

Aktivnost 6

Potapljanje ladjic: Iskalni algoritmi

Namen

Otroci spoznajo, da je čas, ki ga potrebujemo za iskanje določenega podatka, odvisen od tega, kako si podatke organiziramo. Najslabše je, če so neurejeni. Veliko boljše je, če so urejeni, saj lahko tedaj uporabljamo dvojiško iskanje, ki so ga spoznali v prejšnji aktivnosti. Tretji način organizacije, ki jim je verjetno tuj, vendar je najbolj učinkovit, so razpršene tabele.

Z razmišljanjem o tem, kako število potrebnih ugibanj narašča s številom ladij, gradijo intuicijo, ki se skriva za ocenjevanjem zahtevnosti algoritmov.

Trajanje

Ena ura, lahko pa tudi dve, če na koncu izvedeš tudi motivacijsko igro z začetka naslednje aktivnosti.

Potrebščine

- kartoni z velikimi števkami
- vrečka bombonov (po štiri ali pet bombonov za vsako ponovitev igre)
- listi z imeni in števkami (ena kopija za otroke in še ena zate)

Za vsak par otrok

- po eno polo 1A, 1B, 2A, 2B, 3A, 3B. Pripravi tudi nekaj rezervnih pol (1A', 1B', 2A', 2B', 3A', 3B') za pare, ki bi si pomotoma pokazali poli.



Uvodna motivacija

Iskanje števila

1. Izberi petnajst otrok in jih postavi v vrsto pred tablo. Vsak naj naključno izvleče enega od kartonov s številkami.
2. Karton naj drži ob sebi tako, da sošolci, ki bodo ugibali števila, ne vidijo številke, pač pa so pripravljene obrniti karton (in ga obdržati obrnjenega), če jim kdo to plača z bombonom.
3. Izberi si enega od števil, ki so ga izvlekli otroci. Povej, da potrebuješ prostovoljca, ki ti bo moral izročiti karton s tem številom.
4. Prostovoljcu daj štiri ali pet bombonov, s katerimi bo plačeval odkrivanje kartonov. Ko odkrije pravi karton, bo lahko obdržal bombone, ki mu ostanejo.
5. Igra poteka tako, da prostovoljec izroča bombone otrokom s kartoni, dokler ne odkrije pravega ali pa mu zmanjka bombonov (kar je verjetneje).
6. To lahko večkrat ponoviš, da bodo videli, kako težka je naloga. Nalogo vsekakor ponavljaj, če je imel pri otrok srečo. Otroci, ki držijo števila, naj vsakič izberejo nove kartone.

Če so otroci dovolj stari za osnove verjetnostnega računa, lahko tu skupaj razmislite o tem, kako verjetno je, da bo sploh obdržal kakšen bombon (to se zgodi v $4/15$ poskusov) in koliko bombonov obdržijo v poprečju. ($1/15 \times 4 + 14/15 \times 1/14 \times 3 + 14/15 \times 13/14 \times 1/13 \times 2 + 14/15 \times 13/14 \times 12/13 \times 1/12 \times 1 = 1/15 \times (4+3+2+1) = 10/15 = 2/3$ bombona.)

1. Otroci naj vrnejo kartone. Sam naključno izberi petnajst kartonov in jih razdeli otrokom pred tablo tako, da bodo številke urejene po vrsti. Otrokom povej, da so kartoni zdaj urejeni.
2. Spet izbereš otroka, ki bo ugibal število. Če bo ravnal preudarno, lahko vedno pride do iskane številke že s tremi bomboni. Otrokom po potrebi povej, da jim ni potrebno žrtvovati bombona za karton s pravo številko, če vedo, da je pravi, tudi ne da bi ga pogledali.
3. Če otroci ne odkrijejo trika ali če ga ne razumejo vsi, lahko tudi igro z urejenimi števili večkrat ponoviš, spet vsakič z novimi kartoni.

Trik, ki ga morajo odkriti otroci, je bisekcija oz. binarno iskanje, s katerim v vsakem koraku razpolovijo množico števil: najprej pogledajo številko, ki jo ima srednji otrok. Če je prevelika, nadaljujejo z levo, če premajhna z desno polovico. Vsako polovico razpolavljajo naprej, dokler ne naletijo na iskano številko.

