

3. INFORMACIJOS KODAVIMAS KOMPIUTERYJE

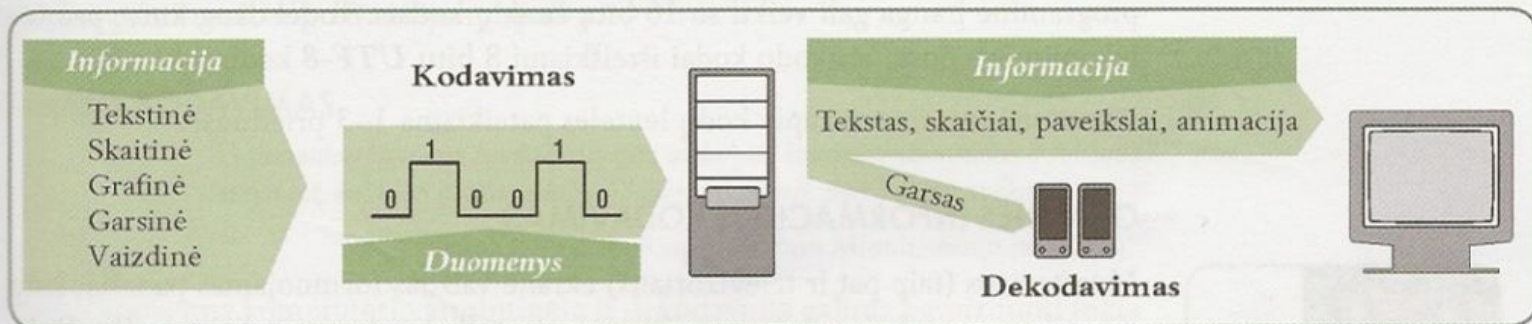
DVEJETAINIS KODAS

Panagrinėkime, kaip kompiuteriui pavyksta susitvarkyti su tokia skirtinga informacija?

Kompiuteris yra elektroninis įrenginys, todėl jis gali dirbti („maitintis“, „mąstyti“, skaičiuoti, pavaizduoti informaciją ekrane) tik su elektros signalais. Taigi norėdami, kad kompiuteris apdorotų skirtingos rūšies informaciją (tekstus, skaičius, garsus, vaizdus, animaciją), ją turime užkoduoti elektros signalais.

Žmogus informacijai užrašyti vartoja tam tikrą rinkinį ženklų – raidžių, skaitmenų, skyrybos, aritmetinių operacijų ženklų, natų. Tuos ženklus žmogus atpažįsta pagal jų formą, regėdamas (arba lytėdamas) juos. Tačiau toks informacijos pateikimo būdas netinka kompiuteriui, nes jam reikėtų techniškai sudėtingos ženklų atpažinimo ir įsiminimo sistemos. Šiai problemai spręsti buvo nutarta visų rūšių informaciją kompiuteriui pateikti tik skaitmenimis. Informacijos kodavimas kompiuteryje naudojant mums įprastą dešimtainę sistemą, t. y. visų dešimties skaitmenų reikšmes, yra sunkiai realizuojamas techniškai, be to, labai brangus uždavinys. Pasirodo, techniškai gana paprastai realizuojamas, nebrangus ir visiškai patikimas yra informacijos kodavimas kompiuteryje **dvejetainiū kodū**. Bet kokia informacija kompiuteryje užrašoma dviem skaitmenimis, tiksliau sakant, būsenomis – 0 ir 1 (nėra srovės impulso arba yra srovės impulsas; neįmagnetinta arba įmagnetinta laikmenos sritis; lazerinis diskas neatspindi šviesos arba atspindi šviesą). Yra dar vienas dvejetainio kodavimo privalumas: atlikti aritmetinius veiksmus su dvejetainiais skaičiais (sudėtį, daugybą) kur kas paprasčiau nei su dešimtainiais. Teoriškai dvejetainio kodo naudojimo idėją XX a. viduryje pagrindė vienas informacijos teorijos pradininkų Klodas Šenonas (angl. *Claude Elwood Shannon*).

Dirbdami kompiuteriu ekrane matome (arba per ausines girdime) dekoduatą informaciją. Tai informacija, specialiomis programomis iš dvejetainio kodo pavers-ta mums suprantama forma. Dekodavimas vyksta naudojant tą patį kodą, pagal kurį informacija buvo užkoduota, tik priešinga kryptimi.



Kiekvienas dvejetainio kodo skaitmuo vadinamas **bitū** pagal anglų kalboje vartojamą terminą *binary digit* (trumpinys *bit*). Galima tai interpretuoti ir kitaip: bitas – tai viena iš dviejų galimų užkoduoti sąvokų: tiesa arba netiesa; taip arba ne. Jei norime užkoduoti keturias pasaulio kryptis, tai kiekvienai jų reikia jau dviejų bitų. Tarkime, pasaulio kryptis galime užkoduoti taip:

00 – rytai, 01 – vakarai, 10 – pietūs ir 11 – šiaurė.

