Psihologia concursurilor de informatică

Cătălin Frâncu

Notă: Am publicat această carte în 1997 la Editura L&S Infomat. În 20 de ani s-au schimbat multe. Materia predată s-a schimbat mult. Concursurile s-au schimbat mult. Limbajele C și C++ au înlocuit aproape complet limbajul Pascal la concursuri. Mai ales, eu cel de acum nu mai sunt de acord cu unele principii, moduri de exprimare și stiluri de programare din această carte. Totuși, ocazional lumea îmi mai cere o copie a ei și mă simt jenat să le trimit un fișier Word (apropo de principii). De aceea, am publicat cartea online. Am păstrat tot conținutul original, reparând doar greșelile evidente de tipar și convertind formatul la LaTeXși imaginile la TikZ/PGF (cuvântul „migălos" abia începe să descrie aceste conversii). — Cătălin, București, 2 martie 2018.

Această carte îi este dedicată fratelui meu Cristi, căruia îi datorez o mare parte din cunoștințele mele în domeniul programării. Un gând bun pentru familia mea și pentru prietenii mei, care mi-au luat toate îndatoririle de pe umeri cât timp am scris cartea de față. Fără înțelegerea și răbdarea lor, nu mi-aș fi putut duce munca la bun sfârșit.

Cuvânt înainte

În ultimii ani s-au tipărit la noi în țară foarte multe culegeri de teorie și probleme de programare. Fiecare din ele acoperă diverse domenii ale informaticii. Unele își propun să inițieze cititorul în tainele diverselor limbaje de programare, altele pun accentul mai cu seamă pe tehnicile de programare și structurile de date folosite în rezolvarea problemelor. În general, cele din prima categorie conțin exemple cu caracter didactic și exerciții cu un grad nu foarte sporit de dificultate, iar celelalte demonstrează matematic fiecare algoritm prezentat, însă neglijează partea de implementare, considerând scrierea codului drept un ultim pas lipsit de orice dificultate.

Desigur, fiecare din aceste cărți își are rostul ei în formarea unui elev bine pregătit în domeniul informaticii. De altfel, citirea volumului de față presupune cunoașterea temeinică a conținutului ambelor tipuri de materiale enumerate mai sus. Totuși, pornind de la observația că scrierea unui program impune atât conceperea algoritmului și demonstrarea corectitudinii, cât și implementarea lui, ambele etape fiind complexe și nu lipsite de obstacole, am considerat necesară scrierea unui nou volum care să trateze simultan aceste două aspecte ale programării.

În afară de aceasta, după cum și titlul lucrării o spune, cartea se adresează pasionaților de informatică și celor care au de gând să participe la concursurile și olimpiadele de informatică. Concursul include apariția unui factor suplimentar care răstoarnă multe din obișnuințele programării „la domiciliu": timpul. Autorul a avut la dispoziție patru ani ca să descopere pe propria piele importanța acestui factor. Și, mai mult decât durata de timp în sine a concursului - care la urma urmei este aceeași pentru toți concurenții - contează capacitatea fiecăruia de a gestiona bine acest timp.

Dacă în fața calculatorului de acasă, cu o sticlă de Coca-Cola alături și casetofonul mergând, este într-adevăr un lucru lăudabil să justificăm matematic fiecare pas al algoritmului, să nu ne lăsăm înșelați de intuiție și să scriem programul fără să ne grăbim, alocându-ne o jumătate din timp numai pentru depanarea lui, în schimb în timp de concurs lucrurile stau tocmai pe dos. De demonstrații riguroase nu se mai ocupă nimeni,

intuiția este la mare preț și de nenumărate ori este criteriul care aduce victoria, iar timpul de-abia dacă este suficient pentru implementarea programului, despre depanare nemaiîncăpând discuții. În multe cazuri, cele două etape ale programării - conceperea și implementarea algoritmului - încep să se bată cap în cap. Uneori avem la dispoziție un algoritm foarte puternic, dar nu știm cum s-ar putea implementa, alteori acest algoritm nu face față volumului maxim de date de intrare, iar alteori ne dăm seama că am putea foarte ușor să scriem un program, dar nu suntem în stare să demonstrăm că el ar merge perfect. Foarte des se renunță la implementarea algoritmilor de complexitate optimă, care sunt alambicați și constituie adevărate focare de „bug"-uri, preferându-se un algoritm mai lent dar care să se poată implementa rapid și fără dureri de cap. Mulți olimpici pierd clasa a IX-a, poate chiar și pe a X-a descoperind aceste lucruri. Cartea de față își propune să le mai ușureze drumul.

