сложной функцией их состояния в начале цикла.

Наш анализ логической необратимости будет осуществлен на трех различных уровнях. Аргументы первого уровня заключаются просто в утверждении о том, что существующие вычислительные устройства сильно зависят от логически необратимых действий, и, следовательно, любая машина, воспроизводящая логическую организацию существующих машин, будет обладать логической необратимостью, а значит, и физической необратимостью.

На втором уровне мы рассматриваем специальный класс компьютеров, использующих логические функции только одной или двух переменных. После выполнения машинного цикла состояние каждого из N элементов становится функцией состояния не более двух двоичных элементов до выполнения цикла. Допустим теперь, что компьютер логически обратим. Тогда машинный цикл отображает 2^N возможных начальных состояний на полное пространство 2^N состояний, а не только на некоторое подпространство. В 2^N возможных состояниях каждый бит имеет равную вероятность оказаться в состоянии **единицы** и в состоянии **ноль**. Следовательно, обратимый компьютер может использовать только те функции, таблицы истинности которых содержат равное число состояний **единица** и **ноль**. Таким образом, допустимыми функциями являются тождество и отрицание, а также **исключающее или** и его отрицание. Однако эти функции не образуют полного набора [8] и не позволяют образовать все остальные функции истинности.

На третьем уровне мы рассматриваем более общие устройства. Например, устройство с тремя входами и тремя выходами, т. е. небольшой трехбитный компьютер. Пусть p , q и r — переменные до выполнения машинного цикла. Рассмотрим конкретную функцию истинности, которая заменяет r на $p \cdot q$, если $r = 0$, и заменяет r на $\overline{p \cdot q}$, если $r = 1$. Во время машинного цикла переменные p и q не изменяются. Можно сказать, что r определяет выбор программы, а p и q являются переменными, на которых выполняется выбранная программа. Это устройство логически обратимо, поскольку результат на выходе однозначно определяет входные данные.

Тем не менее, оно пригодно для реализации такой операции как **и**, которая сама по себе не является логически обратимой. Но компьютер сохраняет достаточно входной информации для обеспечения обратимости. Интересно отметить, однако, что программа на самом деле не «сохраняется»; какая она была, можно понять только дедуктивным путем.

Теперь рассмотрим компьютер, предназначенный для более общих целей, который при выполнении программы проходит, как правило, через множество машинных циклов. На первый взгляд может показаться, что логическую обратимость легко получить, просто сохраняя входные данные в некоторой части машины. Однако мы будем называть машину логически обратимой тогда и только тогда, когда обратимым является каждое ее действие. Это означает, что каждый раз, когда вычисляется функция истинности двух переменных, необходимо сохранять дополнительную информацию о величинах, над которыми производятся действия, вне зависимости от того, нужна эта информация или нет. Стирание, эквивалентное **установке в единицу**, обсуждавшемуся в введении, не допускается. Потребуется, следовательно, длинная программа, которая приведет в полный беспорядок положения битов несущественной информацией о промежуточных результатах. Более того, если потребуется использовать обратимую функцию трех переменных, такую, как **и**, необходимо будет дополнить исходную программу отдельным состоянием **нуль** для каждой операции **и**, поскольку «отклонение», программирующее машину, не сохраняется в процессе выполнения **и**. Это значит, что машина должна иметь гигантские возможности для хранения как дополнительных битов с «отклонениями», так и дополнительных битов с выходными данными. Можно ли обеспечить такие возможности для реализации обратимости всех промежуточных действий? Если наша машина способна в обычном смысле выполнять непрерывную программу, то понятно, что сохранение всей информации о промежуточных действиях невозможно.

Не будем, однако, соблазняться таким простым выходом из положения. Вполне вероятно, что удастся построить машину, полезную в общепринятом смысле, но не имеющую возможности выполнять непрерывную программу. Допустим, что эта машина использует логически необратимые функции истинности. Необратимые функции могут быть превращены в обратимые, как уже было проиллюстрировано, путем «погружения» их в функции большего числа переменных. Более крупные функции требуют, однако, дополнительных входов для управления и дополнительных выходов для сохранения информации, обеспечивающей необратимость. Мы утверждаем теперь, что, хотя более крупная машина и обратима, она

не может служить полезным вычислительным устройством в обычном смысле этого слова.