Trimis bitais galima užkoduoti ne daugiau kaip aštuonias skirtingas galimas reikšmes, pavyzdžiui, savaitės dienas. Viena reikšmė (111) liks nepanaudota:

| | |
|----------------|-----|
| Pirmadienis | 000 |
| Antradienis | 001 |
| Trečiadienis | 010 |
| Ketvirtadienis | 011 |
| Penktadienis | 100 |
| Šeštadienis | 101 |
| Sekmadienis | 110 |

Bitų skaičių padidinę iki keturių, galėsime užkoduoti jau šešiolika skirtingų galimų reikšmių, iki penkių – jau trisdešimt dvi skirtingas galimas reikšmes, ir t. t.:

| Bitų skaičius | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | ... | n |
|--|---|---|---|----|----|----|-----|-----|-----|------|------|------|-----|-------|
| Galimų koduoti skirtingų reikšmių skaičius | 2 | 4 | 8 | 16 | 32 | 64 | 128 | 256 | 512 | 1024 | 2048 | 4096 | ... | 2^n |

TEKSTO KODAVIMAS

Tekstui užrašyti naudojame didžiąsias ir mažąsias raides, skaitmenis, skyrybos, aritmetinių operacijų ir kitokius ženklus. Kompiuteryje kiekvienas teksto ženklas koduojamas dvejetainiu kodu naudojant specialias **koduotės** (kodų lenteles). Jose nurodomi tam tikrų ženklų dvejetainiai kodai. XX a. septintajame dešimtmetyje Amerikos nacionaliniame standartų institute buvo sukurta vadinamoji **ASCII** (angl. *American Standard Code for Information Interchange*) koduotė, kuri paplito daugelyje šalių. Dauguma dabartinių koduočių buvo sukurtos papildant (arba pakeičiant) **ASCII** kodų lentelę.

Lietuvoje priimta 8 bitų tarptautinio standarto koduotė **ISO 8859-13**. Ši koduotė nepriklauso nuo operacinės sistemos. Lietuvos standartai apibrėžia koduotes ir atskiroms operacinėms sistemoms: **DOS-775**, **Windows-1257**. Naudodamiesi šiomis koduotėmis, galime rašyti ne tik lietuviškai, bet ir anglų, danų, estų, latvių, lenkų, norvegų, suomių, švedų ir vokiečių kalbomis. Konsorciumas *Unicode* sukūrė 16 bitų koduotę **unikodas**, kurioje nurodyti visų pasaulio kalbų (jų dabar yra daugiau kaip 2 200) svarbiausių rašto ženklų dvejetainiai kodai. Palaipsniui visame pasaulyje pereinama prie šios koduotės, tačiau dar ne visa aparatinė ir programinė įranga gali veikti su 16 bitų ženklų kodais. Todėl daug kur, ypač telekomunikacijose, unikodo kodai išreiškiami 8 bitų **UTF-8** kodų sekomis.

Išsamesnė informacija apie kodų lenteles pateikiama 1–3 prieduose.

GRAFINĖS INFORMACIJOS KODAVIMAS

Monitoriaus (taip pat ir televizoriaus) ekrane vaizdas formuojamas panašiai kaip mozaika – iš mažiųjų stačiakampių spalvotų taškelių, kurie vadinami **pikseliais**. Juos galima apžiūrėti lupa.

Kompiuteryje skirtingo tipo vaizdai koduojami skirtingai.

Taškinių (rąstrinės) grafikos paveikslai paprastai gaunami perkeliant vaizdą į kompiuterį skaitmeniniais fotoaparatais, skaitytuvais ir kitais panašiais prietaisais, taip pat piešiant paveikslus piešimo programomis, pavyzdžiui, programa *Piešimas* (*Paint*). Tokie paveikslai dažniausiai laikomi failuose su prievardžiais **BMP**, **PNG**, **GIF**, **TIF**, **JPEG** ir kt. Į failą iš eilės įrašoma paveikslo kiekvieno taško spalva.