S-a presupus cunoscut limbajul de programare Pascal, cu toate instrucțiunile și procedurile sale standard. În carte există multe surse în limbajul C standard. Am preferat acest lucru, deși la concursuri se recomandă programarea în Pascal, pentru că am sperat că un elev familiarizat cu limbajul Pascal va citi fără dificultate o sursă C și pentru că am dorit ca această carte să fie și un exercițiu de C, al cărui număr de utilizatori la nivelul liceului este destul de redus. Surse în Pascal există numai acolo unde se urmărește punerea în evidență a unei anumite subtilități a limbajului.

Tehnicile de programare s-au presupus cunoscute în esență, astfel încât am trecut direct la unele optimizări, la exemple de folosire a lor și la compararea lor, respectiv la prezentarea unor criterii în funcție de care să optăm pentru folosirea fiecăreia. De asemenea, am renunțat la definirea termenilor de graf, arbore, vector, matrice, listă, stivă și a tuturor celorlalte structuri de date de bază. Am considerat interesantă prezentarea pe larg numai a *heap*-urilor și a tabelelor de dispersie (*hash*), care sunt mai rar folosite și de aceea mai puțin cunoscute. În sfârșit, am presupus cunoscută noțiunea de complexitate a unui algoritm, deoarece în toate problemele se face calculul complexității.

Cartea prezintă interes și pentru problemele pe care le cuprinde. Ele nu sunt banale (de fapt, majoritatea au avut onoarea de a da bătăi de cap concurenților la olimpiade) și pot fi lucrate acasă de către elevi pentru menținerea în formă. Tocmai de aceea, am urmărit ca, ori de câte ori am propus o problemă spre rezolvare, enunțul să fie dat în aceeași formă pe care ar fi avut-o la un concurs: clar, detaliat, cu specificarea formatului datelor de intrare și ieșire și cu un exemplu sau două. Singura diferență este că, de regulă, la concursuri și olimpiade se precizează numai timpul limită admis pentru un test; în carte am considerat folositor să se specifice și o complexitate optimă a algoritmului care rezolvă problema, deoarece timpul de execuție variază în funcție de resursele calculatorului și de

limbajul de programare folosit, deci e un criteriu mai puțin semnificativ. Timpul destinat implementării unei probleme este în general egal cu cel care s-a acordat la concursul unde a fost propusă respectiva problemă.

În sfârșit, consider că programatorii care își propun să scrie aplicații de mari dimensiuni ar avea destul de multe lucruri de învățat din acest volum, deoarece am inclus și detalii privind structuri de date mai neobișnuite sau gestionarea economică a memoriei.

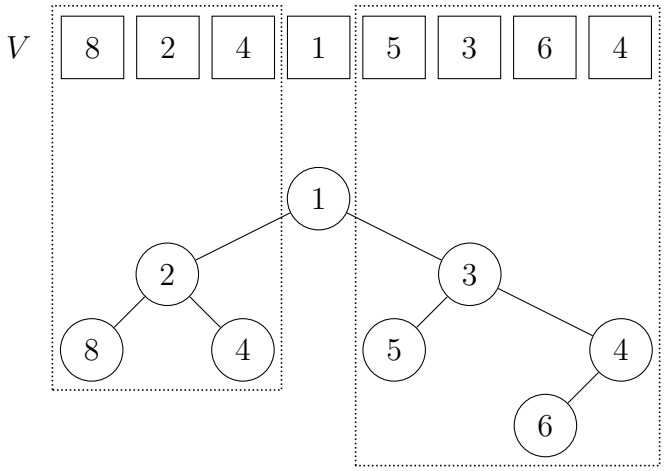
Sper să nu închideți această carte cu sentimentul că mai bine n-ați fi deschis-o.

Cătălin Frâncu

0.1 Problema 17

Iată un nou exemplu de problemă care admite două rezolvări: una evidentă, dar neeficientă și una mai puțin evidentă, dar cu mult mai eficientă.