Прежде всего, чтобы обеспечить пространство для сохранения дополнительных входных и выходных данных, требуется знать, сколько раз необходимо выполнить каждую операцию исходной (необратимой) машины. Полезность компьютера, однако, заключается в том, что он является несколько большим, чем настольным справочным устройством; он может также выполнять множество программ, которые не предусматривались при его разработке. Такая машина с расширенными функциями должна включать в себя некоторое количество положений битов для каждого встроенного устройства порядка числа шагов программы и требует обеспечения числа переключений во время загрузки программы, сравнимого с числом переключений в самой программе. Настройка управления во время загрузки программы, которая обычно заключается в определении длинного ряда битов как состояния **нуль**, является как раз одной из разновидностей необратимой логической операции, чего мы пытались избежать. Таким образом, наша громоздкая машина не использует необратимых операций во время выполнения программы только за счет дополнительной необратимости на этапе загрузки.

4. Логическая необратимость и рождение энтропии

Остается подробно изучить связь между логической необратимостью и изменениями энтропии. Например, снова рассмотрим операцию **установки в единицу**. Обобщение рассуждений на более сложные логические операции будет тривиальным.

Допустим сначала, что операция **установки в единицу** уже выполнена на каждом из битов ансамбля. Эта ситуация напоминает ансамбль спинов, направленных вдоль положительной оси z . В тепловом равновесии биты (или спины) могут находиться в двух равновероятных состояниях. Наши специально приготовленные системы обнаруживают гораздо больше упорядоченности, а следовательно, более низкие по сравнению с характерными для состояния равновесия температуру и энтропию. При адиабатическом размагничивании используется как раз такое спиновое состояние, и в процессе дезориентации спинов система получает энтропию от окружающей среды и решетка, в узлах которой находятся спины, размораживается. Ансамбль упорядоченных битов ведет себя аналогично. По мере повышения температуры и забывания начального состояния окружающая

среда размораживается. Отметим, что наиболее важным моментом здесь является не то, что в начальном состоянии все биты ансамбля согласованны, а только то, что существует единственное, хорошо определенное начальное состояние для набора битов. Хорошо определенное начальное состояние соответствует, по обычному определению энтропии $S = k \ln W$, нулевой энтропии. Информационные степени свободы могут посредством тепловой релаксации перейти в любое из 2^N состояний (для N -битового ансамбля), при этом энтропия увеличится до $kN \ln 2$.

Отметим, что наши рассуждения не обязательно зависят от часто упоминаемой связи между энтропией и информацией. Мы просто рассматриваем каждый бит как локализованный в физической системе, возможно, со многими степенями свободы, помимо информационной. Однако для каждого возможного физического состояния, которое будет интерпретироваться как **нуль**, имеется практически идентичное возможное физическое состояние, в котором физическая система представляет **единицу**. Следовательно, для системы в состоянии **единицы** доступно в два раза меньше состояний, чем для системы, которая может находиться в **единице** или в **нуле**. (Мы будем игнорировать случай, когда **единица** и **нуль** представлены состояниями с различной энтропией. Этот случай требует более сложной аргументации, но приводит к таким же результатам.)

При выполнении операции **установки в единицу** происходит обратный процесс. В начале каждый бит находится в одном из двух состояний, а в конце имеет место хорошо определенное состояние. Изучим эту операцию подробнее.

Рассмотрим статистический ансамбль состоящий из битов, находящихся в тепловом равновесии. Если все биты переводятся в **единицу**, то число состояний, занимаемых ансамблем сокращается вдвое. Энтропия, следовательно, уменьшается на $k \ln 2 = 0.6931k$ на бит. Энтропия замкнутой системы, т. е. компьютера с собственными источниками энергии, не может уменьшаться; следовательно, она должна проявиться в качестве эффекта нагревания где-нибудь еще, восполняя $0.6931kT$ на бит. Конечно, это минимально возможное нагревание, и наши рассуждения не гарантируют, что этот минимум реально достигим.

