

II. 8. Karakterek kódolása

A számjegyek, betűk, írásjelek és egyéb jelek (összefoglaló szóval karakterek) számítógépes ábrázolására bináris kódokat használunk. Az ábrázolandó karakterek száma nem haladja meg a 256-ot, így az ábrázoláshoz elegendő 2^8 lehetőség, azaz egy bájt. Ezért a számítógépeken egy karaktert egy bájton ábrázolunk. Természetesen fontos, hogy ugyanaz a bitkombináció ugyanazt a karaktert jelentse mindenütt a világon. Ezt biztosítják a kódszabványok. A világon ma két fő számítógépes kódszabvány létezik, az egyik az EBCDIC (Extended Binary Coded Decimal Interchange Code), ami a korábban használt BCD kód kiterjesztett változata (a BCD egy hat bites, 64 különböző karaktert ábrázolni tudó, ma már nem használt kódrendszer), a másik a kiterjesztett ASCII kód (American Standard Code for Information Interchange). A kiterjesztett ASCII rendszerben 0 és 127 között a rögzített, 128 és 255 között a nemzetfüggő, átdefiniálható alakú jelek helyezkednek el, így például a magyar ékezetes betűk. A szöveg tényleges feldolgozása a kódokkal történik. (A többnyelvű írás igénye miatt valószínű ki fogja őket szorítani az unikód.)

Ezek a kódszabványok a betűkön kívül a decimális számjegyeket, néhány írásjelet, vezérlőkaraktereket és adatátviteli vezérlőjeleket is tartalmaznak.

A vezérlőkarakterek valamilyen akciót váltanak ki abból a berendezésből, amelyre küldjük őket.

Azokat a karaktereket, amelyek a nyomtatón megjeleníthetők, nyomtatható, a többi nem nyomtatható karaktereknek nevezzük.