Основы теории информации

Количество информации как мера уменьшения неопределённости знания

Подход к информации как к мере уменьшения неопределённости знания позволяет количественно измерять информацию.

Классический пример: монета, которую мы бросаем на ровную поверхность.

С равной вероятностью произойдёт одно из двух возможных событий – монета окажется в одном из двух положений: "орёл" или "решка". События равновероятны, так как при возрастающем числе опытов число выпадений "орла" и "решки" постепенно сближаются.

Перед броском существует **неопределённость нашего знания** (возможны 2 события), и как упадёт монета – в точности предсказать невозможно. После броска наступает **полная определённость**, т.к. мы видим, что монета в данный момент находится в определённом положении.

Это сообщение (событие) приводит **к уменьшению неопределённости нашего знания в 2 раза,** т.к. из двух возможных равновероятных событий реализовалось одно.

Основы теории информации

В окружающей действительности часто встречаются ситуации, когда может произойти число событий, большее чем два. Причем не всегда события равновероятны.

Очевидно, что чем больше начальное число возможных событий, тем больше начальная неопределённость нашего знания и тем большее количество информации будет содержать сообщение о результатах опыта.

Таким образом, математическая функция, связывающая количество возможных событий ***N*** и количество информации ***I*** , заключающееся в произошедшем событии, должна обладать тремя свойствами:
1. С ростом ***N*** функция должна монотонно возрастать.
2. При ***N*** равном 1, т.е. когда неопределенности знания нет (только одно возможное событие), эта функция равна 0 (количество информации равно нулю).
3. Если два независимых набора возможных событий ***N*** и ***M*** объедены в один набор, из которого могут одновременно реализоваться пара событий, то неопределенность знания объединенного набора равна сумме неопределенностей исходных наборов событий ***N*** и ***M.***

Существует только одна такая функция – логарифмическая.

Основы теории информации

Если все из возможных событий ***N равновероятны,*** то количество информации ***I*** выражается формулой:

$$I = \log N$$

Выбор основания логарифма диктуется тем, какую единицу измерения количества информации мы хотим использовать. Чаще всего применяется основание 2. При этом единица количества информации называется двоичной единицей или ***битом***, и представляет собой информацию, содержащуюся в одном событии из двух равновероятных событий (***N=2***).

Другими словами, ***бит*** – это такое количество информации, которое содержит сообщение, уменьшающее неопределённость знания в 2 раза.

Компьютер оперирует числами в двоичной системе счисления, поэтому в кратных единицах измерения количества информации используется коэффициент 2^n.

Основы теории информации

Задолго до появления компьютеров вообще, и тем более использующих двоичную систему счисления, в 1928 году американский инженер Ричард Хартли подметил закономерность и предложил меру для измерения количества информации:

$$I = \log_2 N$$

где ***N*** - количество равновероятных событий; ***I*** - количество бит в сообщении о том, что любое из ***N*** событий произошло.
Иногда формулу Хартли записывают так:

$$I = \log_2 N = \log_2 \left({}^{1}\!/\!_{p} \right) = \log_2 p^{-1} = -\log_2 p$$

т.к. каждое из ***N*** событий имеет равновероятный исход ***p = 1 / N***, то ***N = 1/p***.
На первый взгляд, кажется, что получается парадокс: количество информации выражается отрицательным числом. Однако если вспомнить, что вероятность – это величина в диапазоне от 0 до 1, то все становится понятным.

Задача.
Шарик находится в одной из трех урн: A, B или C. Определить, сколько бит информации содержит сообщение о том, что он находится в урне B.
Решение.
Такое сообщение содержит $I = \log_2 3 = 1{,}585$ бита информации.

Основы теории информации

Разновероятные события.

Не все ситуации имеют одинаковые вероятности реализации. Существует много таких ситуаций, у которых вероятности реализации различаются. Например, если бросают несимметричный предмет.

Еще один бытовой пример - "правило бутерброда".

"Однажды в детстве я уронил бутерброд. Глядя, как я виновато вытираю масляное пятно, оставшееся на полу, старший брат успокоил меня:

- Не горюй, это сработал закон бутерброда.
- Что еще за закон такой? - спросил я.
- Закон, который гласит: "Бутерброд всегда падает маслом вниз". Впрочем, это шутка, - продолжал брат.- Никакого закона нет. Просто бутерброд действительно ведет себя довольно странно: большей частью масло оказывается внизу.
- Давай-ка еще пару раз уроним бутерброд, проверим, - предложил я. - Все равно ведь его придется выкидывать.

Проверили. Из десяти раз бутерброд восемь раз упал маслом вниз.

И тут я задумался: а можно ли заранее узнать, как сейчас упадет бутерброд маслом вниз или вверх?

Наши опыты прервала мать…"

(Отрывок из книги "Секрет великих полководцев", В.Абчук).

Основы теории информации

В 1948 г. американский инженер и математик К. Шеннон предложил формулу для вычисления количества информации для событий с различными вероятностями.

Если ***I*** - количество информации,

N - количество возможных событий,

p_i - вероятности отдельных событий,

то количество информации для событий с различными вероятностями можно определить по формуле:

$$I = -\sum_{i=1}^{N} p_i \cdot \log_2 p_i,$$

где ***i*** принимает значения от 1 до ***N***.

Формулу Хартли теперь можно рассматривать как частный случай формулы Шеннона:

$$I = -\sum_{i=1}^{N} p_i \cdot \log_2 p_i = -\frac{1}{N}\sum_{i=1}^{N} \log_2\left(\frac{1}{N}\right) = -\frac{1}{N} \cdot N \cdot \log_2\left(N^{-1}\right) = \log_2 N$$

При равновероятных событиях получаемое количество информации максимально.

Основы теории информации

Задачи.

1. Определить количество информации, получаемое при реализации одного из событий, если бросают:

а) несимметричную четырехгранную пирамидку;

б) симметричную и однородную четырехгранную пирамидку.

Решение.

а) Будем бросать несимметричную четырехгранную пирамидку.