Рассмотренная нами операция установки применялась к ансамблю в тепловом равновесии. В действительности хотелось бы понять, что происходит в конкретной вычислительной схеме, работающей с информацией, которая еще не нагрелась, но состоит из хорошо определенных состояний **единица** или **нуль**. Пусть для начала операция установки выполняется на случайной цепи, состоящей из **единиц** и **нулей**. Можно, как обычно,

использовать эквивалентность статистического ансамбля усреднению по времени и сделать вывод о том, что диссипация на одну операцию установки для временной последовательности равна диссипации для нагретого ансамбля.

Компьютер, однако, вряд ли работает со случайными данными. Одно из двух состояний бита может появляться чаще другого или даже, если эти вероятности совпадают, может иметь место корреляция между соседними битами. Другими словами, установленные значения не отражают максимально возможного объема информации. Рассмотрим крайний случай, когда на входе установлены только **единицы** и не требуется выполнять никаких операций. Очевидно, что не происходит изменений энтропии и отсутствует диссипация. И наоборот, если на все начальные состояния — **нули**, то они не несут информации и их установка в **единицы** не приводит к изменениям энтропии. Отметим, однако, что операция установки, используемая когда на входе все **единицы** (ничего не делать), не может применяться в том случае, когда на входе все **нули**, а мы хотим перевести их в **единицы**. Теперь положение аналогично фазовому преобразованию между двумя фазами в равновесии, и процесс может быть сделан обратимым без увеличения энтропии вселенной, достаточно только изобрести специально для этой задачи некоторую процедуру установки. Таким образом, в случае, когда начальные состояния не обладают максимально возможным разнообразием, неизбежное повышение энтропии в операции **установки** может быть нейтрализовано, но только с использованием наших знаний о входных данных и соответствующей реализации этой операции.

Обобщение на другие логически необратимые операции очевидно и будет проиллюстрировано только одним примером. Рассмотрим совсем небольшой компьютер с тремя двоичными элементами p , q и r . Машинный цикл заменяет p на r , q на r и r на $p \cdot q$. Всего имеется восемь возможных состояний, и в тепловом равновесии они появляются с равной вероятностью. На сколько может уменьшиться энтропия за машинный цикл? Начальные и конечные состояния машины показаны на рис. 5. Состояния α и β появляются с вероятностью $1/8$ каждое, состояния γ и δ — с вероятностью $3/8$ каждое. Начальная энтропия была равна

$$S_i = k \ln W = -k \sum \rho \ln \rho = -k \sum \frac{1}{8} \ln \frac{1}{8} = 3k \ln 2.$$

Конечная энтропия есть

$$S_f = -k \sum \rho \ln \rho = -k \left(\frac{1}{8} \ln \frac{1}{8} + \frac{1}{8} \ln \frac{1}{8} \right) + \frac{3}{8} \ln \frac{3}{8} + \frac{3}{8} \ln \frac{3}{8}.$$

Разность $S_i - S_f$ равна $1.18k$. Следовательно, минимальная диссипация для начального состояния, не несущего полезной информации, равна $1.18kT$.

перед циклом				после цикла			конечное состояние
p	q	r		p_1	q_1	r_1	
1	1	1	→	1	1	1	α
1	1	0	→	0	0	1	β
1	0	1	→	1	1	0	γ
1	0	0	→	0	0	0	δ
0	1	1	→	1	1	0	γ
0	1	0	→	0	0	0	δ
0	0	1	→	1	1	0	γ
0	0	0	→	0	0	0	δ

Рис. 5. Устройство с 3 входами и 3 выходами, которое 8 возможных состояний отображает на 4 различных состояния

Возникает вопрос, действительно ли энтропия снижается логически необратимой операцией. Если мы на самом деле отображаем возможные начальные состояния **нуль** и возможные начальные состояния **единица** на одно и то же пространство, т. е. пространство состояний **единица**, никакого вопроса нет. Но возможно, что после выполнения операции может остаться некоторая небольшая разница между системой, которая уже была в состоянии **единицы**, и той, которая должна быть переведена в нее. Если такая разница существует в течение некоторого промежутка времени, в этом нет большого вреда, но как было показано в обсуждении недиссипативного субгармонического осциллятора, нельзя проигнорировать кумулятивный процесс, во время которого различия между всевозможными состояниями **единица** становятся все больше и больше в соответствии с их предысторией. Следовательно, физическое отображение «много в одно», которое и является источником изменений энтропии, не должно осуществляться во всех деталях во время машинного цикла, выполняющего логическую функцию. Но оно обязательно должно иметь место, и это все, что относится к исследованию выделения тепла.

