

PRESEK

List za mlade matematike, fizike, astronome in računalnikarje

ISSN 0351-6652

Letnik 30 (2002/2003)

Številka 4

Strani 208-212

Jože Pahor:

JE SLIKA RES VEČ KOT TISOČ BESED

Ključne besede: računalništvo, shranjevanje slik.

Elektronska verzija: <http://www.presek.si/30/1522-Pahor.pdf>

© 2003 Društvo matematikov, fizikov in astronomov Slovenije

© 2010 DMFA - založništvo

Vse pravice pridržane. Razmnoževanje ali reproduciranje celote ali posameznih delov brez poprejšnjega dovoljenja založnika ni dovoljeno.

JE SLIKA RES VEČ KOT TISOČ BESED

Pravijo, da pove vsaka slika več kot tisoč besed. Pa je to res? Bi znali opisati obraz razbojnika Cefizlja tako natančno, da bi ga butalski policaj prepoznal na prvi pogled?

Danes, v dobi računalnikov, se pogosto pogovarjamo tudi z njimi. Kako shraniti v računalniku besedilo, kako shraniti sliko, da jo bo računalnik na zahtevo spet lahko pokazal na zaslonu? Besedilo nam bo delalo bržkone manj preglavic kot slika, zato začnimo z njim.

Pri pogovoru z računalnikom se moramo prilagoditi njegovim lastnostim. Računalnik pomni, misli in govori po dvojiško. V vsaki svoji spominski celici ustvari lahko le stanje, ki pomeni 0, in stanje, ki pomeni 1. Nekaj dvojiške govornice že poznamo, namreč dvojiški zapis števil. Niz 101 pomeni 5, niz 101010 pa 42. Shranimo obe števili v zaporedni spominski celici. Celotni niz 101101010 lahko preberemo na različne načine, saj ne vemo, kje je razmejitev. Rezultat je lahko spet 5 in 42, morda dobimo 11 in 10 (1011 in 01010) ali kaj tretjega. Da ne bo zmede, se je treba vnaprej dogovoriti, kakšna števila bomo uporabljali in kako jih bomo skladiščili.

Najprej si omislimo večje predalčke, zloge oziroma byte, ki jih sestavlja po osem celic oziroma bitov. Če namenimo vsakemu številu po dva byta, lahko zapisujemo števila od 0 do 65535. Če je dvojiško število vredno manj kot 16 bitov, ga dopolnimo spredaj z ničlami:

0000 0000 0000 0101 je po naše 5.

Kadar bi radi pisali še večja števila, se vnaprej dogovorimo za tri ali štiri osembitne predale oziroma byte. S štirimi byti pridemo do $2^{32} - 1 = 4\,294\,967\,295$.

Zdaj poskusimo z besedilom. Slovenska abeceda pozna 25 črk. Da bi jih zapisali po dvojiško, potrebujemo 25 različnih kombinacij ničel in enic ter dogovor, katero črko pomeni vsaka od teh kombinacij. S petbitnimi nizi zmoredemo 32 različnih kombinacij, torej dovolj. Lahko bi začeli takole: 00001 pomeni **a**, 00010 pomeni **b**, 00011 pomeni **c** in tako naprej. Niz 00000 smo prihranili, da napove presledek med dvema besedama. Šestim preostalim nizom priredimo piko, vejico in še kaj. Žal bi bil tak dogovor uporaben le znotraj Slovenije. Marsikje šumnikov ne potrebujejo, manjkale pa bi jim oznake za njihove črke. Poleg tega petbitni kosi tudi niso pisani na kožo računalniku, ki ljubi byte, osembitne predalčke. Zato

bomo uporabili po en byte za vsak znak abecede. Pri 256 različnih nizih je prostora za vse znake, ki jih je nekdanj zmožgal pisalni stroj. Zapisati znamo velike in male črke, pa številke, ločila. Nekateri znaki so ukazi: pojdi v novo vrsto, pozvoni, vrni se. Poseben znak pove, da je konec našega besedila.

Dogovor je zapis *American Standard Code for Information Interchange*, ali krajše ASCII, torej zapis znakov za informacijsko izmenjavo po ameriško¹. Sporočilo PRESEK se po tem dogovoru napiše 80 82 69 83 69 75. Ustrezni dvojiški zapis spravimo v 48 spominskih celic². Naši šumniki, veliki in mali, nosijo precej visoke številke, ki pa niso enotno standardizirane.