допустим, что грани пирамидки такие, что отношение их площадей можно представить пропорцией: 4 : 2: 1: 1. Тогда, вероятность отдельных событий будет такова:

$$p_1 = 1/2,\ \ p_2 = 1/4,\ p_3 = 1/8,\ \ p_4 = 1/8,$$

Вычислим по формуле Шеннона количество информации, получаемой после реализации одного из этих событий:

$$I = -(1/2*\log_2 1/2 + 1/4*\log_2 1/4 + 1/8*\log_2 1/8 + 1/8*\log_2 1/8) =$$
$$= 1/2 + 2/4 + 3/8 + 3/8 = 14/8 = 1{,}75 \text{ (бит)}.$$

б) Теперь рассчитаем количество информации, которое получится при бросании симметричной и однородной четырехгранной пирамидки:

$$I = \log_2 4 = 2 \text{ (бит)}.$$

Основы теории информации

Самостоятельно решить задачи:

1. Вероятность первого события составляет 0,5, а второго и третьего по 0,25. Какое количество информации мы получим после реализации одного из них?

2. Какое количество информации будет получено при игре в рулетку с 32-мя секторами?

3. Сколько различных чисел можно закодировать с помощью 8 бит?

Основы теории информации

Рассмотренный выше способ определения количества информации, получаемого в событиях, которые уменьшают неопределенность наших знаний, рассматривает информацию с позиции ее содержания, новизны и понятности для человека. С этой точки зрения, в опыте по бросанию кубика одинаковое количество информации содержится в сообщениях "два", "вверх выпала грань, на которой две точки" и в зрительном образе упавшего кубика:

одно из 6 равновероятных событий $I = \log_2 6 = \log_2 (3*2) = \log_2 2 + \log_2 3 = 1 + 1{,}585 = 2{,}585$ *бит*.

При передаче и хранении информации с помощью различных технических устройств информацию следует рассматривать как последовательность знаков (цифр, букв, кодов цветов точек изображения), не рассматривая ее содержание.

Основы теории информации

При таком подходе, термин «**событие»** ассоциируют с понятием **«сообщение»**.
Сообщение состоит из некоторого набора символов используемой знаковой системы – **алфавита.** Появление одного из символов алфавита в сообщении (**элементарное дискретное сообщение**) можно рассматривать как одно из состояний события. Если появление символов равновероятно, то можно рассчитать, сколько бит информации несет каждый символ. Информационная емкость знаков определяется их количеством в алфавите. Чем из большего количества символов состоит алфавит, тем большее количество информации несет один знак. **Полное число символов алфавита принято называть мощностью алфавита.**

Тогда бит – это количество информации, содержащееся в одном дискретном сообщении источника равновероятных сообщений с объемом алфавита равным двум.

Каждая буква русского алфавита (если считать, что е=ё) несет информацию **5 бит ($32 = 2^5$).**

Основы теории информации

Количество информации, которое содержит сообщение, закодированное с помощью знаковой системы, равно количеству информации, которое несет один знак, умноженному на число знаков в сообщении.
Следовательно, слово на русском языке из 5 букв в появившемся тексте (сообщении) содержит количество информации 5x5 = 25 бит.
При таком подходе в результате сообщения о результате бросания кубика, получим различное количество информации в зависимости от числа точек на выпавшей грани. Чтобы его подсчитать, нужно умножить количество символов (точек) на количество информации, которое несет один символ (точка).

На кубике в общей сложности 21 точка, следовательно количество информации одной точки $I = log_2 21 \sim 4{,}37(бит)$. Значит, если выпала грань с двумя точками, то количество информации равно ~ *8,74 (бит).*

Основы теории информации

Единица и ноль

Хотя внутренний язык некоторых компьютеров первого поколения был основан на десятичной системе счисления, начиная с 50-х годов практически во всех цифровых вычислительных машинах применялась уже двоичная система. Наличие всего двух символов значительно упрощало и удешевляло схемы, построенные на основе этой системы. Микроскопические электронные переключатели в центральном процессоре современного компьютера принимают только два состояния - они либо проводят ток, либо нет, представляя тем самым значения 0 и 1. Для схем, построенных на десятичной системе, потребовалось бы 10 различных состояний. Двоичная система соответствует также алгебраической системе логики, разработанной в XIX в. английским математиком Джорджем Булем. В рамках этой системы высказывание может быть либо истинным, либо ложным, подобно тому как переключатель может быть либо открытым, либо закрытым, а двоичный разряд - равен 1 или 0.

Если расположение переключателей соответствует булевым функциям, то образованные из этих переключателей схемы могут выполнять как арифметические, так и логические операции. Такая арифметическая схема, называемая двоичным сумматором. Функции сумматоров определяются самим их названием: они суммируют двоичные числа, следуя правилам, аналогичным правилам сложения десятичных чисел.

Основы теории информации

Почему именно основание два ?

Система счисления с основанием два более всего соответствует способу представления информации в компьютере. На самом деле компьютеры "понятия не имеют" ни о каких буквах, цифрах, командах или программах, поскольку представляют собой сложные электрические схемы, которые способны распознавать лишь напряжение и силу тока.

Для упрощения логики микросхем инженеры отказались от измерения силы тока (слабый, средний, большой и очень большой) и различают лишь два состояния (есть ток и нет тока). Собственно, в этом и состоит разница между аналоговой техникой и цифровой. Состояния "есть ток" и "нет тока" можно выразить и иначе, например: да или нет, ИСТИНА или ЛОЖЬ, true или false, 1 или 0. В соответствии с общепринятым соглашением единица равна true или истине, или «да». Неоспоримое преимущество двоичной системы счисления: с помощью единиц и нулей можно описать состояние отдельного элемента электрической схемы (есть ток или нет тока).

В электричестве может только два вида неисправностей:
Ток есть там, где его не должно быть, и тока нет там, где он должен быть.

Основы теории информации

Биты, байты и полубайты

Решив представлять данные последовательностями единиц и нулей, минимальную единицу информации, содержащую один двоичный разряд, назвали битом (bit, от binary digit - двоичная цифра). В связи с тем, что первые компьютеры были способны обрабатывать одновременно лишь по восемь битов, считалось вполне естественным писать код, используя восьмиразрядные двоичные числа, называемые байтами (byte). Половина байта (4 бита) называется полубайтом.