5. Подробный анализ бистабильной ямы

В дополнение к предварительному общему рассмотрению дадим теперь более подробный анализ переключения в системе, представленной

в виде бистабильной потенциальной ямы, как показано на одномерном рис. 1, с барьером, большим по сравнению с kT . Допустим, кроме того, что переключение осуществляется под действием дополнительной силы, увеличивающей энергию одной ямы по сравнению с другой, но сохраняющей при этом барьер, который должен быть преодолен тепловым возбуждением. (Достаточно большая сила может просто полностью уничтожить один из минимумов. Поэтому переключающие силы будут считаться малыми.) Рассмотрим теперь статистический ансамбль систем с двойными ямами с неравновесным распределением, и зададимся вопросом, как быстро будет достигнуто равновесие. Этот вопрос детально исследован в статье [2], и здесь мы ограничимся весьма простым кинетическим анализом, который приводит к тому же ответу. Пусть n_A и n_B — числа членов ансамбля в ямах A и B соответственно. U_A U_B — энергии минимумов каждой ямы, а U — высота барьера между ними. Тогда скорость, с которой частицы переходят из ямы A в яму B будет иметь вид $\nu n_A \exp[-(U - U_A)/kT]$. Поток из B в яму A равен тогда $\nu n_B \exp[-(U - U_B)/kT]$. Два множителя, определяющих частоты, принимаются равными, поскольку их разность несущественна по сравнению с разностью экспонент. Имеем:

$$\begin{aligned} \frac{dn_A}{dt} &= -\nu n_A \exp[-(U - U_A)/kT] + \nu n_B \exp[-(U - U_B)/kT], \\ \frac{dn_B}{dt} &= \nu n_A \exp[-(U - U_A)/kT] - \nu n_B \exp[-(U - U_B)/kT]. \end{aligned} \quad (5.1)$$

Можно рассматривать уравнения (5.1) как линейное преобразование (n_A, n_B) в $\left(\frac{dn_A}{dt}, \frac{dn_B}{dt}\right)$. Характеристические числа этого преобразования суть

$$\lambda_1 = 0, \quad \lambda_2 = -\nu \exp[(U - U_A)/kT] - \nu \exp[-(U - U_B)/kT].$$

Собственное значение $\lambda_1 = 0$ соответствует не зависящей от времени заселенности ямы. Это равновесное распределение

$$n_A = n_B \exp \frac{U_B - U_A}{kT}.$$

Остающееся отрицательное собственное значение должно быть поставлено в соответствие отклонениям от равновесия, а $\exp(-\lambda_2 t)$ определяет скорость исчезновения этих отклонений. Время релаксации τ выражается, следовательно, через величину U_0 , которая является средним U_A

и U_B

$$\frac{1}{\tau} = \lambda_2 = \nu \exp[-(U - U_0)/kT] \times \\ \times \{ \exp[-(U_0 - U_A)/kT] + \exp[-(U_0 - U_B)/kT] \}. \quad (5.2)$$

Средняя энергия U_0 в уравнении (5.2) сокращается, и уравнение (5.2) выполняется вне зависимости от определения U_0 . Обозначая $\Delta = \frac{1}{2}(U_A - U_B)$, получаем

$$\frac{1}{\tau} = 2\nu \exp[-(U - U_0)/kT] \operatorname{ch}(\Delta/kT). \quad (5.3)$$

