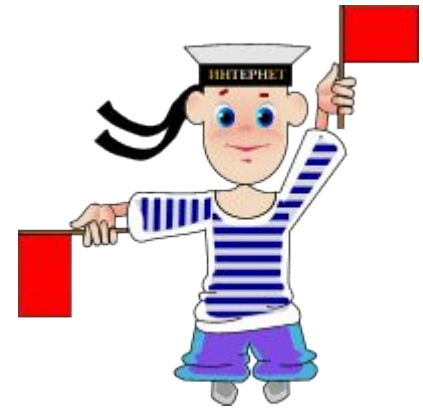


Алфавитный метод измерения информации

Автор: Харичева О.С., учитель информатики и ИКТ гимназии № 99 г.Екатеринбурга.
В работе использованы иллюстрации из книги А. А. Дуванова «Азы информатики»
(художник А. А. Русс) с разрешения автора.

Кодирование — перевод информации в удобную для передачи или хранения форму.



Сообщение кодируется с помощью некоторого набора знаков.

Если для данного набора установлен порядок следования знаков, то он называется **алфавитом**.



Конспект

Кодирование текстовой информации

Присвоение символу конкретного двоичного кода – это вопрос соглашения, которое фиксируется в кодовой таблице.

Конспект



Одним из первых стандартов кодирования русских букв на компьютерах был код КОИ-8 («Код обмена информационный – 8 битный»). Эта кодировка применяется в компьютерах с операционной системой UNIX.



Наиболее распространенная кодировка – это стандартная кириллическая кодировка **Microsoft Windows**, обозначаемая сокращением **CP1251** («CP» означает «Code Page»). Все **Windows** – приложения, работающие с русским языком, поддерживают эту кодировку.



Фирма Apple разработала для компьютеров **Macintosh** свою кодировку русских букв (**Mac**)



Стандарты кодировок:

1. **КОИ-8** (Код обмена информационный – 8-битный) - **UNIX**
2. **CP1251** (Code Page) - **Microsoft Windows**
3. **CP866** - **MS-DOS**
4. **Mac** - **Macintosh**
5. **ISO8859-5** (International Standards Organization)



Таблица кодировки символов

Двоичный код	Десятичный код	КОИ8	CP1251	CP866	Mac	ISO
0000 0000	0					
.....						
0000 1000	8	Удаление последнего символа (клавиша Backspace)				
.....						
0000 1101	13	Перевод строки (клавиша Enter)				
.....						
0010 0000	32	Пробел				
0010 0001	33	!				
.....						
0101 1010	90	Z				
.....						
0111 1111	127	□				
.....	128	-	Ъ	А	А	К
.....						
1100 0010	194	Б	В	-	-	Т
.....						
1100 1100	204	Л	М	:	:	Ь
.....						

Задача

2004-A2 Считая, что каждый символ кодируется одним байтом, оцените информационный объем следующего предложения:

«Мой дядя самых честных правил, Когда не в шутку занемог, Он уважать себя заставил И лучше выдумать не мог.»

- 1) 108 бит 2) 864 бит 3) 108 килобайт 4) 864 килобайт

- 1) Подсчитать количество символов: 108 включая кавычки и пробелы.
- 2) Перевести в байты: 108 символов = 108 байт.
- 3) Перевести в биты: 108 байт \times 8 = 864 бита.



Ответ: 2

В последнее время появился **новый** международный стандарт **Unicode**, который отводит на каждый символ не один байт, а два, и поэтому с его помощью можно закодировать не 256 символов, $2^{16}=65536$ различных символов. Эту кодировку поддерживает платформа Microsoft начиная с Windows&Office97.



Задача

2007-A1 Считая, что каждый символ кодируется 16-ю битами, оцените информационный объем следующей пушкинской фразы в кодировке Unicode:

Привычка свыше нам дана: Замена счастию она.

- 1) 44 бита 2) 704 бита 3) 44 байта 4) 704 байта

- 1) Подсчитать количество символов: 44 включая знаки препинания и пробелы.
- 2) Перевести в биты: 44 символа = 704 бита.



Ответ: 2

ЕГЭ

2005 Считая, что каждый символ кодируется одним байтом, оцените информационный объем следующего предложения из пушкинского четверостишия:

Певец-Давид был ростом мал, Но повалил же Голиафа!

- 1) 400 бит 2) 50 бит 3) 400 байт 4) 5 байт

1) Подсчитать количество символов и байт: 50.

2) Перевести байты в биты: 50 байт = 400 бит.



Ответ: 1

ЕГЭ

2006 Считая, что каждый символ кодируется двумя байтами, оцените информационный объем следующего предложения в кодировке Unicode:

Один пуд – около 16,4 килограмм.

1)32 Кбайта 2)512 бит 3)64 бита 4)32 байта

1) Подсчитать количество символов: 32.

2) Подсчитать количество байт: 64.

3) Перевести в биты: 64 байта = 512 бит.



Ответ: 2

ЕГЭ

2008 В кодировке Unicode на каждый символ отводится два байта. Определите информационный объем слова из двадцати четырех символов в этой кодировке.

1)384 бита 2)192 бита 3)256 бит 4)48 бит

1) Подсчитать количество байт: 48.

2) Перевести в биты: 48 байт = 384 бита.



Ответ: 1

ЕГЭ

2008-A1 Автоматическое устройство осуществило перекодировку информационного сообщения на русском языке, первоначально записанного в 16-битном коде Unicode, в 8-битную кодировку КОИ-8. При этом информационное сообщение уменьшилось на 400 бит. Какова длина сообщения в символах?

1) 25 2) 50 3) 100 4) 400

Ответ: 2