Tudi debelo knjigo lahko varčno zapišemo. Prva knjiga zadnje čase priljubljenega Gospodarja prstanov šteje 535 strani. Vsaka polna stran ima 35 vrstic, dolgih po 65 znakov. Vsak znak zavzema po osem spominskih celic. To terja skupaj 9 milijonov in nekaj nad 700 tisoč spominskih celic. Ker šteje po osem celic za en byte, je zapis dolg nekaj manj kot 1,2 MB (mega bytov)³. To je na videz mnogo. Vendar pa na današnje CDje lahko shranimo 650 ali 700 MB. Na tak CD bi lahko zapisali preko 500 Gospodarjev prstanov. Na knjižnih policah bi to pomenilo vsaj dvajset metrov ali nekaj sto kilogramov. Trdi disk v samem računalniku je še zmogljivejša shramba od CDja. Današnji trdi diski običajno drže 40 000 do 80 000 MB. Z besedili potemtakem ne bo težav.

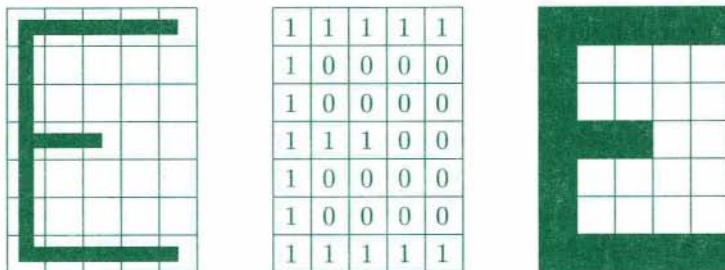
Na koncu knjige Gospodar prstanov najdemo nekaj zemljevidov. Kako shraniti sliko? Bo tudi v deželi računalnikov veljala več kot tisoč besed?

Poskusili bomo spremeniti preprosto sliko v niz iz enic in ničel. Naj bo to kar slika črke E. Prekrijmo sliko z mrežo, ki šteje pet polj v širino in sedem polj v višino.

¹ Novejši je UNICODE, kjer znaki za posamezne črke zajemajo po 16 spominskih celic. To pomeni, da razpolaga s preko 65 500 nizi, zato so nabori znakov lahko veliko daljši. Tako lahko v enem naboru zapišemo vse posebne znake različnih jezikov (šumniki, akcenti, grška abeceda), lahko pa tudi zapišemo pisave z veliko znaki (arabske, kitajske pismenke itd.).

² Oznaki za P in R nista sosednji števili, ker je med njima v angleški abecedi črka Q.

³ V računalništvu je $1k=1024$ in $1M = 1024 \times 1024 = 1\,048\,576$.



Slika 1. Levo: Črka E sedi v nekaterih poljih. Sredina: Vpišemo enice v polja, kjer najdemo košček črke E, v prazna polja pa 0. Desno: Pri obnavljanju slike počrtnimo polja, označena z 1.

V nekaterih poljih najdemo košček črke, druga polja so spet prazna. Otapa jmo sliko polje za poljem in označimo polna polja z znakom 1, prazna pa z znakom 0. Celoten niz ničel in enic pove vse o sliki:

11111 10000 10000 11100 10000 10000 11111 .

Zaradi preglednosti smo uporabili presledke med zapisi posameznih vrstic. Računalnik seveda presledkov ne pozna. Ko bere celotni niz brez presledkov, torej

11111100001000011100100001000011111 ,

mora vedeti, koliko polj si sledi v vsaki vrsti in koliko je vseh vrst. Da ne bo zmede, namenimo vsakemu od obeh podatkov po osem spominskih celic, bitov, torej po en byte, in jih postavimo na začetek slikovnega zapisa kot 00000101 00000111, kar je 5 in 7 po naše. Ker poznamo pravila, ki so nas vodila pri zapisu slike, iz celotnega zapisa ne bo težko znova obnoviti slike. Prvih osem polj pove vodoravno razsežnost slike, drugih osem polj pa navpično. Produkt obeh razsežnosti, ki je pri nas 35, pove število nadaljnjih polj, ki jih je potrebno prebrati, da zremo vse o sliki. Celoten zapis slike je zdaj dolg 51 bitov. Pri uporabi znakov ASCII smo za zapis črke E potrebovali le 8 bitov.

Z obnovljeno sliko nismo zadovoljni, saj ne posnema dovolj natančno izvornika. Očitno je mreža pregroba. Polj v vrstah naj bo zato nekaj sto ali celo tisoč, povečajmo pa tudi število vrst. Oba podatka povemo, preden opišemo vsebino slike. Po en byte za vsakega bo premalo, naj bodo po trije.