В первом порядке по переключающей силе, обуславливающей различие между U_A и U_B , $(U - U_0)$ не изменяется, и уравнение (5.3) можно переписать как

$$\frac{1}{\tau} = \frac{1}{\tau_0} \operatorname{ch}(\Delta/kT), \quad (5.4)$$

где τ_0 — время релаксации для симметричной потенциальной ямы, когда $\Delta = 0$. Это уравнение демонстрирует полезность данного устройства. Время релаксации τ_0 характеризует промежуток времени, необходимый для нагревания бистабильного устройства, и представляет собой максимальное время, через которое с устройством можно начать работать. С другой стороны, τ является минимальным временем переключения. Таким образом, $\operatorname{ch}(\Delta/kT)$ представляет максимальное число переключений за время жизни информации. Оно может быть большим, следовательно, устройство может быть полезным. Даже если Δ достаточно велико для нарушения приближения первого порядка, в котором $(U - U_0)$ — константа, экспоненциальная зависимость $\operatorname{ch}(\Delta/kT)$ от Δ в уравнении (5.3) перевесит изменения в $\exp[(U - U_0)/kT]$, и τ_0/τ останется быстро растущей функцией Δ .

Отметим, что Δ равна половине энергии, рассеивающейся в процессе переключения. Тепловые распределения вероятностей в каждой яме будут практически одинаковыми до и после переключения, единственная разница в том, что конечная яма на 2Δ ниже исходной. Эта разница в энергиях рассеивается и соответствует энергии для половины площади петли гистерезиса энергии, обычно связывающейся с переключением. Уравнение (5.4) подтверждает, таким образом, хорошо известный эмпирически факт, что увеличение скорости переключения может быть достигнуто только ценой повышения диссипации на одно переключение. Уравнение (5.4) выполняется, однако, только для специального класса моделей и не может применяться в общем случае. Чтобы показать это, рассмотрим альтернативную модель. Допустим, что информация сохраняется положением частицы на

прямой и что $x = \pm a$ соответствуют **нулю** и **единице**. Не предполагается существование никакого барьера, но случайное диффузионное движение частиц принимается достаточно медленным, так что положения будут сохраняться на приемлемо продолжительном промежутке времени. (Эта модель, вероятно, имеет больше общего с поведением ферритов и ферроэлектриков, когда переключение связано с движением доменной стенки, чем предыдущая модель с бистабильной ямой. Разница в энергиях между полностью и частично переключенными ферритами чрезвычайно мала, и имеет место низкая мобильность доменной стенки, удерживающей частицу вблизи ее начального состояния в отсутствие переключающих сил, а это начальное состояние может с равным успехом быть как частично, так и полностью переключенным.) В этом случае частицы диффундируют на типичное расстояние s за время $\tau \sim s^2/2D$, где D — постоянная диффузии. Расстояние, соответствующее потере информации, есть $s \sim a$, время релаксации — $\tau_0 \sim a^2/2D$. В присутствии силы F частица движется со скоростью μF , где мобильность μ определяется отношением Эйнштейна как D/kT . Чтобы передвинуть частицу переключающей силой F на расстояние $2a$, требуется время τ_s :

$$\mu F \tau_s = 2a, \quad (5.5)$$

или

$$\tau_s = 2a/\mu F. \quad (5.6)$$

Диссипация энергии 2Δ есть $2aF$. Это следует из уравнений:

$$\tau_s = 2a^2/\mu\Delta, \quad (5.7)$$

$$\tau_s/\tau_0 = 4kT/\Delta, \quad (5.8)$$

которые демонстрируют такой же характер зависимости τ_s от Δ как и для случая с барьером, но не включает экспоненциальное изменение с показателем Δ/kT . Если не принимать во внимание остальные аргументы, становится ясно, что бистабильный элемент энергии уравнения (5.4) нужно предпочесть диффузионно стабилизированному элементу уравнения (5.8).

Приведенные выше примеры позволяют глубже понять необходимость диссипации энергии, неявно подтверждающуюся рассуждениями, использующими свойства энтропии. В операции **сохранения в единицу** требовалось установить систему в состояние **единицы** вне зависимости от ее начального состояния. Это делалось посредством понижения энергии состояния **единица** относительно состояния **нуль**. Частица переходит в

состояние с низшей энергией, рассеивая при этом избыток энергии, имевшийся с начальном состоянии.