ENUNȚ: Fie V un vector. Arborele cartezian atașat vectorului V este un arbore binar care se obține astfel: Dacă vectorul V este vid (are 0 elemente), atunci arborele cartezian atașat lui este de asemenea vid. Altfel, se selectează elementul de valoare minimă din vector și se pune în rădăcina arborelui, iar arborii cartezieni atașați fragmentelor de vector din stânga (respectiv din dreapta) elementului minim se pun în subarborele stâng, respectiv drept al rădăcinii. Iată, de exemplu, care este arborele cartezian al următorului vector cu 8 elemente:



În figură au fost încadrate prin dreptunghiuri punctate porțiunile din stânga, respectiv din dreapta elementului minim, împreună cu subarborii atașați. Trebuie observat că arborele cartezian atașat unui vector poate să nu fie unic, în cazul în care există mai

multe elemente de valoare minimă. Vom impune ca o condiție suplimentară ca elementul care va fi trecut în rădăcină să fie cel mai din stânga dintre minime (cel cu indicele cel mai mic). Astfel, arborele cartezian este unic.

Cerința problemei este ca, dându-se un vector, să i se construiască arborele cartezian.

Intrarea: Fișierul de intrare `INPUT.TXT` conține pe prima linie valoarea lui N ($N \leq 10.000$), iar pe a doua N numere naturale mai mici ca 30.000, separate prin spații.

Ieșirea se va face în fișierul text `OUTPUT.TXT` sub următoarea formă:

$T_1 \quad T_2 \quad T_3 \quad \ldots \quad T_N$

unde T_i este indicele în vector al elementului care este părintele lui $V[i]$ în arborele cartezian. Dacă $V[i]$ este rădăcina arborelui, atunci $T_i = 0$.

Exemplu: Pentru exemplul dat mai sus, fișierul `INPUT.TXT` este:

```
8
8 2 4 1 5 3 6 4
```

După cum reiese din figură, tatăl elementului 8 este elementul 2, adică al doilea în vector; tatăl elementului 2 este elementul 1, adică al 4-lea în vector; tatăl elementului 5 este elementul 3, adică al 6-lea în vector ș.a.m.d. Fișierul de ieșire este deci:

```
2 4 2 0 6 4 8 6
```

Complexitate cerută: $O(N)$.

Timp de implementare: 45 minute - 1h.

Timp de rulare: 2 secunde.

REZOLVARE: Nu întâmplător s-a impus o complexitate liniară pentru rezolvarea acestei probleme. Altfel, ea ar fi trivială în $O(N^2)$, prin următoarea metodă: scriem o procedură care parcurge vectorul și caută minimul, apoi se reapelează pentru bucățile de vector aflate în stânga, respectiv în dreapta minimului. Pentru a demonstra că această variantă de rezolvare are complexitate pătratică, să ne imaginăm cum s-ar comporta ea pe cazul:

$$V = (N \quad N-1 \quad N-2 \quad \ldots \quad 2 \quad 1) \tag{1}$$

La primul apel, procedura ar face N comparații pentru a parcurge vectorul (deoarece elementul minim este ultimul în vector) și s-ar reapela pentru porțiunea din vector care cuprinde primele $N-1$ elemente. La al doilea apel, ar face $N-1$ comparații și s-ar reapela pentru primele $N-2$ elemente etc. În concluzie, numărul total de comparații făcute este

$$N + (N-1) + (N-2) + \cdots + 1 = \frac{N(N+1)}{2} \tag{2}$$

de unde rezultă complexitatea. O problemă interesantă, pe care îi vom lăsa plăcerea cititorului să o rezolve, este de a demonstra că această versiune **nu poate atinge o complexitate mai bună decât** $O(N \log N)$ și de a arăta care sunt cazurile cele mai favorabile pe care se obține această complexitate.

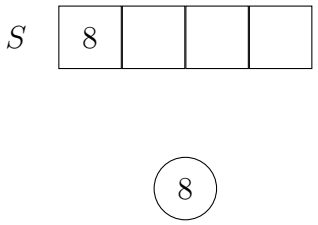
A doua metodă este și ea destul de ușor de înțeles și de implementat. Ceea este mai greu de acceptat este că ea are complexitate liniară, așa cum vom încerca să explicăm la sfârșit. Iată mai întâi principiul de rezolvare: vom porni cu un arbore cartezian vid și, la fiecare pas, vom adăuga câte un element al vectorului V la acest arbore, astfel încât structura obținută să rămână un arbore cartezian. La al k-lea pas, vom adăuga elementul $V[k]$ în arbore și vom restructura arborele în așa fel încât să obținem arborele cartezian atașat primelor k elemente din V. Trebuie să ne concentrăm atenția asupra a două lucruri:

1. Cunoscând arborele cartezian atașat primelor k-1 vârfuri și elementul $V[k]$, cum se obține arborele cartezian atașat primelor k vârfuri?
2. Cum reușim să actualizăm de N ori arborele astfel încât timpul total consumat să fie liniar?

