

Кодирование текстовой информации

Информация, выраженная с помощью естественных и формальных языков в письменной форме, обычно называется **текстовой информацией**.

Для обработки текстовой информации на компьютере необходимо представить ее в двоичной знаковой системе.

Для кодирования каждого знака требуется количество информации равное 8 битам и каждому знаку необходимо поставить в соответствие уникальный двоичный код из интервала 00000000 до 11111111 (в десятичном коде от 0 до 255).

Присваивание знаку конкретного двоичного кода – это вопрос соглашения, который фиксируется в кодовой таблице.

Структура существующих кодовых таблиц:

Десятичные коды с 0 до 33 – соответствуют операциям;

С 33 до 127 – являются интернациональными и соответствуют знакам латинского алфавита, цифрам, знакам арифметических операций и знакам препинания;

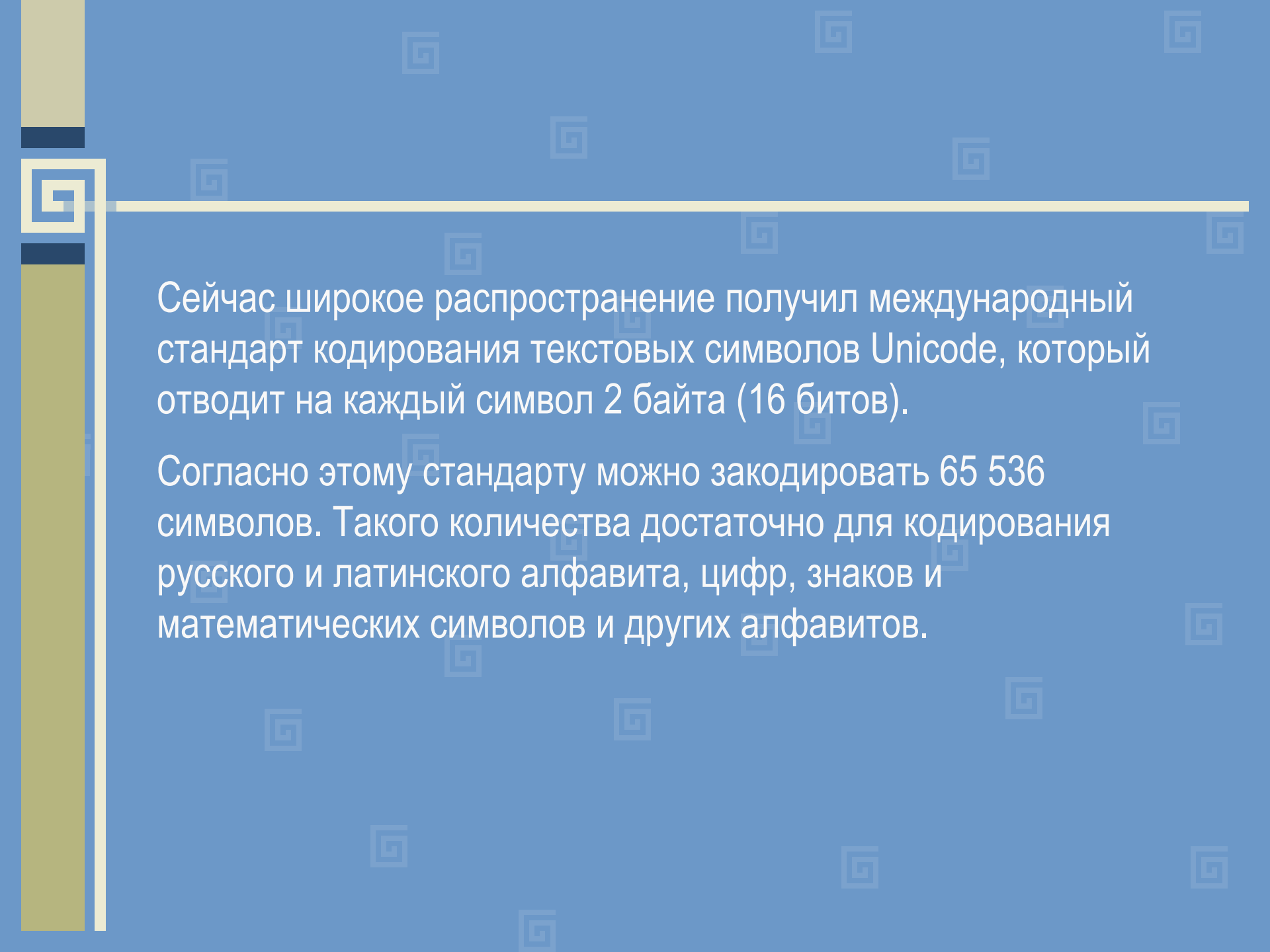
С 128 до 255 – являются национальными.

В настоящее время существуют пять различных кодировочных таблиц для русских букв :

Windows, MS-DOS, Mac, ISO, KOI-8

Тексты, созданные в одной кодировке не будут правильно отображаться в другой.

За перекодировку текста отвечают специальные программы – конверторы.



Сейчас широкое распространение получил международный стандарт кодирования текстовых символов Unicode, который отводит на каждый символ 2 байта (16 битов).

Согласно этому стандарту можно закодировать 65 536 символов. Такого количества достаточно для кодирования русского и латинского алфавита, цифр, знаков и математических символов и других алфавитов.

Задание: Закодируйте фразу

*Я изучаю информатику
успешно.*

Одним из первых стандартов кодирования русских букв на компьютерах был **КОИ8** ("Код обмена информацией, 8-битный").

—		Г	Г	└	└	┌	┌	┐	┐	┘	■	■	■	■	■
128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143
▒	▒	▒	Г	■	●	√	≈	≤	≥	nbsp	Ј	◦	²	•	÷
144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159
=		F	ё	П	Г	Г	П	Г	Е	Ц	Ц	Г	Ш	Г	Г
160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175
Г	Г	Г	Ё	Г	Г	Г	П	Г	Г	Ц	Ц	Г	Г	Г	©
176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191
Ю	а	б	ц	д	е	ф	г	х	и	й	к	л	м	н	о
192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207
п	я	р	с	т	у	ж	в	ь	ы	з	ш	э	щ	ч	ъ
208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223
Ю	А	Б	Ц	Д	Е	Ф	Г	Х	И	Й	К	Л	М	Н	О
224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239
П	Я	Р	С	Т	У	Ж	В	Ь	Ы	З	Ш	Э	Щ	Ч	Ъ
240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255