6. Три источника ошибок

В этой главе мы рассмотрим, в тесной связи с предыдущими рассуждениями, относительную значимость нескольких возможных источников ошибок в процессе вычислений. Прежде всего, реальное время, отпускаемое на переключение, конечно, и релаксация в нужное состояние не может завершиться полностью. Если T_s — реальное время, в течение которого приложена переключающая сила, и τ_s — время релаксации из уравнения (5.4), то $\exp(-T_s/\tau_s)$ есть вероятность того, что переключения не произойдет. Второй источник ошибок подробно рассмотрен в статье Дж. А. Свенсона [5] и заключается в том, что τ_0 конечно, и информация будет исчезать, хотя предполагается, что она спокойно хранится в начальном состоянии. Относительная важность указанных ошибок является предметом компромиссов при разработке прибора. Время T_s , отпущенное на переключение, всегда может быть увеличено, что сделает релаксацию переключения более полной. Однако полное время, доступное для программы, меньше, чем τ_0 — времени релаксации для хранимой информации, т. е. увеличение времени переключения понижает число шагов в максимально возможной программе.

Третий источник ошибок заключается в том, что даже если система может полностью релаксировать за время переключения, обязательно останется часть ансамбля порядка $\exp(-2\Delta/kT)$, оставшаяся в нежелательном начальном состоянии. Для использования в дальнейших рассуждениях назовем эту ошибку больцмановской. Мы покажем, что не имеет значения, какой бы компромисс между двумя первыми видами ошибок ни был достигнут при разработке, больцмановская ошибка не будет доминировать. Проведем грубое сравнение ошибок, не рассматривая различия в истории получения информации.

Дадим верхнюю оценку больцмановской ошибки, считая, что переключение имело место в каждом машинном цикле предыстории каждого бита. Это и будет оценкой сверху для больцмановской ошибки, которая, как мы покажем, пренебрежимо мала по сравнению с другими ошибками. Вероятность больцмановской ошибки на одно переключение есть $\exp(-2\Delta/kT)$. За время переключения непереключенные биты возвращаются обратно со скоростью $\exp(-t/\tau_0)$. Следовательно, за время T_s непереключенные биты имеют вероятность T_s/τ_0 потерять информацию.

Если больцмановская ошибка доминирует, то

$$T_s/\tau_0 < \exp(-2\Delta/kT). \quad (6.1)$$

Рассмотрим для определенности бистабильную яму уравнения (5.4). Тогда (6.1) принимает вид

$$\frac{2T_s}{\tau_s} \exp(-\Delta/kT) < \exp(-2\Delta/kT), \quad (6.2)$$

или, что эквивалентно

$$\frac{T_s}{\tau_s} < \frac{1}{2} \exp(-\Delta/kT). \quad (6.3)$$

Рассмотрим теперь релаксацию переключенного состояния. Ошибка за счет неполной релаксации есть $\exp(-T_s/\tau_s)$, что, согласно уравнению (6.3), удовлетворяет

$$\exp(-T_s/\tau_s) > \exp[-\frac{1}{2} \exp(-\Delta/kT)]. \quad (6.4)$$

Аргумент правой стороны этого неравенства $\frac{1}{2} \exp(-\Delta/kT)$ меньше $\frac{1}{2}$. Следовательно, правая сторона велика по сравнению с $\exp(-2\Delta/kT)$, больцмановской ошибкой, экспонента которой определено больше единицы. Мы показали, таким образом, что, если больцмановская ошибка доминирует над распадом информации, она в свою очередь должна подавляться неполной релаксацией при переключении.

Другой возможный путь к этому выводу состоит в демонстрации того, что накопленная больцмановская ошибка, с учетом максимального числа переключений, разрешенного уравнением (5.4), мала по сравнению с единицей.

Рассмотрим теперь диффузионно стабилизированный элемент уравнения (5.8). Вместо уравнения (5.4) находим для него соотношение

$$\exp(-T_s/\tau_s) > \exp[(-\Delta/4kT) \exp(-2\Delta/kT)], \quad (6.5)$$

где правая сторона снова больше больцмановской ошибки — $\exp(-2\Delta/kT)$. В этом случае также можно использовать альтернативный подход, основанный на накоплении ошибки.