С помощью восьми двоичных разрядов можно представить до 256-ти различных значений. Почему? Рассмотрим разряды: если все восемь битов установлены (единица), значение составит 255 (128+64+32+ 16+8+4+2+ 1), если не установлен ни один бит (все они равны нулю), значение составит нуль. А в диапазоне от нуля до 255-ти как раз и содержатся 256 возможных вариантов.

Основы теории информации

Двоичные числа

Компьютер использует наборы нулей и единиц для представления всего, с чем он работает. Машинные коды представляют собой пакеты нулей и единиц, понятных, центральному процессору и другим микросхемам. Отдельные наборы нулей и единиц можно перевести обратно в числа, понятные людям, но было бы ошибкой полагать, что это числа и что они на самом деле имеют такие значения.

Например, процессор Intel 8086 интерпретировал набор битов 10010101 как команду. Конечно, это число можно представить в десятичном виде (149), но для людей это не имеет никакого смысла.

Иногда числа представляют собой команды, иногда - значения, иногда - программный код. Одним из стандартизованных кодовых наборов является ASCII . В нем каждая буква или знак препинания имеет семиразрядное двоичное представление. Например, строчная буква "a" представлена двоичным числом 01100001. Хотя это значение можно преобразовать в десятичное число 97 (64 + 32 + 1), понимать его следует не как число, а как букву. Поэтому когда говорят, что буква "a" в стандарте ASCII представлена числом 97, на самом деле имеют в виду, что 97 – это номер (код) буквы «a» в кодировке ASCII.

Основы теории информации

От десятичных к двоичным

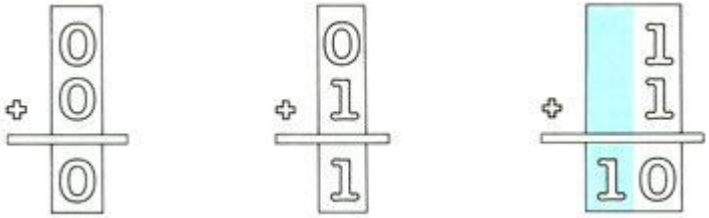
В двоичной системе счисления, как и в десятичной, значение цифры определяется ее положением относительно других цифр данного числа, т.е. ее позицией. Скажем, в десятичной системе цифра 1 означает число 1, но если справа от нее приписать два нуля, то она уже будет означать 100. На этом простом правиле держится вся арифметика: прежде чем складывать или вычитать числа, их нужно расположить так, чтобы равнозначные их разряды были выровнены по соответствующим столбцам.

В десятичной системе чем дальше слева от десят[10 1] [8 4 2 1]ходится цифра, тем в большую степень нужно[0] [0]вание системы (число 10), чтобы получ[1] [1]й цифры. В двоичной системе (основание 2) сдвиг[2] [10]ифры на одну позицию влево означает у[3] [11]ницу показателя степени, в которую нужно[4] [100]2. Так, 2 в степени 0 равно 1, 2 в степени 1 равно[5] [101] равно 4 и т. д. Чтобы найти десятичный эквив[6] [110] двоичного числа, достаточно просто заметить, в [7] [111]ициях расположены единичные разряды, и сло[8] [1000]. [9] [1001] [10] [1010]

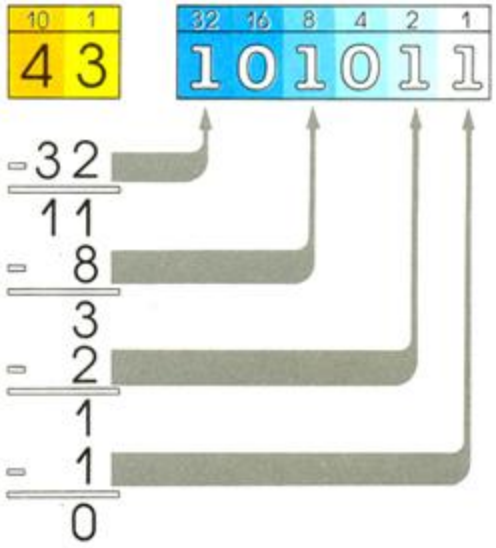
Основы теории информации

Чтение двоичных чисел. Поскольку в двоичной системе лишь две цифры, значение разрядов числа возрастает как степень двойки и двоичные числа быстро превращаются в длинные цепочки из 0 и 1. Сложив значения разрядов, в которых находятся единицы, мы получим десятичный эквивалент числа. Так, например, двоичное число 101 - это 4 плюс 1, т. е. десятичное число 5.

Правила сложения



Преобразование десятичного числа в двоичное. Чтобы перевести десятичное число в двоичное, нужно сначала вычесть из него число, равное максимально возможной степени двойки, а затем все время вычитать максимальные степени двойки уже из остатка, ставя единицу в тех позициях, где вычитание возможно, и 0 там, где нет. Цепочка вычитаемых чисел для десятичного числа 43 - это 32 (есть), 16 (нет), 8 (есть), 4 (нет), 2 (есть) и 1 (есть).