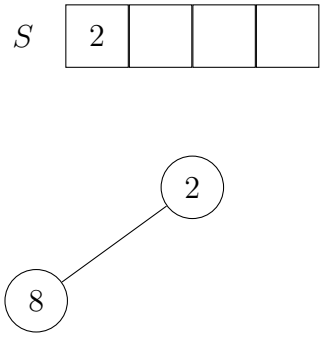
Pentru a răspunde la prima întrebare, pe lângă vectorii V și T, mai este necesară o stivă S, în care vom stoca elemente ale vectorului V. Inițial, stiva este vidă. Atunci când un nou element X sosește, el va fi introdus în stivă imediat după ultimul număr din stivă care are o valoare mai mică sau egală cu X. Toate elementele care se aflau înainte în stivă pe poziții mai mari sau egale cu poziția pe care a fost inserat X vor fi eliminate din stivă, iar elementul care se afla exact pe poziția lui X va deveni fiul stâng al lui X. X însuși va deveni fiul drept al predecesorului său în stivă. La fiecare moment, primul element din stivă este rădăcina arborelui cartezian.

Pentru a înțelege mai bine principiul de funcționare a stivei, să analizăm mai de aproape exemplul din enunț.

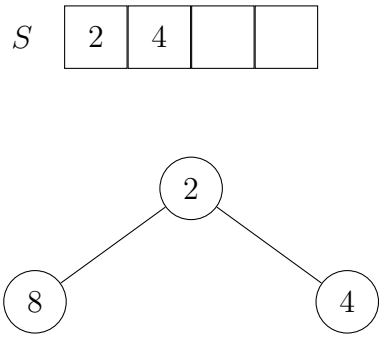
La început stiva este vidă. Primul element din V are valoarea 8, drept care îl vom pune în stivă, iar arborele cartezian va avea un singur nod:



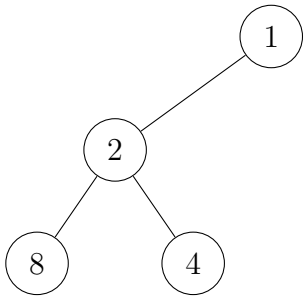
Următorul element sosit este 2. Acesta este mai mic decât 8, deci trebuie introdus înaintea lui în stivă. El va fi deci primul element din stivă și rădăcina arborelui cartezian la acest moment. Concomitent, 8 va fi eliminat din stivă și va deveni fiul stâng al lui 2:



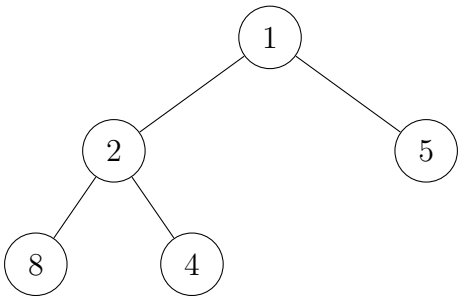
Se observă că arborele obținut este tocmai arborele cartezian atașat secvenței $(V[1], V[2])$. Următorul element este 4, care este mai mare decât 2, deci trebuie adăugat în vârful stivei. Nici un element nu este eliminat din stivă, iar 4 devine fiul drept al lui 2:



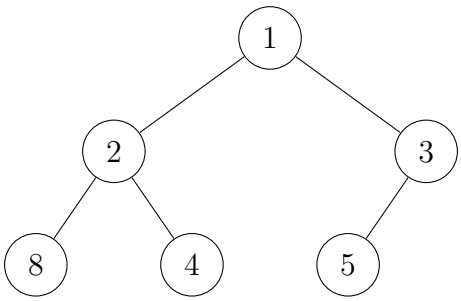
Următorul element sosit este 1, care este mai mic decât toate numerele din stivă. Stiva se va goli, iar numărul 2 (cel peste care se va scrie 1) va deveni fiul stâng al lui 1:



Deja arborele începe să semene cu forma sa finală. Urmează elementul 5, care va fi adăugat în stivă și „atârnat" în dreapta lui 1:

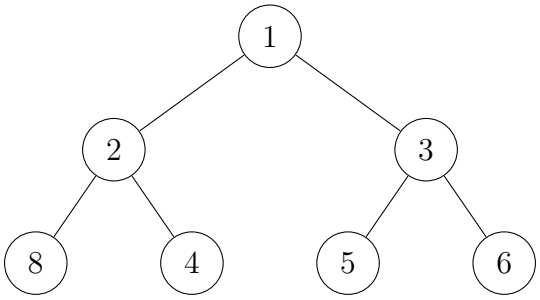


Elementul 3 este mai mare ca 1, căruia îi va deveni fiu drept, dar mai mic ca 5, pe care îl va elimina din stivă:

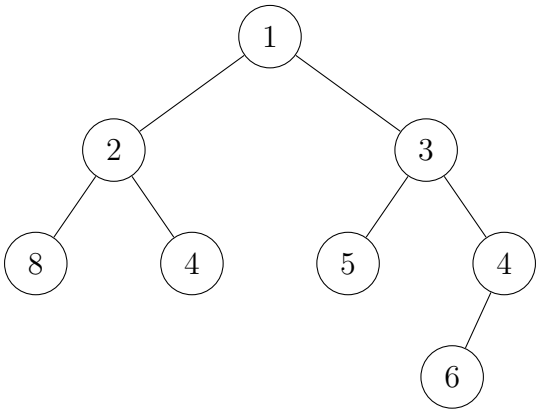


Următorul număr, 6, va fi adăugat la extremitatea dreaptă a arborelui și în vârful

stivei:



În sfârșit, elementul 4 va fi fiul drept al lui 3 și îl va elimina din stivă pe 6, care îi va deveni fiu stâng:



Se observă că arborele a ajuns tocmai la forma sa corectă. Trebuie acum să ne ocupăm de un detaliu de implementare. Pentru a afla poziția pe care trebuie inserat un element în stivă avem două metode:

1. Putem să căutăm în stivă de la dreapta la stânga (ar fi mai corect spus „de la vârf spre bază") până dăm de un element mai mic decât cel de inserat; programul folosește această metodă și îi vom discuta în final eficiența.

2. Putem face o căutare binară în stivă, întrucât elementele din stivă au valori crescătoare de la bază spre vârf (lăsăm demonstrația acestei afirmații în seama

cititorului). O căutare binară într-un vector de k elemente poate necesita, în cazul cel mai nefavorabil, $\log k$ comparații. În cazul cel mai nefavorabil, când vectorul V este sortat crescător, elementele vor fi introduse pe rând în stivă și nu vor mai fi scoase, deci la fiecare pas se vor face $\log k$ comparații, unde k ia valori de la 1 la N. Complexitatea care rezultă este mai slabă decât cea cerută:

$$\begin{aligned} O(\log 1 + \log 2 + \cdots + \log N) = O(\sum_{k=1}^{N} \log k) = \\ = O(\log \prod_{k=1}^{N} k) = O(\log N!) = O(N \cdot \log N) \end{aligned} \tag{3}$$

Acesta este unul din puținele cazuri în care căutarea binară este mai ineficientă decât cea secvențială.

Pentru ușurința programării, sursa C de mai jos reține în stiva S nu valorile elementelor, ci indicii lor în vectorul V (deoarece aceștia sunt ceruți pentru construcția vectorului T).

```
#include <stdio.h>
#define NMax 10001

int V[NMax], /* Vectorul */
    T[NMax], /* Vectorul de tati */
    S[NMax], /* Stiva */
    N;       /* Numarul de elemente */

void ReadData(void)
{ FILE *F=fopen("input.txt","rt");
  int i;

  fscanf(F,"%d\n",&N);
  for (i=1; i<=N;)
    fscanf(F, "%d", &V[i++]);
}

void BuildTree(void)
{ int i,k,LenS=0;

  S[0]=0; /* Pentru ca initial T[1] sa fie 0 */
  for (i=1; i<=N; i++)
    { /* Cauta pozitia pe care va fi inserat V[i] */
      k=LenS+1;
```

```
25        while (V[S[k-1]]>V[i]) k--;
26        /* Face corecturile in S si T */
27        T[i]=S[k-1];
28        if (k<=LenS) T[S[k]]=i;
29        /* i este ultimul element din stiva, deci... */
30        S[LenS=k]=i;
31      }
32  }
33
34  void WriteSolution(void)
35  { FILE *F=fopen("output.txt","wt");
36    int i;
37    for (i=1; i<=N;)
38      fprintf(F,"%d ",T[i++]);
39    fprintf(F,"\n");
40  }
41
42  void main(void)
43  {
44    ReadData();
45    BuildTree();
46    WriteSolution();
47  }
```