Если мы будем рассматривать более реалистичные модели машин, в которых переключающие силы приложены к связанным устройствам, как сделано, например, для бездиодной логики магнитного ядра [4], окажется, что провести четкое аналитическое исследование типов ошибок гораздо труднее. Тем не менее мы убеждены, что их классификация будет подобна вышеприведенной.

7. Заключение

Информационные степени свободы компьютера взаимодействуют с тепловым резервуаром, представленном в данном случае остальными степенями свободы. Это взаимодействие играет две важных роли. Во-первых, оно действует как накопитель энергии, рассеиваемой в процессе вычислений. Такая диссипация имеет непреодолимый минимум, что является следствием выполнения компьютером необратимых операций. Во-вторых, взаимодействие становится источником шума, приводящего к возникновению ошибок. В частности, тепловые флуктуации дают предположительно переключенному элементу небольшую вероятность остаться в начальном состоянии, даже если переключающая сила действовала в течение достаточно длительного времени. На примере двух простых моделей было показано, что этот источник ошибок подавлен одним из двух других источников:

1) Неполным переключением, связанным с недостаточно продолжительным промежутком времени, отпущенным на переключение.

2) Распадом информации, вызванным тепловыми флуктуациями.

Разумеется, очевидно, что тепловой шум и требования к диссипации энергии находятся на таких масштабах, что могут не приниматься во внимание при рассмотрении современной вычислительной техники. Рассчитанная диссипация, однако, является абсолютным минимумом возможной величины. Реальные устройства, работающие на высоких скоростях и размеры которых далеки от минимальных, требуют, вероятно, значительно большей диссипации энергии для обеспечения стирания несущественных деталей предыстории вычислений.

8. Благодарности

Некоторые из обсуждавшихся вопросов были впервые поставлены Е. Р. Пиоре несколько лет назад. На ранних стадиях [2, 5] этот проект продвигался в основном усилиями покойного Джона Свенсона. Общение с Гордоном Лешером было весьма существенным для развития идей, представленных в данной статье.

Список литературы

- [1] L. Brillouin, *Science and Information Theory*, Academic Press Inc., New York, 1956.
- [2] R. Landauer and J. A. Swanson, *Phys. Rev.* 121 (1961) 1668.
- [3] K. Mendelssohn, *Progress in Cryogenics*, Vol. 1, Academic Press. Inc., New York, 1959. Chapter I by D. R. Young, p. 1.
- [4] L. B. Russell, *IRE Convention Record*, p. 106 (1957).
- [5] J. A. Swanson, *IBM Journal*, 4 (1960) 305.

Хотелось бы воспользоваться возможностью выделить два утверждения из статьи Свенсона, на которые не было обращено должного внимания в опубликованной версии.

1) Большое число частиц в оптимальном элементе (~ 100) является следствием низких энергий на одну частицу (или клетку), задействованную в типичном кооперативном эффекте, используемом в компьютере. Нет проблем в хранении информации в положении единственной частицы при комнатной температуре, если энергия активации ее движения достаточно велика (порядка нескольких эВ).

2) Оптимальный объем Свенсона в общем случае не слишком отличается от общепринятого требования для U , именно $vt \exp(-U/kT) \ll 1$, что может быть получено без использования теории информации. Это значит, что использование громоздких и усложненных методов кодирования не позволит увеличить объем хранимой информации. Очевидно, что желательно избавиться от этих усложнений, т. к. лишь небольшое увеличение каждого элемента по сравнению с «оптимальным» значением сделает его пригодным для переноса информации и без использования усложнений.

- [6] R. L. Wigginton, *Proceedings of the IRE*, 47 (1959) 516.
- [7] A. W. Lo, представлено в *IRE Transactions on Electronic Computers*.
- [8] D. Hilbert and W. Ackermann, *Principles of Mathematical Logic*, Chelsea Publishing Co., New York, p. 10 (1950).