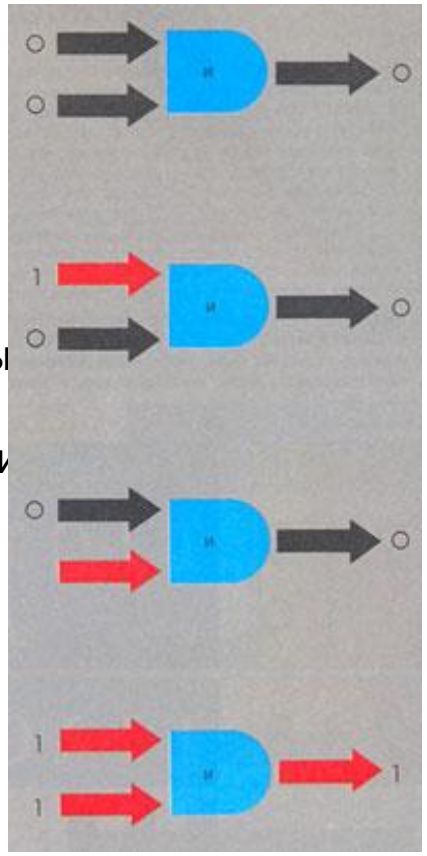


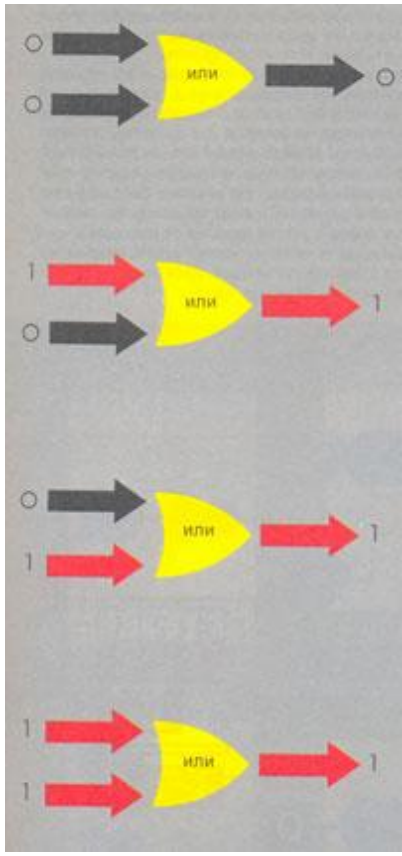
Основы теории информации

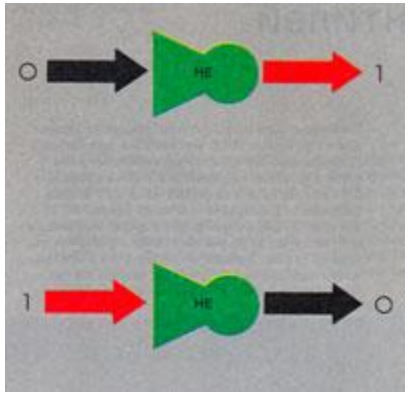
Логические схемы

Во всех современных компьютерах применяется логическая система, изобретенная Джорджем Булем. Тысячи микроскопических электронных переключателей в кристалле интегральной схемы сгруппированы в системы «вентилей», выполняющих логические операции, т.е. операции с предсказуемыми результатами. На рисунках показаны элементарные логические вентили И, ИЛИ и НЕ.

Все остальные логические схемы компьютера могут быть построены на основе вентилей этих трех типов. Соединенные в различные комбинации, логические вентили дают возможность компьютеру решать задачи с помощью закодированных импульсов его двоичного языка. На вход каждого логического вентиля поступают электрические сигналы высокого и низкого уровней напряжения, которые он интерпретирует в зависимости от своей функции и выдает один выходной сигнал также либо низкого, либо высокого уровня. Эти уровни соответствуют одному из состояний двоичной системы: да - нет, единица - нуль, истина - ложь.



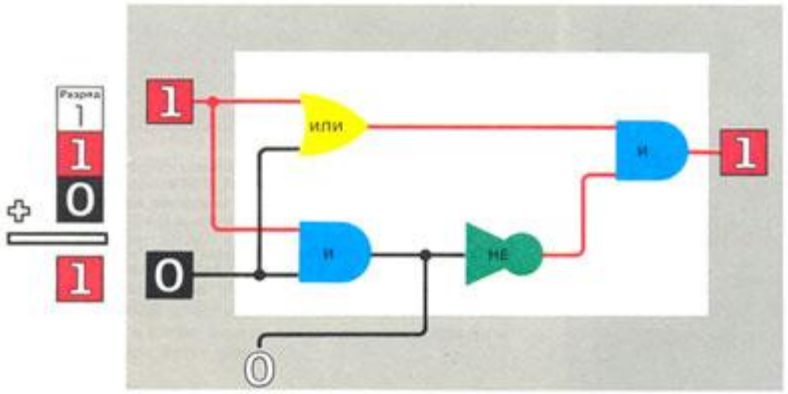


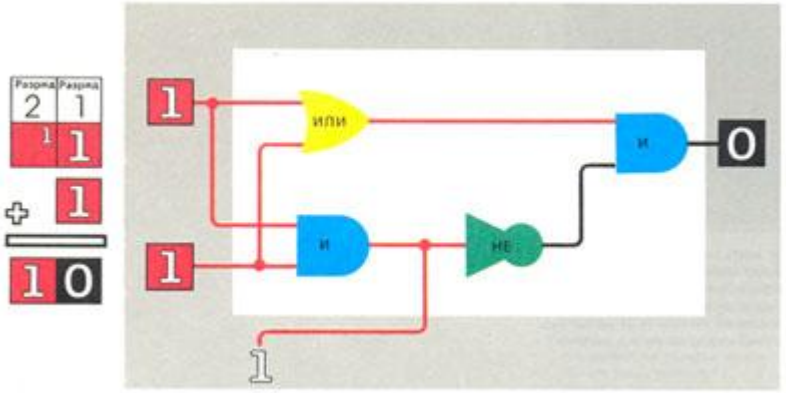


Основы теории информации

Сумматор

Вентили **И**, **ИЛИ** и **НЕ** соединяются в различные комбинации, которые образуют, электронные схемы, называемые полусумматорами и полными сумматорами. С помощью таких схем компьютер производит двоичное сложение. С некоторыми изменениями эти схемы могут применяться также для вычитания, умножения и деления.