Taškinės grafikos paveikslo vaizdo kokybė priklauso nuo monitoriaus ekrano **skiriamosios gebos** (ekrane esančių taškų skaičiaus): kuo taškų daugiau, tuo jų matmenys mažesni ir vaizdas mūsų akiai atrodo vientisesnis.

| | | | | |
|----|----|----|----|----|
| 1 | 2 | 3 | 4 | 5 |
| 6 | 7 | 8 | 9 | 10 |
| 11 | 12 | 13 | 14 | 15 |

Paveikslo taškų eilės numeriai

Brėžiniams, schemoms, šriftams koduoti patogiau naudoti **vektorinę grafiką**. Šios grafikos paveikslo vaizdas ekrane taip pat formuojamas iš taškų, tačiau kompiuterio atmintinėje laikoma ne kiekvieno taško spalva, o atkarpų (vektorių), kuriomis galima suskaidyti vaizdą, ilgis, kryptis ir spalva. Dėl vektorinės grafikos vaizdo kodavimo ypatumų, vaizdą didinant ar mažinant, jo kokybė nesikeičia. Tokie paveikslai braižomi vektorinės grafikos programomis (vadinamosiomis **braižyklėmis**), pavyzdžiui: *OpenOffice.org Draw*, *CorelDraw*. Populiariausi vektorinės grafikos failų formatai yra **CDR, EPS, WMF**.

Smalsiems

Daikto spalvą lemia nuo jo į akį atspindimi spinduliai. Dienos šviesoje balti atrodo daiktai, kurie atspindi visų rūšių šviesos spindulius. Ir atvirkščiai, juodi beveik visus spindulius sugeria. Daiktai matomi spalvoti, kai jie nevisiškai sugeria spindulius.

Spalvoto vaizdo ekrane kiekvieno pikselio spalva gaunama skirtingomis proporcijomis maišant tris pagrindines spalvas (raudoną, žalią ir mėlyną). Toks spalvų modelis vadinamas **RŽM** (spalvų pavadinimų pirmųjų raidžių santrumpa: *R* – raudona, *Ž* – žalia, *M* – mėlyna, angl. *RGB*). Lenteleje nurodytos aštuonios spalvos, gautos maišant tris grynąsias spalvas (jei pagrindinė spalva yra derinyje, atitinkamame langelyje pažymėtas vienetą, jei nėra – nulis).

| Pikselio spalva \ Pagrindinė spalva | Juoda | Mėlyna | Žalia | Žydra | Raudona | Purpurinė | Geltona | Balta |
|-------------------------------------|-------|--------|-------|-------|---------|-----------|---------|-------|
| Raudona | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| Žalia | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| Mėlyna | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |

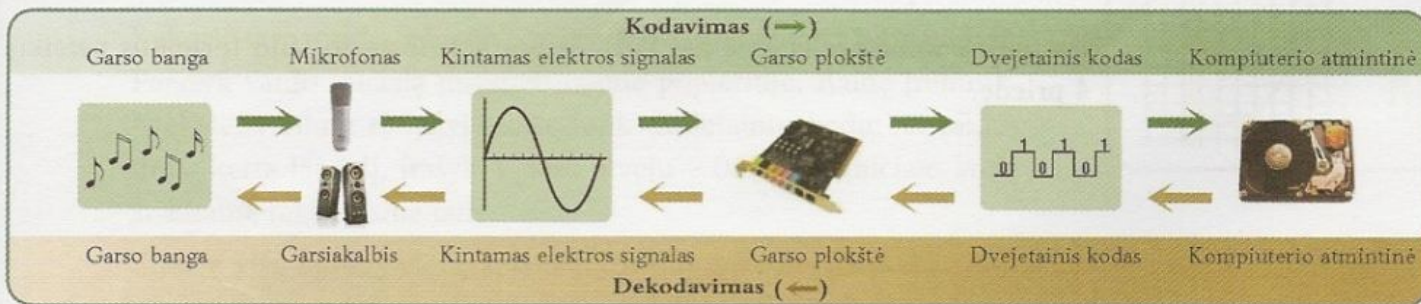
Jei vaizdas būtų sudarytas tik iš šių aštuonių spalvų, tai kiekvienam vaizdo pikseliui laikyti kompiuterio atmintinėje reikėtų 3 bitų. Tikroviškam vaizdui reikia daug spalvų. Jos gaunamos nurodant kiekvienos pagrindinės spalvos intensyvumą intervale [0; 255]. Pavyzdžiui, spalvai (■) gauti pagrindinių spalvų intensyvumas turi būti toks: raudonos – 100, žalios – 150, mėlynos – 200. Kiekvienos pagrindinės spalvos reikšmei užkoduoti kompiuterio atmintinėje reikia 8 bitų. Iš viso pasirinktai spalvai užkoduoti reikia 24 bitų. Vaizdą koduojant 24 bitais galima gauti maždaug 16 milijonų spalvų.

GARSO KODAVIMAS

Aš sumaniau išgirstus žodžius sugauti, uždaryti į švininius vamzdžius ir ten juos laikyti tol, kol man patiks, o kai atidengsiu dangtelį, kad žodžiai suskambėtų.