Otrokom povej, da je pri petnajstih številih vedno možno najti pravo z največ tremi bomboni. Najprej odkrijemo osmo število. Morda bomo imeli srečo in je že pravo; v tem primeru smo porabili samo en bombon. Če je preveliko, je pravo število med sedmimi na levi, če premajhno, med sedmimi na desni: v vsakem primeru pa nam ostane le še sedem števil. Med temi sedmimi vedno vprašamo po četrtem. Če imamo

srečo, je pravo in smo porabili dva bombona. Če ni, pa je preveliko in se pravo število skriva med tremi števili na levi, ali pa premajhno in je pravo število med tremi na desni. V vsakem primeru pa nam po dveh porabljenih bombonih ostanejo le še tri števila. S tretjim bombonom vprašamo po srednjem številu izmed preostalih treh. Če je število pravo, ga imamo. Če ni, pa vemo, da je pravo število na levi ali na desni in ga lahko vzamemo, ne da bi ga pogledali.

Povedano prikaži. Izberi otroka, za katerega veš, da je razumel recept (če bo imel težave, mu pomagaj). Števila, ki naj ga ugiba, si ne izberi naključno: izberi število, ki je na sedmem mestu. Skladno z navodili mora otrok izbrati osmi karton. Število bo preveliko: vsi otroci od vključno osmega otroka naprej niso več kandidati, zato naj počepnejo. Nato bo izmed ostalih sedmih izbral četrtega. Ta bo prevelik, zato naj počepne, hkrati z njim pa vsi otroci levo od njega. Izmed ostalih treh bo izbral srednjega (to je, šesti otrok). Številka bo premajhna, zato počepne on in otrok na levi. Edini, ki po tem še stoji, je sedmi otrok.

Vprašaj otroke, ali podoben način iskanja uporabljajo tudi kje drugje, recimo, ko iščejo ime v abecednem seznamu imen ali pa knjigo v knjižnici. Ugotovili bodo, da so ta način iskanja pravzaprav že uporabljali, vendar tega niso vedeli.

Morda bodo odkrili še drug trik: če je številka majhna, se ne splača deliti tako, da začnemo s srednjim, temveč začnemo nekoliko bolj levo in obratno. Tudi, ko v telefonskem imeniku iščemo kakega Demšarja, ga ne odpremo na sredini, temveč bolj na začetku. Tudi ko na knjižni polici iščejo Grafenauerja, ne začnejo na sredini, temveč levo. Če se kateri izmed otrok domisli te rešitve, povej, da se uporablja tudi to, vendar bomo tu delali drugače. Če otroci sami ne bodo dobili te ideje, pa jih ni potrebno begati z njo.

Iskanje imena

1. Primerno izberi tri otroke in jim daj sezname s številkami in imeni. Povej, da so to telefonske številke pri nekem novem ponudniku mobilne telefonije, ki ima manj uporabnikov in so zato številke samo štirimestne.
2. Drugi otroci naj številke ne vidijo, prav tako naj izbrani trije otroci ne vidijo, da imajo različne liste: pri enem so urejeni po številkah, pri drugem po imenih, pri tretjem pa so naključno pomešani.
3. Otroke s seznamami izzovi, kdo prej poišče Božidarjevo telefonsko številko, številko, ki jo ima Vesna, Martina, Jure... Če se izogneš imenom, ki se pojavljajo na vrhu neurejenega seznama in seznama urejenega po številkah, bo navadno (ali vedno) najhitrejši otrok, čigar seznam je urejen po imenih.
4. Zdaj sprašuj po številkah: kdo ima telefonsko številko 7007? Kaj pa 5371, 9131, 4223? Pri vseh številkah bi moral biti najhitrejši otrok, čigar seznam je urejen po številkah, izjema pa je zadnja, ki se slučajno pojavi na začetku neurejenega seznama.
5. Igro lahko nadaljuješ, dokler otroci ne uganemo, v čem je trik.

Pogovor

Igra spet kaže, koliko nam pri iskanju pomaga, če so reči urejene.

Otrokom povej, da je na seznamih, s katerimi so se igrali, 336 telefonskih števil, kar je primerljivo s telefonskih imenikov v mobitelih njihovih staršev. Kako, da mobitel tako hitro poišče številko? So sezname, ki jih ima urejeni? (Po čem? Lahko ima tudi dva seznama oz. z nekim trikom, ki ga uporabljajo računalnikarji, seznam, ki se ga da pogledati na dva načina, tako da je enkrat urejen po imenu in enkrat po številki.)

Računalniki so hitri. Bi lahko iskali tudi po neurejenih seznamih?