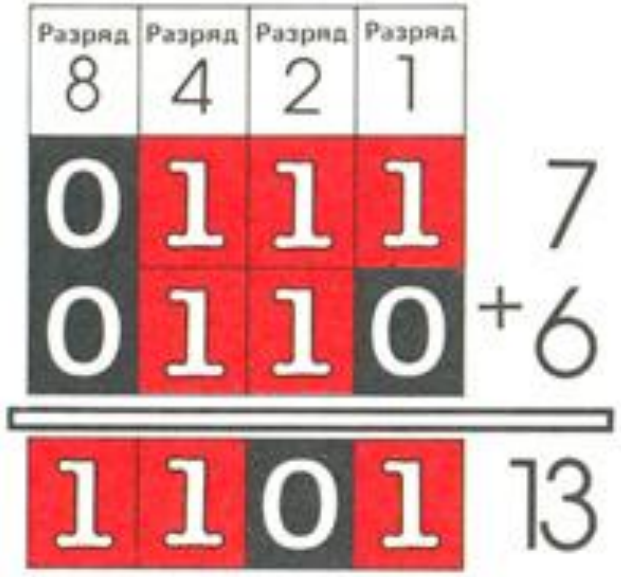




Подобно тому как вентили соединяются в сумматоры, отдельные сумматоры можно связать в единую схему, образующую каскадный сумматор - устройство, в котором на каждую пару битов приходится один сумматор.

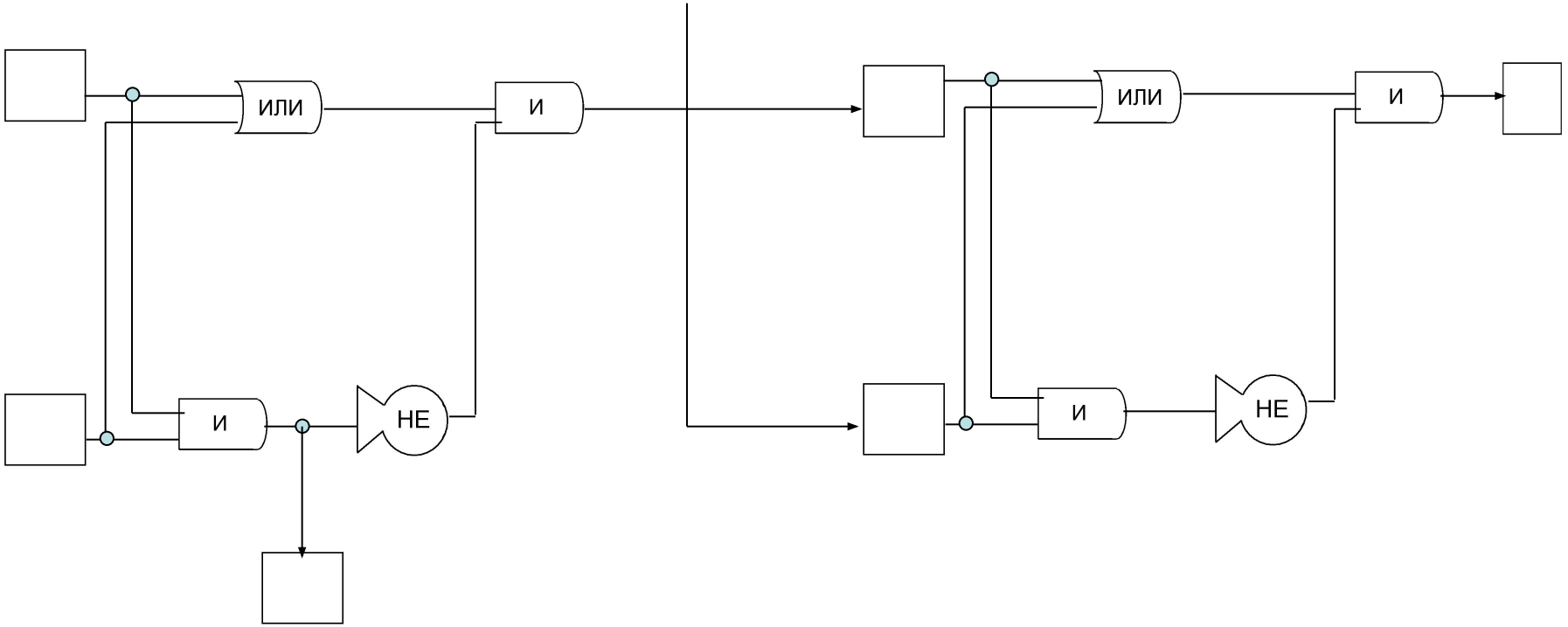
Основы теории информации

В примере два четырехбитных числа (справа) складываются при помощи каскада из 4 сумматоров. Для самого младшего разряда используется полусумматор, который может генерировать бит переноса, но сам не получает его. Остальные сумматоры - полные. Цепочку сумматоров можно продолжить так, чтобы это позволило складывать двоичные числа с требуемым количеством разрядов.