Rudolfas Erichas Raspé „Barono Miunhauzeno nuotykių“

Garso kodavimą kompiuterio atmintinėje ir dekodavimą galima pavaizduoti tokia schema:

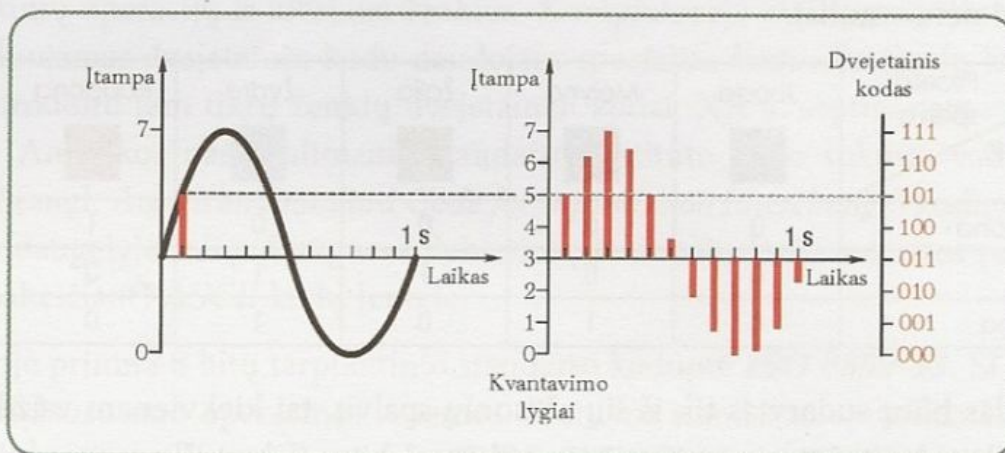


Oru sklindantis garsas, patekęs į mikrofoną, pakeičiamas elektros signalu, kurio įtampa ir dažnis kinta laike. Garso plokštėje šis signalas **skaitmėninamas** – verčiamas dvejetainiu kodu ir po to įrašomas į kompiuterio atmintinę.

Skaitmeninimas vyksta periodiškai matuojant signalo įtampą. Matavimų rezultatas – trumpi impulsai, kurie atkartoja pradinio signalo formą (paveikslėlyje žymimi raudonais stulpeliais). Matavimų skaičius per sekundę vadinamas **išrankos dažniu**. Paveikslėlyje vaizduojama dvylika matavimų, t. y. išrankos dažnis yra 12 Hz (hercų).

Intervalas tarp įtampos mažiausios ir didžiausios reikšmių dalijamas į vienodus lygius (atkarpas), kurie vadinami **kvantavimo lygiais**. Kiekvienam kvantavimo lygiui priskiriamas dvejetainis kodas. Paveikslėlyje pavaizduoto signalo įtampos mažiausia reikšmė yra nulis voltų, didžiausia – septyni voltai. Šiame intervale parinkti aštuoni kvantavimo lygiai, todėl galimiems impulsams koduoti reikia trijų bitų. Jei gretimi įtampos lygmenys skirtųsi puse volto, tada iš viso būtų 16 kvantavimo lygmenų ir impulsams koduoti reikėtų jau keturių bitų.

Kiekvienam impulsui priskiriamas artimiausias jo vertę atitinkančio kvantavimo lygio dvejetainis kodas. Visų matavimų rezultatai (dvejetainiai kodai) įrašomi į kompiuterio atmintinę.



Šiuolaikinės garso plokštės geba išskirti 65 536 signalo reikšmes, t. y. koduoti kiekvieną impulsą 16 bitų. Kuo didesnis kvantavimo lygių skaičius ir išrankos dažnis, tuo geresnė garso kodavimo kokybė. Garso dažnis, kurį skiria žmogaus ausis, gali kisti nuo 20 iki 20 000 Hz. Norint, kad skaitmeninis garsas būtų artimesnis gyvam, reikia skaitmeninimui naudoti dar didesnę išrankos dažnį. Todėl garsui skaitmeninti naudojamas išrankos dažnis viršija 40 kHz.

Vienas populiariausių garso failų formatų yra **WAV**. Jis leidžia įrašyti garsą didžiausiu tikslumu. **MIDI** formatas skirtas instrumentinei muzikai. Labai populiarius yra **MP3** formatas, kokybiškai koduojantis muziką. Šiuo formatu įrašoma tik ta muzikinės informacijos dalis, kurią skiria žmogaus ausis, kita informacija atmetama. **MP3** formatas sumažina informacijos kiekį tūkstančius kartų, todėl kelių minučių trukmės melodija ar daina lengvai gali būti siunčiama net elektroniniu paštu.

Išsamesnė informacija apie kai kuriuos grafikos ir garso failų formatus pateikta 4 priede.