Acum să analizăm și complexitatea acestui algoritm. În primul rând, ea nu poate fi mai bună decât $O(N)$, pentru că aceasta este complexitatea funcțiilor de intrare și ieșire. Procedura `BuildTree` se compune dintr-un ciclu `for` în care se execută patru operații în timp constant și o instrucțiune repetitivă `while`. Numărul total de operații în timp constant care se execută în procedură este prin urmare $O(N)$. Problema este: care este numărul total maxim de evaluări ale condiției logice din bucla `while`? Aparent, bucla while se execută de $O(N)$ ori, deci numărul total de evaluări ar fi $O(N^2)$. Să aruncăm totuși o privire mai atentă.

Fiecare evaluare a condiției din bucla `while` are ca efect decrementarea lui k și, implicit, eliminarea unui element deja existent în stivă. Pe de altă parte, fiecare element este introdus în stivă o singură dată și deci nu poate fi eliminat din stivă decât cel mult o dată. Așadar numărul maxim de elemente ce pot fi eliminate din stivă pe parcursul executării procedurii `BuildTree` este $N-1$, deci numărul total de evaluări ale condiției este $O(N)$. De aici rezultă că programul are complexitate liniară.

0.2 Problema 18

ENUNȚ: Se dă un vector nesortat cu elemente numere reale oarecare. Considerând că vectorul ar fi sortat, se cere să se găsească distanța maximă între două elemente consecutive ale sale, fără însă a sorta efectiv vectorul.

Intrarea: Fișierul `INPUT.TXT` conține pe prima linie numărul N de elemente din vector ($N \leq 5.000$). Pe următoare linie se dau numerele separate prin spații.

Ieșirea: Pe ecran se va tipări un mesaj de forma:

Distanța maximă este D

Exemplu: Pentru fișierul de intrare cu conținutul

```
4
5 3.2 2 3.7
```

răspunsul trebuie să fie

```
Distanta maxima este 1.3
```

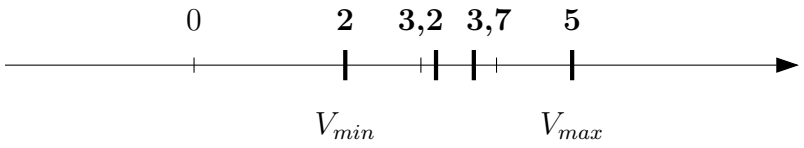
Timp de implementare: 30 minute.

Timp de rulare: 2-3 secunde.

Complexitate cerută: $O(N)$.

REZOLVARE: Desigur că primul lucru la care ne gândim este să sortăm vectorul și să îl parcurgem apoi de la stânga la dreapta, căutând distanța maximă între două elemente consecutive. Complexitatea unui asemenea algoritm este $O(N \log N)$. Nici această soluție nu este rea, iar la un concurs, comisiei de corectare i-ar veni destul de greu să găsească teste care să departajeze un algoritm în $O(N \log N)$ de unul liniar, chiar și pentru $N = 5.000$. Totuși, vom arăta care este algoritmul liniar; în primul rând de dragul „artei", iar în al doilea rând pentru că nu este cu mult mai greu de implementat decât o sortare.

Primul lucru care trebuie făcut este găsirea maximului și a minimului din vector; să notăm aceste valori cu V_{max} și V_{min}. Aceste operații se fac în timp liniar, eventual chiar la citirea datelor din fișier. Apoi se împarte intervalul $[V_{min}, V_{max}]$ de pe axa reală în $N - 1$ intervale egale. Iată cazul exemplului din enunț, unde $V_{min} = 2$ și $V_{max} = 5$:



Lungimea fiecărui interval va fi deci de

$$D = \frac{V_{max} - V_{min}}{N - 1} \qquad \text{(în cazul nostru } D = 1\text{)} \tag{4}$$