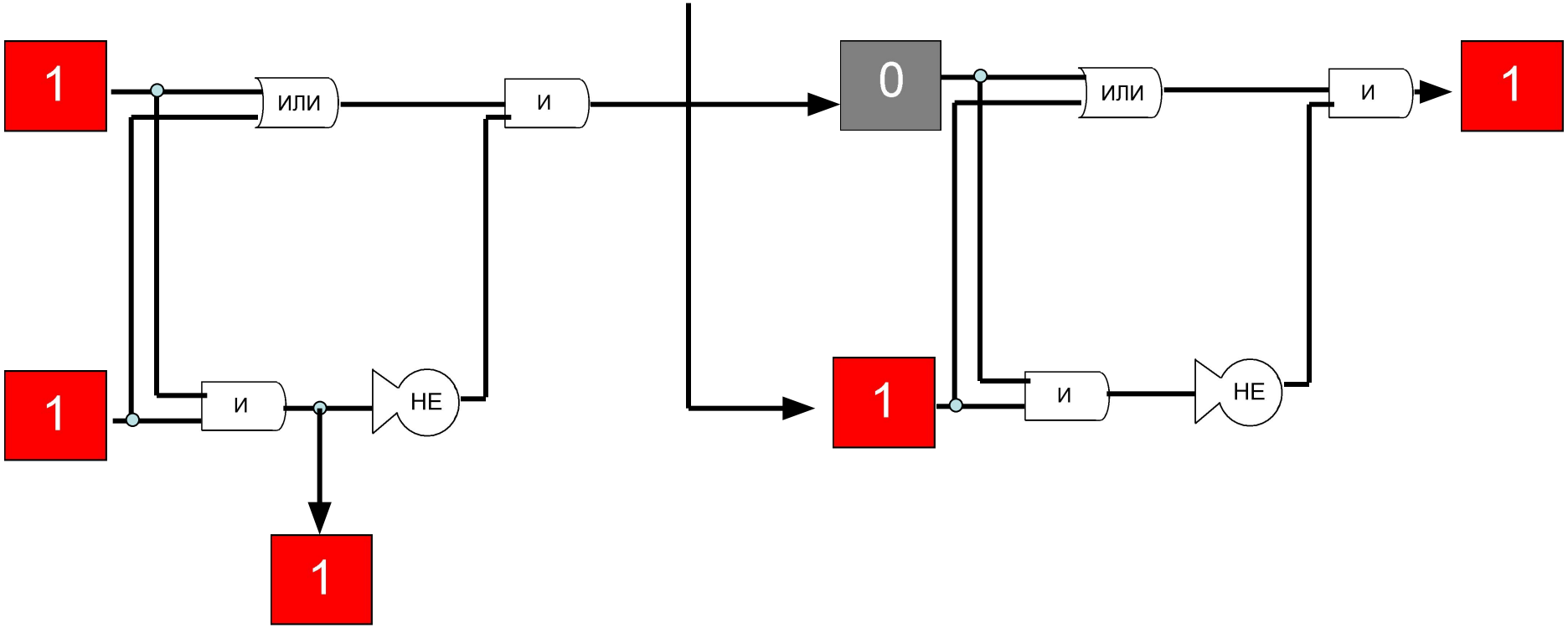
Основы теории информации

Логическая схема полного сумматора



Основы теории информации

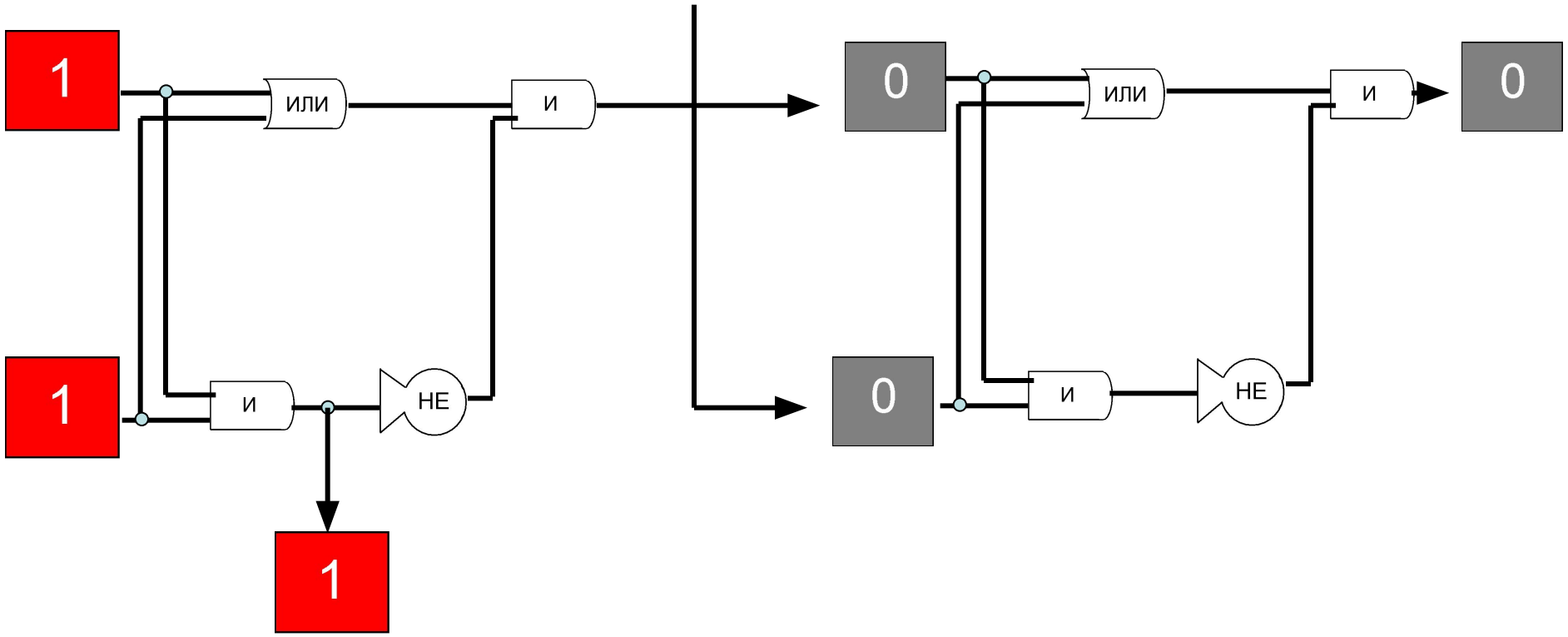
1. Логическая схема полного сумматора



На входах «1», из низшего разряда переходит «1»