To bi bilo prepočasno. Predstavljajte si, da bi imeli v trgovini 10.000 različnih produktov. Ko pridemo na blagajno, blagajničarka bere kodo izdelkov in računalnik mora v seznamu poiskati tistega z ustrezno številko. Če seznam ne bi bil urejen, bi včasih hitro našel pravo (tako kot učenec, ki je slučajno takoj dobil ime osebe s telefonsko številko 4223), če bi imel smolo, pa bi moral pregledati 10000 števil. V poprečju bi jih pregledal 5000 in če bi se z vsako ukvarjal le tisočinko sekunde, bi mu vsak izdelek v poprečju vzel pet sekund. Bi hodili v trgovino s tako počasno blagajno?

Blagajna išče kodo izdelkov. Kdo na tem svetu še veliko išče? Google! Google ve za kako milijardo strani, vendar mora tisto, ki jo iščemo, najti v trenutku.

Vprašaj otroke, če si lahko izmislijo za iskanje še kaj boljšega kot je razpolavljanje urejenih seznamov. Najbrž si ne bodo. Povej, da bodo zdaj odigrali potapljanje ladjic na tri načine: prva dva bosta podobna iskanju v neurejenem seznamu in iskanju z razpolavljanjem, tretja pa bo drugačna in še hitrejša od razpolavljanja.

Preden začneš igro, se v pogovoru prepričaj, da učenci niso pozabili, kako deluje razpolavljanje.

Potapljanje ladjic

Linearno iskanje po neurejeni tabeli

1. Otroci naj se razdelijo v pare. En otrok iz para dobi list 1A, drugi pa 1B. Listov si ne smejo kazati med seboj! (Lista 1A' in 1B' sta namenjena otrokom, ki bi radi odigrali več iger ali pa so ponesreči pokazali svoj list nasprotniku. Ostali listi so za naslednje igre.)
2. Vsak od otrok si izbere in obkroži eno od ladij v gornji vrsti in pove nasprotniku njeno številko. Nasprotnikova naloga je zadeti to ladjo.
3. Otroka izmenično poskušata zadeti nasprotnikovo ladjo tako, da govorita črke pod ladjami. Nasprotnik vedno pove številko ladje, ki je bila zadeta. Učenec, ki je streljal, naj si številko zapiše v ladjo.
4. Zmaga otrok, ki prvi zadane nasprotnikovo ladjo. Tudi po zmagi pa otroka igro odigrata naprej, dokler ni potopljena še zmagovalčeva ladja.
5. Otroci naj preštejejo, kolikokrat so morali ugibati, da so zadeli nasprotnikovo ladjo. Štejejo naj vse napačne strele in tudi pravih.

Otroke spodbudi, naj bodo čim učinkovitejši (če je to potrebno – navsezadnje gre za tekmovanje!)

Recimo, da ima Ana list 1A in izbere ladjo 4917. To številko pove svoji nasprotnici Bernardi. Bernarda ugiba tako, da pove črko, recimo T. Ana odgovori tako, da pove številko ladje T, torej 4932. Bernarda si v ladjo nad črko T zapiše to številko. Anina ladja bo potopljena, ko bo Bernarda izbrala črko D.

Pogovor po igri

1. Vsak učenec naj pove, kolikokrat je moral ugibati. Številke zapiši na tablo.
2. Učence vprašaj, kolikšno je najmanjše in največje možno število potrebnih ugibanj. (Odgovor je 1 in 25, če predpostavimo, da otroci nikoli ne ciljajo iste ladje večkrat.)

Bisekcija po urejeni tabeli

Parom razdeli liste 2A in 2B. Opozori jih, da so ladje urejene po številkah. Ponovno naj odigrajo igro. Pravila so enaka kot prej, igra pa je seveda lažja.

Pogovor po igri

1. Otroci naj spet sporočijo rezultate. Napiši jih na tablo, tako da jih bo mogoče primerjati s prejšnjimi.

2. Kako so iskali tisti, ki so dosegali najboljše rezultate?
3. Katero ladjo se plača ciljati najprej? S katero ladjo nadaljujemo?
4. Kolikokrat je potrebno ugibati, če uporabljamo takšno strategijo?

Razpršene tabele

Razdeli liste 3A in 3B. Če boš uporabil rezervne pole, moraš pri tej različici igre poskrbeti, da zamenjaš poli obema učencema iz para: poli 3A' in 3B se ne ujemata.

Navodila so enaka kot v prejšnji igri, le ladje so postavljene drugače: iz številke ladje je mogoče razbrati, v katerem stolpcu je. Recimo, da imamo ladjo s številko 2345. Seštejemo $2+3+4+5$, dobimo 14. Zadnja številka vsote je 4, zato se ladja 2345 nahaja stolpcu s številko 4.

