

Николай Петрович Морозов

**КАК ИЗМЕРЯЕТСЯ
ИНФОРМАЦИЯ?**

Сборник задач по информатике

Николай Морозов

**Как измеряется информация?
Сборник задач по информатике**

«Издательские решения»

Морозов Н. П.

Как измеряется информация? Сборник задач по информатике /
Н. П. Морозов — «Издательские решения»,

ISBN 978-5-00-688071-9

Информация измеряется несколькими основными способами, зависящими от контекста её восприятия и обработки. Основные подходы следующие: 1. Алфавитный подход. 2. Содержательный подход (вероятностный метод). 3. Статистический подход (энтропийный метод). В данной книге эти подходы к измерению информации нашли свое отражение в решении более 20 задач и примеров. Это может быть интересно и школьникам и студентам.

ISBN 978-5-00-688071-9

© Морозов Н. П.
© Издательские решения

Содержание

1. Условная мера количества информации (бит)	6
2. Информация по Шеннону (энтропия)	8
3. Количество информации при передаче сообщения	10
4. Примеры практического применения	11
а) Кодирование текста	11
б) Компрессия данных	12
с) Передача цифровых сигналов	13
Заключение	14
5. Решение задач и примеров	15
Задача №1: Кодирование текста	16
Задача №2: Сжатие данных (алгоритм Хаффмана)	17
Конец ознакомительного фрагмента.	18

Как измеряется информация? Сборник задач по информатике

Николай Петрович Морозов

© Николай Петрович Морозов, 2025

ISBN 978-5-0068-8071-9

Создано в интеллектуальной издательской системе Ridero

Измерение информации – важная задача теории информации, поскольку позволяет количественно оценить объем передаваемых сообщений, количество информации, содержащейся в сообщении, и степень неопределенности ситуации.

Рассмотрим основные подходы измерения информации с подробными пояснениями и примерами.

1. Условная мера количества информации (бит)

Информация измеряется в битах (от англ. bit – binary digit). Один бит – это количество информации, необходимое для выбора одного из двух равновероятных исходов.

Формула Хартли:

$$N = 2^i$$

Где, N – количество равновероятных возможных событий; I – количество информации в сообщении о том, что произошло одно из N событий.

Предположим, вы подбрасываете монету. Есть два равновероятных события: выпадение орла или решки ($N=2$). Измерим количество информации:

$$i = \log_2 N \quad \Rightarrow \quad I = \log_2 2 = 1 \text{ бит}$$

Это значит, что результат броска монеты несет ровно один бит информации.

Пример 1:

Выбрали случайную карту из колоды из 52 карт. Сколько информации содержит сообщение о выбранной карте?

$$I = \log_2 52 \approx 5.9 \text{ бит}$$

2. Информация по Шеннону (энтропия)

Шеннон предложил учитывать вероятность каждого события, чтобы точнее измерять информацию. В отличие от формулы Хартли, здесь учитывается случай, когда события не равновероятны

Формула Шеннона:

$$H(X) = - \sum_{i=1}^n p_i \cdot \log_2 p_i,$$

где:

- $H(X)$ – энтропия (количество информации),
- n – общее число событий,
- p_i – вероятность i -го события.

Энтропия показывает среднее количество информации, которое мы получаем при наблюдении события.

Пример 2:

Рассмотрим игральный кубик. Событие выпадения каждой грани имеет одинаковую вероятность $1/6$. Тогда энтропия равна:

$$H(X) = - \sum_{i=1}^6 \frac{1}{6} \cdot \log_2 \frac{1}{6} = - \frac{1}{6} (\log_2 \frac{1}{6} + \dots + \log_2 \frac{1}{6}) = - 6 \cdot \frac{1}{6} \cdot (-\log_2 6) = \log_2 6 \approx 2.58 \text{ бит}$$

Таким образом, бросок кубика несет примерно 2.58 бита информации.

Пример.3.

Рассмотрим другой пример, где вероятности граней неодинаковы:

Вероятности выпавших чисел следующие: $p(1) = \frac{1}{2}$, $p(2) = p(3) = p(4) = p(5) = p(6) = \frac{1}{10}$. Тогда энтропия будет меньше:

$$H(X) = - \left(\frac{1}{2} \log_2 \frac{1}{2} + 4 \cdot \frac{1}{10} \log_2 \frac{1}{10} \right) \approx 1.72 \text{ бит}$$

Здесь энтропия уменьшилась, потому что одно событие стало значительно вероятнее остальных.

3. Количество информации при передаче сообщения

При передаче сообщения важно понимать, сколько информации передается через канал связи. Для этого используют понятие информационной пропускной способности канала.

Информационная пропускная способность канала определяется формулой Шеннона-Хартли:

$$C = B \cdot \log_2 \left(1 + \frac{S}{N} \right),$$

где:

- C — максимальная скорость передачи информации (бит/секунда),
- B — ширина полосы частот канала (Гц),
- S — мощность сигнала,
- N — мощность шума.

Пример 4:

Допустим, имеется канал связи шириной 10 МГц, отношение сигнал/шум равно 15 дБ (соответствует отношению мощностей около 31.6):

$$C = 10^7 \cdot \log_2(1 + 31.6) \approx 10^7 \cdot 4.954 \approx 49.54 \text{ Мбит/с}$$

4. Примеры практического применения

а) Кодирование текста

Текстовые файлы кодируются различными способами. Например, ASCII-кодировка использует 8 бит на символ, UTF-8 – переменное количество байтов (от 1 до 4 байт).

b) Компрессия данных

Алгоритмы сжатия данных (например, ZIP, RAR) используют принципы уменьшения избыточности, основываясь на понятиях энтропии и условной вероятности.

с) Передача цифровых сигналов

Для эффективной передачи данных по каналам связи используется теория информации, позволяющая минимизировать потери информации и повысить надежность передачи.

Заключение

Таким образом, измерение информации осуществляется разными методами в зависимости от контекста и цели анализа.

Основные понятия включают бит, энтропию Шеннона и пропускную способность каналов связи. Эти методы позволяют эффективно оценивать объемы информации, оптимизировать передачу данных и разрабатывать эффективные алгоритмы обработки информации.

5.Решение задач и примеров

Давайте разберём три примера задач, связанных с кодированием текста, сжатием данных и передачей цифровых сигналов. Каждый пример сопровождается подробным пошаговым решением.

Задача №1: Кодирование текста

Определить объём памяти, который потребуется для хранения текста, состоящего из 100 символов, закодированных в формате Unicode UTF-8. Известно, что первые 128 символов ASCII занимают 1 байт, символы кириллицы – 2 байта.

Решение:

1. Определение структуры кодировки UTF – 8:

- Символы ASCII (кодировка первых 128 символов): 1 байт (8 бит).
- Кириллические символы (диапазон 128—255): 2 байта (16 бит).

2. Расчёт объёма памяти:

Предположим, что текст состоит из 50 латинских букв (ASCII) и 50 русских букв (кириллица).

– Латинские буквы (ASCII):

50 символов × 1 байт = 50 байт.

– Русские буквы (UTF-8):

50 символов × 2 байта = 100 байт.

Итого общий объём памяти:

$$V_{\text{общий}} = 50 \text{ байт} + 100 \text{ байт} = 150 \text{ байт}$$

Ответ:

Объём памяти, необходимый для хранения такого текста, составляет 150 байт.

Задача №2: Сжатие данных (алгоритм Хаффмана)

Конец ознакомительного фрагмента.

Текст предоставлен ООО «Литрес».

Прочитайте эту книгу целиком, [купив полную легальную версию](#) на Литрес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.