



Тема. Алфавитный подход к определению количества информации

Тема. Алфавитный подход к определению количества информации

- Сколько символов в компьютерном алфавите?
- Каков объем информации, содержащейся в книге, на аудиокассете, на компакт-диске, в библиотеке?
- Для передачи информации в объеме 10 учебников можно затратить всего одну минуту. Как это сделать?

Научились определять количество информации в сообщениях, уменьшающих неопределенность знаний, рассматривая информацию с позиции человека.

Но...

Вокруг нас везде и всюду происходят информационные обмены. Информацией обмениваются между собой люди, животные, технические устройства, органы человека или животного и т.д. Во всех этих случаях передача информации проходит в виде последовательностей различных сигналов. В вычислительной технике такие сигналы кодируют определенные смысловые символы – буквы, цифры, коды цвета точек и т.д.

Множество используемых в тексте символов называется **алфавитом**.

В информатике алфавит – это не только буквы, но и цифры, и знаки препинания, и другие специальные знаки.

Размер алфавита – количество его символов называется **мощностью алфавита**

$$N = 2^I$$

- формула нахождения количества информации

Для расчета количества
информации по формуле необходимо знать
мощность алфавита

$$N = 2^I$$

- формула нахождения количества информации

Пример 1

Найти объём информации, содержащейся в тексте из 3000 символов, написанном русскими буквами.

Решение:

1) Найдём мощность алфавита: 33 русских прописных букв, 33 русских строчных букв, 21 специальный знак.

$N = 33 + 33 + 21 = 87$ символов.

2) Подставим в формулу и рассчитаем количество информации:

$$87 = 2^I, \quad I \approx 6,4 \text{ бита.}$$

3) Итак, 6,4 бита информации несет один символ в русском тексте.

Так как в тексте 3000 символов, значит, информационный объём данного текста $6,4 \cdot 3000 = 19140$ бит.

Ответ: 19140 бит

$$N = 2^I$$

- формула нахождения количества информации

Пример 2

Найти объём информации, содержащейся в немецком тексте с таким же количеством символов.

Решение:

1) Найдём мощность немецкого алфавита: 26 немецких прописных букв, 26 немецких строчных букв, 21 специальный знак.

$N = 26 + 26 + 21 = 73$ символа.

2) Подставим в формулу и рассчитаем информационный объём одного символа:

$73 = 2^I, I \approx 6,1$ бита.

3) Найдём объём всего текста:

$6,1 \cdot 3000 = 18300$ бит.

Ответ: 18300 бит

$$N = 2^I$$

- формула нахождения количества информации

Сравнивая объёмы информации русского и немецкого текста, мы видим, что на немецком языке информации меньше чем на русском.

Но... содержание не изменилось!

Следовательно, при алфавитном подходе к измерению информации ее количество не зависит от содержания, а зависит от мощности алфавита и количества символов в тексте.

Таким образом, с точки зрения алфавитного объёма, в толстой книге информации больше, чем в тонкой. При этом содержание книги не учитывается.

$$N = 2^I$$

- формула нахождения количества информации

Правило для измерения информации с точки зрения алфавитного подхода:

1. Найти мощность алфавита – N .
2. Найти информационный объём одного символа – I .
3. Найти количество символов в сообщении – K .
4. Найти информационный объём всего сообщения – $K \cdot I$.

$$N = 2^I$$

- формула нахождения количества информации

Пример 3

Найти объём текста, записанного на языке, алфавит которого содержит 128 символов и 2000 символов в сообщении.

Дано: $K = 2000$, $N = 128$.

Найти: I_T .

Решение:

1) $128 = 2^I$, $I = 7$ бит - объём одного символа.

2) $I_T = 7 \cdot 2000 = 14000$ бит - объём сообщения.

Ответ: 14000 бит.

Компьютерный алфавит содержит 256 символов.

Один символ такого алфавита несет 8 бит информации.

~~**8 бит = 1 байт.**~~

Учитывая тот факт, что большинство статей, книг, публикаций и т.д. написаны с помощью текстовых редакторов, то информационный объём любого сообщения можно находить как информационный объём компьютерного текста.

Компьютерный алфавит содержит 256 символов.

Один символ такого алфавита несет 8 бит информации.

~~8 бит = 1 байт.~~

Пример 4

Найти информационный объём страницы компьютерного текста.

14 страницы нашего учебника информатики.

Расчеты производим приблизительно.