De ce s-a făcut această împărțire? Dacă notăm cu D_{max} distanța maximă pe axă între două numere vecine, adică tocmai valoarea pe care o căutăm, se poate demonstra că $D_{max} \geq D$. Într-adevăr, între cele N numere de pe axă se formează $N - 1$ intervale. Dacă presupunem că $D_{max} < D$, rezultă că distanța între oricare două numere consecutive de pe axă este mai mică decât D. De aici deducem că distanța dintre primul și ultimul număr, adică $V_{max} - V_{min}$, este mai mică decât $(N - 1) \times D$. Dar aceasta duce la relația:

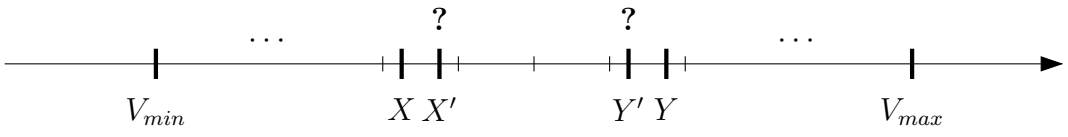
$$V_{max} - V_{min} < (N - 1) \cdot \frac{V_{max} - V_{min}}{N - 1} \implies V_{max} - V_{min} < V_{max} - V_{min} \tag{5}$$

relație care este absurdă; demonstrația afirmației $D_{max} \geq D$ este completă.

Următorul pas pe care îl avem de făcut este să parcurgem încă o dată vectorul de numere și să aflăm pentru fiecare element căruia dintre intervalele de lungime îi aparține. Și această operație se poate face în $O(N)$. Convenim ca dacă un număr X se află exact la limita dintre două intervale, adică

$$X = V_{min} + k \cdot D, \quad 0 \leq k < N \tag{6}$$

el să fie considerat ca aparținând intervalului din dreapta. Aceasta înseamnă că elementul de valoare V_{max} nu aparține nici unuia din cele $N - 1$ intervale, ci celui de-al N-lea interval, $[V_{max}, V_{max} + D]$. Să vedem la ce ne ajută acest lucru. Din moment ce $D_{max} \geq D$, rezultă că este imposibil ca distanța maximă să se producă între două numere din același interval, deoarece distanța în cadrul aceluiași interval nu poate atinge valoarea D. Este deci obligatoriu ca distanța maximă să apară între două elemente din intervale distincte. Să urmărim în figura următoare ce alte proprietăți mai au aceste numere:



Dacă X și Y sunt valorile între care diferența este maximă, este de la sine înțeles că între ele nu mai există nici un număr, deoarece se prespune că X și Y sunt consecutive în vectorul sortat. Aceasta înseamnă însă că X este cel mai mare număr din intervalul său, iar numărul X' nu poate exista acolo unde a fost el figurat. Analog, Y este cel mai mic număr din intervalul său, iar numărul Y' nu poate exista. De fapt, în nici unul din intervalele dintre cele care le cuprind pe X și Y nu poate exista nici un număr.

Prin urmare, diferența maximă se poate produce numai între maximul unui interval și minimul imediat următor. Următorul pas în găsirea soluției presupune aflarea pentru fiecare din cele $N-1$ intervale (sau N dacă îl considerăm și pe ultimul, cel care nu îl conține decât pe V_{max}) a minimului și a maximului. Și acest pas se execută în $O(N)$, deoarece procesarea fiecărui element din vector se reduce la numai două comparări, cu minimul și cu maximul intervalului în care se încadrează el. Vor rezulta doi vectori care în program se vor numi Lo și Hi. Iată care sunt valorile lor pentru exemplul nostru:

Intervalul:	[2, 3)	[3, 4)	[4, 5)	[5, 6)
Lo:	2	3,2	–	5
Hi:	2	3,7	–	5

Deoarece în intervalul [4,5) nu se află elemente, rezultă că elementele corespunzătoare din vectorii Lo și Hi trebuie să aibă o valoare specială care să informeze programul asupra acestui lucru. De exemplu, sursa oferită mai jos folosește următorul artificiu: inițializează vectorul Lo cu valori foarte mari ($V_{max}+1$), astfel încât orice număr „repartizat" într-un interval să modifice această valoare. Similar, vectorul Hi este inițializat cu $V_{min}-1$. Dacă pentru un interval aceste valori se păstrează până la sfârșit, putem trage concluzia că în respectivul interval nu se află nici un număr.