Основы теории информации

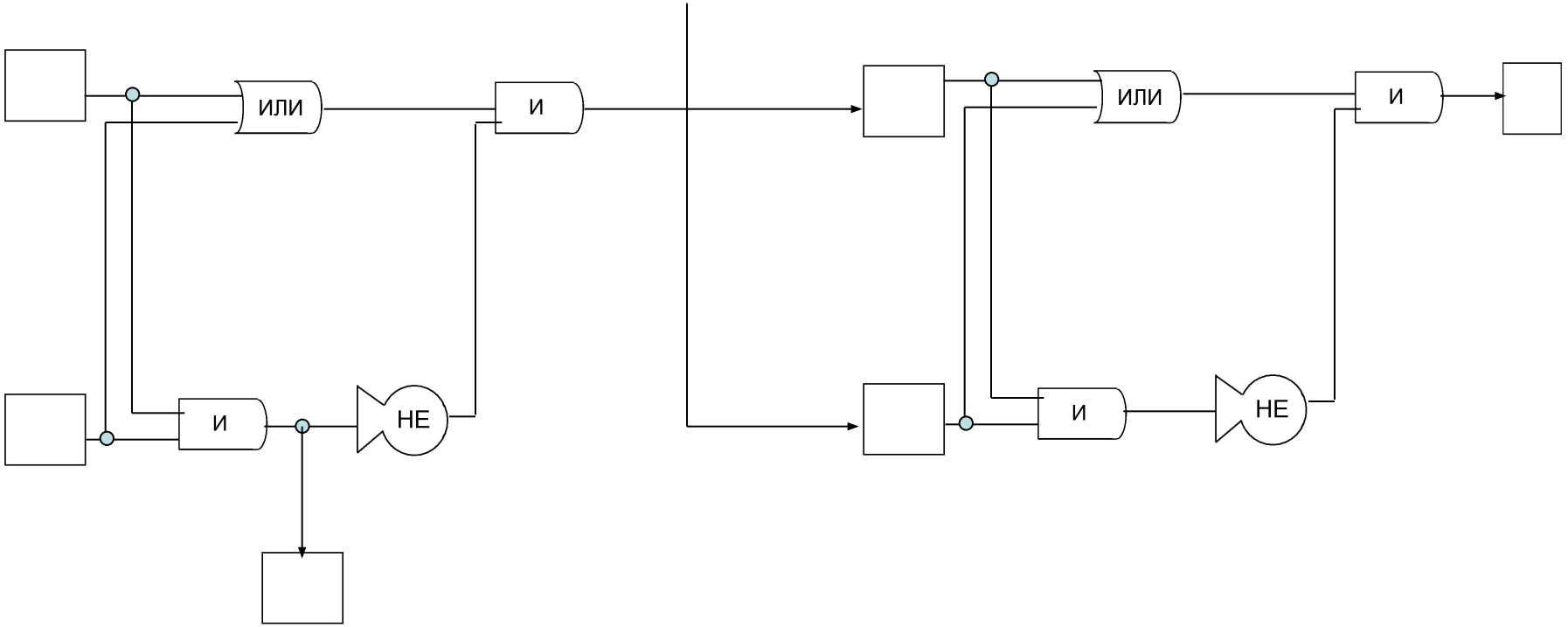
Логическая схема полного сумматора



На входах «1», из низшего разряда переходит «0»

Основы теории информации

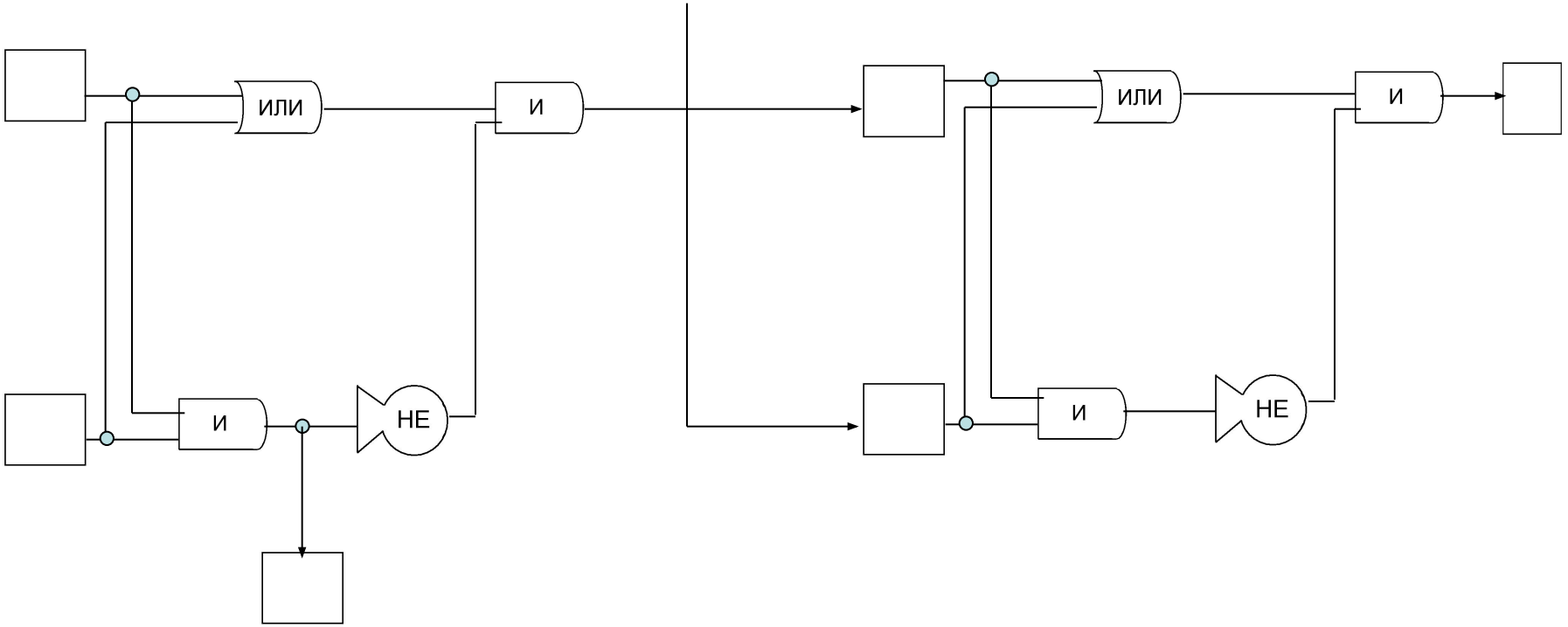
Логическая схема полного сумматора



На входах «1» и «0», из низшего разряда переходит «1»