Решение:

1) Найдём информационный объём одного символа:

$$256 = 2^I, I = 8 \text{ бит} = 1 \text{ байт.}$$

2) Найдём количество символов на странице. Примерно.

$$40 \text{ символов} \cdot 50 \text{ строк} = 2000 \text{ символов.}$$

3) Найдём информационный объём всей страницы:

$$2000 \cdot 1 = 2000 \text{ байт.}$$

Ответ: 2000 байт.

Как это сделать быстро?
Используя все знаки препинания и пробелы, и умножить

Компьютерный алфавит содержит 256 символов.

Один символ такого алфавита несет 8 бит информации.

~~8 бит = 1 байт.~~

Так как информационный объём одного символа несет 1 байт информации, то достаточно подсчитать количество символов в тексте, которое и даст объём текста в байтах.

Компьютерный алфавит содержит 256 символов.

Один символ такого алфавита несет 8 бит информации.

~~8 бит = 1 байт.~~

Для измерения больших объёмов информации используют следующие единицы:

1 килобайт = 1 Кб = 2^{10} байт = 1024 байт

1 мегабайт = 1 Мб = 2^{10} Кб = 1024 Кб

1 гигабайт = 1 Гб = 2^{10} Мб = 1024 Мб

Компьютерный алфавит содержит 256 символов.

Один символ такого алфавита несет 8 бит информации.

~~8 бит = 1 байт.~~

Пример 5

Найти информационный объём небольшой книги в 130 страниц.

Страницы взять из предыдущего примера.

Решение:

2000 байт · 130 = 260000 байт.

Ответ: 260000 байт.

Примеры объёмов информации

Страница книги	2,5 Кб
Учебник	0,5 Мб
БСЭ	120 Мб
Газета	150 Кб
Черно-белый телевизионный кадр	300 Кб
Цветной кадр из трех цветов	1 Мб
1,5-часовой цветной художественный фильм	135 Гб

В 100 Мб можно уместить:

Страниц текста	50000
Цветных слайдов высочайшего качества	150
Аудиозапись	1,5 часа
Музыкальный фрагмент качества CD-стерео	10 минут
Фильм высокого качества записи	15 секунд
Протоколы операций по банковским счетам	За 1000 лет

Скорость передачи информации называется **скоростью информационного потока.**

Выражается в бит/с, байт/с, Кб/с, и т.д.

Скорость информационного потока между техническими устройствами намного выше, чем между людьми.

Обмен информацией при этом происходит по каналам связи.

Основные характеристики каналов связи:

- максимальная скорость передачи информации по каналу связи называется **пропускной способностью канала;**
- **надёжность;**
- **стоимость;**
- **резервы развития.**

Характеристики некоторых каналов СВЯЗИ

Тип связи	Скорость передачи данных (Мб/с)	Помехоустойчивость	Наращиваемость
Электрический кабель: - витая пара - коаксиальный кабель	10-100 до 10	Низкая Высокая	Простая Проблематичная
Телефонная линия	1-2	Низкая	Без проблем
Оптические светодиоды (сверхтонкие силиконовые волокна)	10-200	Абсолютная	Без проблем

Тема. Алфавитный подход к определению количества информации

• **В компьютерном алфавите 256 символов.**

• Каков объём информации, содержащейся в книге, на аудиокассете, на компакт-диске, в библиотеке?

• **Передать по высоко скоростному оптоволокну.** Для передачи информации в объеме 10 учебников можно затратить всего одну минуту. Как это сделать?

Решить задачу:

Перевести объём книги 260000байт в другие единицы измерения.

Решени
е.

$$260000/1024 = 253,90625 \text{ Кб}$$

$$253,90625/1024 = 0,247955 \text{ Мб}$$

Ответ: 253,90625 Кб; 0,247955 Мб.

Домашнее задание

Уровень знания.

- 1) Как определяется количество информации с алфавитной точки зрения? Выучить правило для измерения информации с точки зрения алфавитного подхода.
- 2) Выучить единицы измерения информации.

Уровень понимания.

- 1) В чем отличие алфавитного подхода к измерению информации от вероятностного?
- 2) Выразить
 - 3 Кб в байтах и битах;
 - 81920 бит в байтах и Кб;
 - 3072 Мб в Гб и Кб.
- 3) Мощность некоторого алфавита равна 64 символа. Каким будет объём информации в тексте, состоящем из 100 символов.

Уровень применения.

Определите свою скорость речи и скорость чтения с точки зрения информатики

Творческий уровень.

- 1) Наберите на домашнем компьютере текст, информационный объём которого равен 24000 байт.
- 2) Наберите на домашнем компьютере текст, содержащий не нулевое количество символов и информационный объём которого равен нулю.