УДК 004.01

Ежицкая Дарья Дмитриевна, бакалавр, Сочинский государственный университет, г. Сочи.

ПОЧЕМУ БИТ ЯВЛЯЕТСЯ НАИМЕНЬШЕЙ ЕДИНИЦЕЙ ИЗМЕРЕНИЯ ИНФОРМАЦИИ?

В данной статье проводится комплексное исследование, направленное на установление фундаментальных причин, по которым бит считается минимальной единицей измерения информации. Исследование базируется на методологии логического анализа и опирается на фундаментальные положения теории информации Клода Шеннона, в частности, на понятие информационной энтропии и ее математический аппарат. Доказывается, что бит, как единица выбора между двумя равновероятными состояниями, является наименьшей и неделимой единицей. Статья прослеживает связь между понятием алфавита, мощностью алфавита и минимальной единицей информации, а также дает историческую справку о происхождении термина «бит».

This article presents a comprehensive study aimed at establishing the fundamental reasons why the bit is considered the smallest unit of information. The study draws on the methodology of logical analysis and draws on the fundamental tenets of Claude Shannon's information theory, particularly the concept of information entropy and its mathematical framework. It is demonstrated that the bit, as a unit of choice between two equally probable states, is the smallest and most indivisible unit. The article traces the relationship between the concept of an alphabet, the cardinality of an alphabet, and the smallest unit of information, and also provides a historical overview of the origins of the term "bit."

Ключевые слова: бит, информация, теория информации, Клод Шеннон, алфавит, двоичная система, единицы измерения, энтропия.

Keywords: bit, information, information theory, Claude Shannon, alphabet, binary system, units of measurement, entropy.

Понятие информации является центральным в современной науке, от компьютерных технологий до нейробиологии. Однако для ее точного измерения и обработки необходима фундаментальная, неделимая единица, аналогичная атому в химии или электрону в физике. Такой единицей в информатике и теории информации принято считать бит. Цель данной статьи – дать строгое логико-математическое обоснование тому, почему именно бит, а не какая-либо меньшая величина, считается наименьшей единицей измерения информации. Для достижения этой цели обратимся к базовым понятиям знаковых систем и фундаментальным принципам, заложенным в работах основателя теории информации Клода Шеннона.

1. Алфавит как знаковая система и основа информации

Любое сообщение, будь то текст, число или изображение, строится на основе определенной знаковой системы. Первичным элементом такой системы является алфавит – конечный набор символов (знаков). Важно понимать, что алфавит представляет собой не только набор букв, но и набор цифр, спецсимволов или любых других условных обозначений [1, с. 45].

Ключевой характеристикой алфавита является его мощность – количество символов в нем. Для формирования осмысленных комбинаций (слов, кодов) алфавит должен содержать как минимум два различных символа. Если алфавит состоит лишь из одного символа (например, только «А»), то любая последовательность будет представлять собой одно и то же повторяющееся значение. В такой системе невозможно закодировать различие, а значит, невозможно передать никакое сообщение, кроме факта своего существования. Следовательно, минимально возможная мощность алфавита, пригодного для кодирования информации, равна двум.

2. Двоичный алфавит и рождение бита

Таким образом, самый маленький в мире функциональный алфавит – двоичный. Его символами традиционно являются «0» и «1». Каждая позиция в последовательности, использующей этот алфавит, может находиться только в одном из двух состояний.

Термин для обозначения такой двоичной цифры был введен в научный обиход в 1948 году американским математиком Джоном Тьюки. Он сократил английское выражение «Binary digit» до слова «bit», взяв первые две буквы от «binary» и последнюю букву от «digit» [2, с. 112]. Этот лаконичный термин точно отражал суть понятия.

Однако теоретическое обоснование первостепенной важности двоичной системы было дано в фундаментальном труде Клода Шеннона «Математическая теория связи», опубликованном в том же 1948 году. Шеннон заложил основы теории информации, определив количество информации как меру уменьшения неопределенности [3]. С точки зрения этой теории, информация возникает тогда, когда происходит выбор одного события из множества возможных.

3. Теоретическое обоснование минимальности бита на основе теории Шеннона

Клод Шеннон предложил формулу для вычисления среднего количества информации в сообщении, основанную на понятии энтропии. Для случая, когда события равновероятны, количество информации H в битах, необходимое для устранения неопределенности из N равновероятных событий, выражается формулой:

$$H = log_2 N$$

Рассмотрим простейший случай: выбор между двумя равновероятными событиями (N=2). Например, бросок монеты (орел/решка), ответ на вопрос «да/нет», состояние логического элемента «истина/ложь». Подставив значение в формулу, получаем:

$$H = \log_2 2 = \log_2 2^1 = 1 \text{ бит}$$

Это означает, что сообщение о результате такого выбора содержит ровно 1 бит информации. Теперь зададимся ключевым вопросом: можно ли получить количество информации меньше одного бита? Предположим, что мы пытаемся уменьшить мощность алфавита до одного символа (N=1). Подставив это значение в формулу Шеннона, получим:

$$H = \log_2 1 = \log_2 2^0 = 0 \text{ бит}$$

Ноль бит информации означает полное отсутствие неопределенности и, следовательно, отсутствие информационного сообщения. Система с одним состоянием не может служить для выбора или кодирования различий.