Večina slik pozna poleg belin in črnin tudi vrsto sivih odtenkov. Skalo bo treba razširiti. Naj ostane popolna belina 0, z 255 označimo popolno črnino⁴. Vmes je prostora za 254 odtenkov sivine. To je sicer več, kot jih lahko naše oko razlikuje, vendar je računalniku pri srcu razdelitev na byte. Zapis take slike bo seveda daljši. Vsako polje terja byte in ne bit.

Slika pa ljubi barve. Kako jih spremeniti v ničle in enice? Ugotovili so, da je mogoče katerokoli barvo sestaviti iz treh osnovnih barv: rdeče, zelene in modre. Če ni ne rdeče ne zelene in ne modre, vlada popolna tema. Zapišemo jo z 0, 0, 0. Vse tri barve, vsaka s polno vrednostjo 255, ustvarjajo videz bele barve, ki jo zapišemo kot 255, 255, 255. Povsem rdeče polje ima oznake 255, 0 in 0, seveda vse zapisano z ustreznim nizom štirindvajsetih ničel in enic (sliki na III. strani ovitka).



Slika 2. Belina prehaja postopno v črnino prek vrednosti od 0 do 255.

Zapis slik postaja vse daljši in daljši. Ni težko izračunati, kako dolg bo zapis barvne slike, široke 1000 točk in visoke 800 točk. Približno tak zapis uporabljamo dandanes na zaslonih svojih računalnikov. Zapis takšne slike bo dolg 2 400 000 bytov, kar je skoraj dvakrat toliko kot vse besedilo Gospodarja prstanov. Zdaj poiščimo odgovor na vprašanje, ki smo si ga zastavili. Če predpostavimo, da je povprečna slovenska beseda dolga šest znakov – všteli smo tudi presledek – potem je v Gospodarju prstanov okoli 200 000 besed. Vsaj toliko pa je lahko vredna tudi ena sama slika.

V tehniki in znanosti uporabljamo načrte in grafe. Črte zasedajo le malo prostora; večina slike je bela. Ko načrt ali graf na opisani način spremenimo v niz enic in ničel, bodo prevladovale ničle. Ob veliki množici ničel bo zapis pošastno dolg. Poskusimo z drugačnim zapisom. Spet se lotimo risanja črke E tako, da sledimo naslednjemu nizu ukazov:

Dvigni pero nad papir – za vsak slučaj.

Premakni se v koordinatno izhodišče – potovanje bomo začeli od doma.

Premakni se do točke 2,1 – zdaj smo pripravljeni za risanje.

Spusti pero do papirja – na papirju začni pika.

Premakni se do točke 2,8 – dobimo navpično črto.

Premakni se do točke 7,8 – kar dobro nam gre od rok.

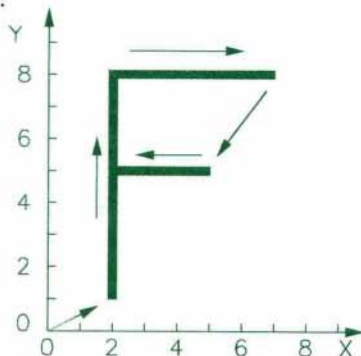
Dvigni pero – del risanja je končan.

Premakni se do točke 5,5.

Spusti pero do papirja.

⁴ V resnici uporabljamo 0 za črnine (ni svetlobe) in 255 za beline (polna svetloba).

Premakni se do točke 2,5 – dobili smo F, ki mu bomo dodali še nogo in ga spremenili v E.



Slika 3. Kako potujemo med risanjem črke E.

Zelo podobno so lahko navodila zapisana v računalniku, le zapis je okrajšan, npr.:

PU (**p**en **u**p = dvigni pero),

PA 0,0 (**p**en **a**bsolute to 0,0 = premik v točko 0,0),

PA 2,2,

PD (**p**en **d**own = spusti pero ali pa **p**ero **d**ol),

PA 2,8 in tako naprej do srečnega konca, ko spravimo pero spet domov.

Poleg ravnih črt so pa tu tudi krive. Te narišemo tako, da jih nadomestimo z množico kratkih daljic. Krog postane, recimo, 40-kotnik.



Slika 4. Petkotnik, desetakotnik in dvajsetkotnik, ki ga vidimo že kot krog.

Opisani način poveljevanja je bil prilagojen risalniku (plotterju), instrumentu, ki je po ukazih računalnika risal načrte pred dvajsetimi leti, torej pred rojstvom sodobnih tiskalnikov. Poleg dveh elektromotorčkov, ki sta vozila pero sem in tja pa papirju, je premogel ravno dovolj vgrajene računalniške pameti, da je razumel nekaj potrebnih ukazov. Risalniki so izumrli, zapis pa se uporablja še danes.

Načinov slikovnega zapisa je še mnogo. Preveč, da bi jih spravili v Presek.

Jože Pahor