**Kódování**

je záznam informace pomocí známého, pokud možno všeobecně rozšířeného kódu (kódy jsou obvykle standardizovány, tj. zveřejněny a používány všude stejně) za účelem jejího praktického a jednoduchého uchovávání a přenosu

např.       Morseova abeceda               – optimalizováno pro přenos světelným signálem

informatika            – záznam textu, obrazu, zvuku optimalizovaný pro přenos elektrickým signálem

– dvě možné úrovně el. signálu – 0 (není) a 1 (je)

– využití číselných soustav (binární, dekadická a hexadecimální)

**bit** (zkr. binary unit, angl. bit = kousek) – zkráceně 1 b

je nejmenší jednotka informace, jeden záznam na záznamovém médiu, jeden impuls na přenosovém médiu, jedna binární číslice, nabývá hodnot 0 nebo 1

Více bitů za sebou lze považovat za binární číslo, které můžeme zapsat i pomocí jiných soustav.

např. 4 bity 1 1 0 0                = 1100 (binární číslo)

= 12 (dekadické číslo)

= 0xC (hexadecimální číslo, pozná se tak, že zápis začíná vždy 0x)

**byte** (= 8 bitů = 23 bitů) – zkráceně 1 B

 Lze pomocí něj zapsat 256 čísel od 0 (00000000) po 255 (11111111).

**méně používané**

2 byte = 1 word (slovo)

2 word = 1 double word (dvojité slovo)

**řecké předpony**

kilo (k), mega (M), giga (G), tera (T) – ve smyslu mocnin dvou

např. kilo = 210 = 1024  mega = 220 = 1048576

**Kódování informace v informatice**

veškeré informace uloženy v souborech, kódovány pomocí číslic 0 a1

Soubory s informací kódovanou známým způsobem jsou opatřeny známou příponou a zpracovávány programy, které toto kódování znají (pokud neurčíte jinak).

Existují programy, které umí zobrazit přímo jednotlivé bity jakéhokoliv souboru, např. Pspad.

**Kódování textu (znaků)**

K zakódování X různých znaků potřebujeme X různých dvojkových čísel. Základní anglická abeceda se skládá z 26 znaků, tj. k jejímu zakódování potřebujeme nejméně 26 dvojkových čísel, nejlépe 0 (00000) až 25 (11001), neboli nejméně 5 bitů, přičemž čísla 26 (11010) až 31 (11111) jsou nevyužita.

Anglická abeceda s rozlišením velkých a malých písmen se skládá z 52 znaků, tj. k jejímu zakódování potřebujeme nejméně 52 dvojkových čísel, nejlépe 0 (000000) až 51 (110011), neboli nejméně 6 bitů, přičemž čísla 52 (110100) až 63 (111111) jsou nevyužita.

Potřebujeme-li zakódovat ještě číslice od 0 do 9, interpunkční a jiná znaménka, zvláštní znaky (mezera, procento, zavináč apod.) a netištitelné znaky (tab, enter), potřebujeme již 8 bitové číslo, které nám dává 255 různých možností od 0 (00000000) po 255 (11111111). Běžně se tedy jeden znak textu zakóduje pomocí 1 bytu. Velikost souboru s neformátovaným textem kódovaným tímto způsobem je pak v bytech rovna počtu znaků textu. Např. soubor Ahoj.txt, který obsahuje slovo Ahoj, má velikost 4 B.

Existuje obrovské množství možností, jak jednotlivým existujícím znakům přiřadit jednotlivá čísla od 0 do 255. V praxi se ovšem používají standardizované převodní tabulky, nejčastěji ASCII (American Standard Code for Information Interchange) tabulka.