4. Логическое доказательство через редукцию алфавита

Для строгого обоснования тезиса о том, что бит является наименьшей единицей информации, применим метод доказательства от противного. Этот метод предполагает, что допускается истинность противоположного утверждения и показываем, что это приводит к логическому противоречию или результату, который противоречит установленным фактам.

Тезис (T). *Бит является наименьшей, неделимой единицей измерения информации.*

Доказательство: предположим, что тезис T ложен. Тогда справедливо противоположное утверждение: бит не является наименьшей единицей измерения информации. Существует некая единица измерения информации, меньшая, чем бит. Обозначим ее как «минимальную информационную единицу» (M), где 1 бит $= k * M$, и $k > 1$ (например, $k = 2, 3, 4$ ...).

Если $\neg T$ истинно, то должна существовать знаковая система (алфавит), основанная на этой меньшей единице – M. Поскольку M меньше бита, мощность соответствующего алфавита (N) должна быть меньше мощности двоичного алфавита. Мощность двоичного алфавита равна 2. Следовательно, мощность алфавита для M должна быть: $1 < N < 2$.

Проанализируем возможность существования алфавита мощностью $1 < N < 2$. Алфавит по определению есть конечное множество дискретных символов [3, с. 56]. Мощность множества (количество элементов) есть целое число $(1, 2, 3$ ...).

Следовательно, не существует целого числа N, которое бы удовлетворяло условию $1 < N < 2$. Мощность алфавита не может быть дробной. Единственные возможные варианты для минимальных алфавитов – это $N = 1$ или $N = 2$.

Таким образом, предположение $(\neg T)$ привело к необходимости существования алфавита с мощностью, не являющейся целым числом, что противоречит самому определению алфавита. Это является логическим противоречием.

Убедимся, что граничные случаи $N = 1$ и $N = 2$ не поддерживают антитезис.

1) Случай $N = 2$. Это двоичный алфавит $\{0, 1\}$. Как следует из формулы Шеннона для равновероятных событий, количество информации в таком сообщении составляет $H = \log_2 2 = \log_2 2^1 = 1\,\text{бит}$. Это соответствует исходному тезису T.

2) Случай $N = 1$. Рассмотрим алфавит, состоящий из одного символа, например, $\{0\}$. Энтропия Шеннона для такой системы равна $H = \log_2 1 = \log_2 2^0 = 0\,\text{бит}$. Это означает, что в системе отсутствует неопределенность. Любое сообщение, составленное из одного символа (например, "0", "00", "000..."), не несет информации, так как оно всегда было предсказуемо и неизменно. Система с $N = 1$ не может служить для кодирования или передачи информации, а лишь для констатации наличия сигнала. Следовательно, эта система не определяет единицу измерения информации, а демонстрирует ее отсутствие.

Поскольку предположение о существовании единицы информации, меньшей, чем бит (антитезис $\neg T$), привело к логическому противоречию (требованию наличия алфавита с нецелой мощностью), а анализ граничных случаев показал, что алфавит с $N = 2$ порождает единицу информации, равную 1 биту, а алфавит с $N = 1$ имеет нулевую информационную емкость, то единственно верным является исходный тезис T.

Что и требовалось доказать.

Проведенный анализ позволяет с уверенностью утверждать, что бит является наименьшей единицей измерения информации. Это обосновано как логически, через анализ свойств алфавита как знаковой системы, так и математически, через фундаментальные положения теории информации Клода Шеннона.

Минимальная мощность алфавита, пригодного для кодирования, равна двум. Выбор одного из двух равновероятных событий содержит ровно 1 бит информации, что математически выражается как $H = \log_2 2 = \log_2 2^1 = 1$. Любая попытка редуцировать систему до одного состояния приводит к нулевой информационной емкости.

Таким образом, бит выступает своеобразным «атомом» информационного мира – неделимой основой, на которой строится все многообразие данных и знаний. Понимание этого принципа является ключевым для дальнейшего изучения информатики, теории кодирования и цифровых технологий.

Список литературы

1. Шеннон К. Работы по теории информации и кибернетике. – М.: Иностранная литература, 1963. – 830 с.
2. Таненбаум Э. Архитектура компьютера. – 6-е изд. – СПб.: Питер, 2013. – 816 с
3. Хэмминг Р. В. Теория кодирования и теория информации. – М.: Радио и связь, 1983. – 176 с.

References

1. Shannon, K. (1963), “Works on Information Theory and Cybernetics.” Moscow: Foreign Literature, 830 p.
2. Tanenbaum, E. (1963), “Computer Architecture.” 6th ed. St. Petersburg: Piter, 2013. 816 p.
3. Hamming, R. W. (1983), “Coding Theory and Information Theory.” Moscow: Radio and Communications, 176 p.