Основы теории информации

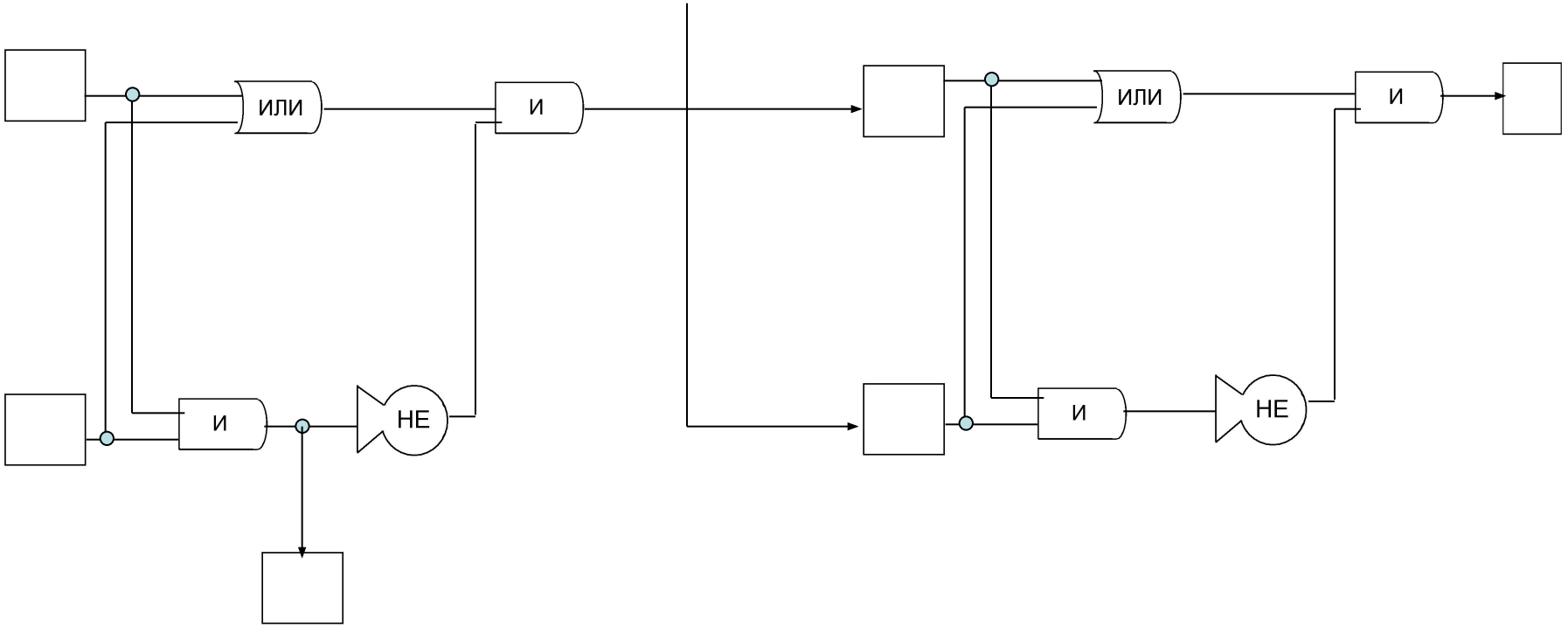
Логическая схема полного сумматора



На входах «1» и «0», из низшего разряда переходит «0»

Основы теории информации

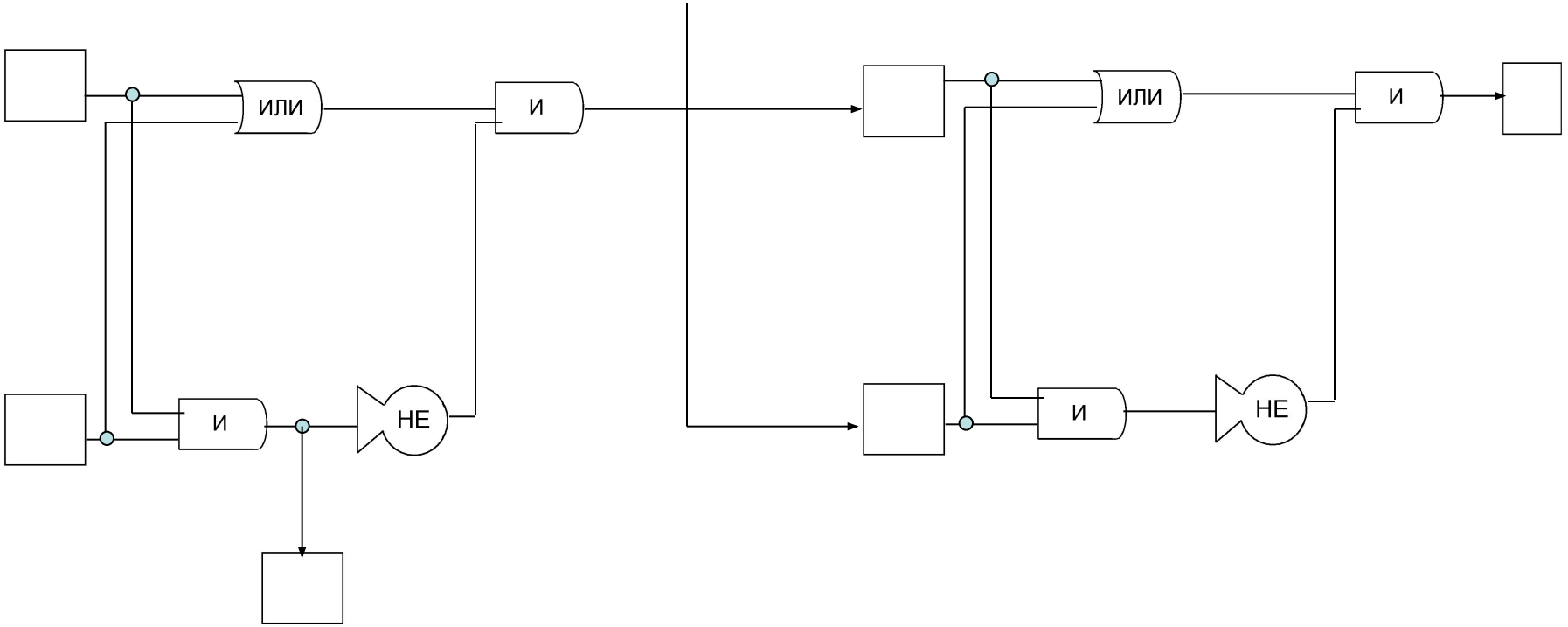
Логическая схема полного сумматора



На входах «0» и «0», из низшего разряда переходит «1»

Основы теории информации

Логическая схема полного сумматора



На входах «0» и «0», из низшего разряда переходит «0»