În continuare, elementele vectorilor Lo și Hi se amestecă formând un nou vector W care de data aceasta este sortat. Sortarea este foarte ușoară, pentru că nu avem decât să așezăm numerele în ordinea $Lo[1]$, $Hi[1]$, $Lo[2]$, $Hi[2]$, ..., $Lo[N]$, $Hi[N]$. Deși la prima vedere pare că noul vector rezultat are $2N$ elemente, de fapt el are numai N elemente, pentru că:

- Dacă într-un interval K există un singur număr, (cazul intervalelor [2,3) și [5,6)) sau există numai numere egale, atunci $Lo[K] = Hi[K]$ și este suficient să copiem în W una singură dintre aceste două valori;
- Dacă într-un interval nu există nici un număr, putem să nu copiem nici o valoare în vectorul W.

Astfel, construcția vectorului W se poate face în timp liniar, mai exact în $O(2N)$. Se observă că la această construcție se poate întâmpla ca unele numere să „dispară", adică să nu fie trecute în vectorul W. De exemplu, dacă între numerele 3,2 și 3,7 ar mai fi existat un număr, 3,5, el nu ar fi fost nici minim, nici maxim pentru intervalul său, deci nu ar fi fost copiat. Totuși, trierea în acest fel a elementelor nu afectează în nici un fel soluția. În cazul nostru, nu se întâmplă să dispară nici un element, deci $W = (2, 3, 2, 3, 7, 5)$.

După ce am construit vectorul W, nu mai avem decât să-l parcurgem de la stânga la dreapta și să tipărim diferența maximă întâlnită între două numere consecutive (repetăm, vectorul W este sortat), această ultimă etapă necesitând și ea un timp liniar. Nu se poate obține o complexitate inferioară celei liniare, întrucât citirea datelor presupune ea însăși N operații.

Ca un detaliu de implementare, odată ce au fost construiți vectorii Lo și Hi, vectorul V nu mai este necesar, deci putem construi chiar în el vectorul W, pentru a economisi memorie.

```
#include <stdio.h>
#define NMax 5000
float V[NMax+1], Lo[NMax+1], Hi[NMax+1];
float Delta, Max, Min;
int N;

void ReadData(void)
/* Citeste vectorul si afla maximul si minimul */
{ FILE *F=fopen("input.txt", "rt");
  int i;

  fscanf(F, "%d\n", &N);
  fscanf(F, "%f", &V[1]);
  Max=Min=V[1];

  for (i=2; i<=N; i++)
    {
      fscanf(F, "%f", &V[i]);
      if (V[i]>Max) Max=V[i];
        else if (V[i]<Min) Min=V[i];
    }
  fclose(F);
}

void Split(void)
{ int i, K;
```

```
27
28    Delta = (Max-Min)/(N-1);
29    /* Se initializeaza vectorii Lo si Hi */
30    for (i=0; i<=N; Lo[i]=Max+1, Hi[i++]=Min-1);
31
32    /* Se construiesc intervalele */
33    for (i=1; i<=N; i++)
34      {
35        K = (V[i]-Min)/Delta;
36        if (V[i]<Lo[K]) Lo[K]=V[i];
37        if (V[i]>Hi[K]) Hi[K]=V[i];
38      }
39  }
40
41  void Rebuild(void)
42  /* Rescrie vectorul V, pentru a economisi memorie,
43     pastrand numai capetele intervalelor */
44  { int i, M=0;
45
46    for (i=0;i<N; i++)
47      {
48        if (Lo[i] != Max+1) V[++M]=Lo[i];
49        if (Hi[i] != Min-1 && Hi[i] != Lo[i]) V[++M]=Hi[i];
50      }
51    N=M;
52  }
53
54  void FindGap(void)
55  /* Acum cautarea distantei maxime se face
56     secvential, vectorul fiind sortat */
57  { int i;
58    float Gap=0;
59
60    for (i=2; i<=N; i++)
61      if (V[i]-V[i-1] > Gap)
62        Gap = V[i]-V[i-1];
63    printf("Distanta␣maxima␣este␣%0.3f\n", Gap);
64  }
65
66  void main(void)
67  {
68    ReadData();
69    if (Max==Min)
70      puts("0");
71    else
```

```
72      {
73        Split();
74        Rebuild();
75        FindGap();
76      }
77  }
```

Cuprins