ASCII kódovací tabulka

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 32 |  | 60 | < | 88 | X | 116 | t | 144 | � | 172 | ¬ | 200 | Č | 228 | ä |
| 33 | ! | 61 | = | 89 | Y | 117 | u | 145 | ‘ | 173 | ­ | 201 | É | 229 | ĺ |
| 34 | „ | 62 | > | 90 | Z | 118 | v | 146 | ’ | 174 | ® | 202 | Ę | 230 | ć |
| 35 | # | 63 | ? | 91 | [ | 119 | w | 147 | “ | 175 | Ż | 203 | Ë | 231 | ç |
| 36 | $ | 64 | @ | 92 | \ | 120 | x | 148 | ” | 176 | ° | 204 | Ě | 232 | č |
| 37 | % | 65 | A | 93 | ] | 121 | y | 149 | • | 177 | ± | 205 | Í | 233 | é |
| 38 | & | 66 | B | 94 | ^ | 122 | z | 150 | – | 178 | ˛ | 206 | Î | 234 | ę |
| 39 | ‚ | 67 | C | 95 | \_ | 123 | { | 151 | — | 179 | ł | 207 | Ď | 235 | ë |
| 40 | ( | 68 | D | 96 | ` | 124 | | | 152 | � | 180 | ´ | 208 | Đ | 236 | ě |
| 41 | ) | 69 | E | 97 | a | 125 | } | 153 | ™ | 181 | µ | 209 | Ń | 237 | í |
| 42 | \* | 70 | F | 98 | b | 126 | ~ | 154 | š | 182 | ¶ | 210 | Ň | 238 | î |
| 43 | + | 71 | G | 99 | c | 127 |  | 155 | › | 183 | · | 211 | Ó | 239 | ď |
| 44 | , | 72 | H | 100 | d | 128 | € | 156 | ś | 184 | ¸ | 212 | Ô | 240 | đ |
| 45 | – | 73 | I | 101 | e | 129 | � | 157 | ť | 185 | ą | 213 | Ő | 241 | ń |
| 46 | . | 74 | J | 102 | f | 130 | ‚ | 158 | ž | 186 | ş | 214 | Ö | 242 | ň |
| 47 | / | 75 | K | 103 | g | 131 | � | 159 | ź | 187 | » | 215 | × | 243 | ó |
| 48 | 0 | 76 | L | 104 | h | 132 | „ | 160 |  | 188 | Ľ | 216 | Ř | 244 | ô |
| 49 | 1 | 77 | M | 105 | i | 133 | … | 161 | ˇ | 189 | ˝ | 217 | Ů | 245 | ő |
| 50 | 2 | 78 | N | 106 | j | 134 | † | 162 | ˘ | 190 | ľ | 218 | Ú | 246 | ö |
| 51 | 3 | 79 | O | 107 | k | 135 | ‡ | 163 | Ł | 191 | ż | 219 | Ű | 247 | ÷ |
| 52 | 4 | 80 | P | 108 | l | 136 | � | 164 | ¤ | 192 | Ŕ | 220 | Ü | 248 | ř |
| 53 | 5 | 81 | Q | 109 | m | 137 | ‰ | 165 | Ą | 193 | Á | 221 | Ý | 249 | ů |
| 54 | 6 | 82 | R | 110 | n | 138 | Š | 166 | ¦ | 194 | Â | 222 | Ţ | 250 | ú |
| 55 | 7 | 83 | S | 111 | o | 139 | ‹ | 167 | § | 195 | Ă | 223 | ß | 251 | ű |
| 56 | 8 | 84 | T | 112 | p | 140 | Ś | 168 | ¨ | 196 | Ä | 224 | ŕ | 252 | ü |
| 57 | 9 | 85 | U | 113 | q | 141 | Ť | 169 | © | 197 | Ĺ | 225 | á | 253 | ý |
| 58 | : | 86 | V | 114 | r | 142 | Ž | 170 | Ş | 198 | Ć | 226 | â | 254 | ţ |
| 59 | ; | 87 | W | 115 | s | 143 | Ź | 171 | « | 199 | Ç | 227 | ă | 255 | ˙ |

Kódy 0-31 jsou přiřazeny netištitelným znakům (např. 9 = tab, 10 = enter). Kódy 48-57 jsou přiřazeny číslicím 0-9 (kód číslice se dá tedy určit jako 48 + číslice, např. číslice 5 bude kódována jako 48+5 = 53 = 110101). Význam kódů 128-255 závisí na použitém typu písma. V českém písmu budou mít tato čísla jiný význam než v řeckém. Pokud v elektronickém textu uvidíte neznámé znaky tam, kde evidentně nemají co dělat, může to být způsobeno právě špatně vybraným písmem. Autor textu např. zakódoval český znak č pomocí českého písma jako kód 190, zatímco prohlížeč textu kód 190 dekódoval pomocí řeckého písma jako řecký znak alfa.

Osmibitové kódování znaků již dnes není dostačující, neboť je potřeba moci jedním kódem zakódovat i znaky všech národních abeced (včetně azbuky nebo asijských písem), aby je bylo možno v jednom textu kombinovat. Proto bylo vyvinuto 16tibitové kódování **Unicode**, kterým je možno zakódovat až 216 (přibližně 64 tisíc) znaků. Význam kódů 0-128 je stejný jako ve standardu ASCII. Velikost souboru s neformátovaným textem kódovaným pomocí Unicode je v bytech rovna dvojnásobku počtu znaků textu. Některé dnešní programy ještě nepodporují kódování Unicode (nebo musíte tuto možnost nejdříve ručně povolit), takže znak kódovaný pomocí Unicode zobrazí jako dva jiné znaky.

Textové soubory kódované uvedenými způsoby lze vytvářet, ukládat a otevírat pomocí textových editorů, např. Notepadu (poznámkový blok). Notepad pracuje tak, že bity jakéhokoliv souboru, který v něm otevíráte (vč. souborů, které text neobsahují, ale přesto si vyberete, že je chcete otevírat v Notepadu) rozdělí po osmi a každou osmici přeloží podle ASCII tabulky na příslušný znak, který pak zobrazí.

Příklad

Text = Ahoj!123

Kódování podle ASCII

A      h               o               j                 !        1                2                3

binárně                       = 0100 0001 0110 1000 0110 1111 0110 1010 0010 0001 0011 0001 0011 0010 0011 0011

dekadicky                   = 65    104            111            106       33         49              50              51

hexadecimálně          = 0x41               0x68          0x6F         0x6A      0x21    0x31          0x32          0x33