Pogovor po igri

1. Analiziraj rezultate.
2. Katere ladje je najlažje najti? (Tiste, ki so same v svojem stolpcu.) Katere je najtežje najti? (Tiste iz stolpcev z veliko ladjami.)
3. Recimo, da bi sestavljali svojo polo in bi hoteli narediti čim težjo igro. Kaj bi morali storiti? (zmislili bi si takšne številke, da bi bile vse ladje v istem stolpcu.) Kako bi jo olajšali? (Tako, da bi bilo v vseh stolpcih enako število ladij.)
4. Kateri od treh postopkov je najhitrejši? Zakaj?
5. Kakšne so prednosti in slabosti posameznih postopkov? (Drugi je hitrejši od prvega, vendar prvi ne zahteva, da so ladje urejene po velikosti, torej ga bomo uporabili v primerih, ko urejanje po velikosti iz kakega razloga ni praktično ali možno. Tretji je navadno hitrejši od prvih dveh, vendar se lahko, kadar imamo smolo, zgodi, da je zelo počasen. V najslabšem primeru so lahko po naključju vse ladje v istem stolpcu; v tem primeru je postopek enako počasen kot prvi.)
6. Razmisli, koliko ugibanj bi bilo pri binarnem iskanju potrebnih za sto ladij (šest), tisoč (devet), milijon (devetnajst). (Ko se število ladij podvoji, je potrebno le eno dodatno ugibanje. Pri razmišljanju o tem spomni otroke na prejšnjo aktivnost.)
7. Kaj se zgodi, če ladje z določeno številko ni? Koliko strelav bo potrebnih, da to odkrijemo? (Pri linearnem iskanju jih potrebujemo 25, pri binarnem iskanju največ 5, pri razpršitveni funkciji pa toliko, kolikor je ladij v ustreznem stolpcu, v katerem bi bila ta ladja, če bi obstajala.)

8. Otrokom povej, da imenujemo takšno tabelo "razpršena tabela", saj so ladje razpršene po njej glede na nekaj, kar izračunamo iz njihovih števil.
9. Če poznajo pojem funkcije, jim povej še, da funkciji, ki iz številke ladje izračunajo številko stolpca, pravimo razpršitvena funkcija. Računalniki ne "razpršujejo" ladij, temveč Google lahko (ali v resnici počne to, ne vemo) razpršuje besede, po katerih iščemo, tako da iz vsake besede izračuna neko številko. Ko računamo takšne številke si prizadevamo poiskati razpršitveno funkcijo, ki bo čim boljše "razprševala". Kot smo se naučili ob ladjah, si ne želimo razpršitvene funkcije, ki bi prevečkrat dodelila isto številko.

Za kaj gre?

Ko programi obdelujejo podatke, so ti lahko shranjeni na različne načine. Programer ima na voljo več osnovnih podatkovnih struktur (tabele, sezname, drevesa...), ki jih izbira glede na to, kaj bo program počel s podatki, katere operacije bodo najpogostejše, koliko pomnilnika bo vse skupaj zahtevalo...

V tej aktivnosti otroci spoznajo, kako se tri podatkovne strukture – neurejena, urejena in razpršena tabela – vedejo z vidika časa, potrebnega za iskanje določenega podatka.

Računalniki shranjujejo ogromne količine podatkov, zato je pomembno, da za iskanje po njih uporabljajo hitre algoritme. Eden takšnih je binarno iskanje, pri katerem so podatki urejeni, zato lahko z vsakim preverjanjem razpolovimo množico kandidatov. V trgovini bi za iskanje izdelka potrebovali le 14 primerjanj, kar je bistveno manj kot 5000. Če trgovina podvoji svojo ponudbo, to pomeni le eno primerjanje več.

Tretja strategija je razprševanje (*hashing*; razpršitvena tabela je v angleščini *hash table* in razpršitvena funkcija *hash function*). Podatki so razmetani v veliko število predalčkov; predalček, v katerega gre podatek, je izračunan iz podatka samega. Običajno imamo predalčkov dovolj preveč, da lahko razpršitvena funkcija – ki je navadno bolj zapletena od te, s katero so se igrali otroci – tako dobro razmeče podatke, da je v vsakem predalčku le eden ali največ nekaj. Na ta način računalnik v trenutku, z eno samo ali največ nekaj poizvedbami, najde, kar išče.

Med naštetimi načini iskanja se največ uporablja tretja, razen kadar je pomembno, da so podatki na kakšen način urejeni ali pa je nesprejemljivo, da je iskanje v nekaterih ponesrečenih primerih počasnejše.