



dopermutovanost znamená, že zbytková posloupnost je sestupná a nejdlejší možná. Začít permutovat zbytek od začátku je také jednoduché. Stačí si uvědomit, že jediné co je třeba udělat, je setřídít ji vzestupně. To bylo ale kamenem úrazu některých řešení, neboť posloupnost třídili některým z třídících algoritmů místo aby si všimli, že ze sestupné posloupnosti vytvoříme vzestupnou tak, že obrátíme pořadí jejich prvků.

Záměna přilehlého prvku je již triviální. Problém může nastat pouze tehdy, když žádný přilehlý prvek neexistuje. To se ale stane jenom tehdy, když jsme na vstupu dostali lexicograficky nejvyšší možnou permutaci.

Paměťová i časová složitost je  $O(N)$ .

Jan Bulánek & Zbyněk Falt

### 18-5-3 Číňanské volby

Představme si následující „paralelní“ algoritmus na vyhodnocení voleb: každý Číňan si najde do dvojice nějakého Číňana, který chce volit někoho jiného. Pokud má některý kandidát nadpoloviční většinu hlasů, někteří z jeho příznivců zůstanou nespárovaní; u takto nalezeného potenciálního vítěze si ještě ověříme, zda skutečně vyhrál (spočítáme si počet jeho hlasů).

Druhou část tohoto algoritmu jistě zvládneme v lineárním čase a s konstantní pamětí, zbývá si rozmyslet, jak realizovat tu první. Zjevně nás nezajímá, jak jsou Číňané spárování, stačí nám vědět, kolik jich zůstane nespárovaných a pro koho nespárování hlasují. Budeme si tedy udržovat dvě čísla –  $K$  (číslo kandidáta, pro nějž hlasují nespárování Číňané), a  $C$  (počet nespárování Číňanů). Postupně procházíme všechny Číňany. Pokud aktuální Číňan hlasuje pro kandidáta  $K$ , nejde ho spárovat, a proto zvýšíme  $C$  o jedna. Pokud hlasuje pro někoho jiného (kandidáta  $X$ ), rozlišíme dva případy: jestliže je  $C > 0$ , pak tohoto voliče spárujeme s jedním z voličů kandidáta  $K$ , tedy snížíme  $C$  o jedna. Jestliže je  $C = 0$ , nelze aktuálního Číňana spárovat, a proto dosadíme  $K = X$  a  $C = 1$ .

Tento algoritmus používá konstantní množství paměti a seznam hlasů projde dvakrát, časová složitost je tedy  $O(N)$ . Pokud bychom chtěli být přesnější, je třeba vzít do úvahy to, že na uložení čísel potřebujeme  $O(\log N)$  bitů, což pro velká  $N$  nelze zanedbat tedy paměťová složitost je  $O(\log N)$  a časová  $O(N \log N)$ .

Zdeněk Dvořák

### 18-5-4 Detektýv

S důkladností takřka šerlokovskou prozkoumáme několik možných řešení, až usvědčíme to nejrychlejší. Označme si (věrní písmenkům ze zadání)  $N$  délku stopovaného řetězce,  $k$  počet podezřelých sekvencí,  $p_1, \dots, p_k$  délky těchto sekvencí a  $P = p_1 + \dots + p_k$  jejich celkovou délku.

0. *pokus* (jak by ho vymyslel strážník Vopička): Budeme hledat každou sekvenci zvlášť, a to tak, že si po vstupu „pojedeme okénkem“ délky  $p_i$  a vždy porovnáme, jestli se okénko rovná  $i$ -té sekvenci. Kdybychom si okénko ukládali jako cyklické pole, zvládli bychom ho posunout v konstantním čase, ale stejně nás nemine čas  $O(p_i)$  na porovnání. Celkově trvá  $O(Np_1 + \dots + Np_k) = O(NP)$  a navíc potřebujeme  $k$ -krát volat rewind.

1. *pokus* (inspektor Neverley): Damned, na hledání výskytů jednoho řetězce přeci můžeme použít algoritmus KMP

z té vaší cookbook, takže jeden průchod zvládneme v čase  $O(N + p_i)$ , celkově tedy  $O(Nk + P)$  s  $k$  rewindy. That's it.

2. *pokus* (policejní rada Žák): V kuchařce je přeci i algoritmus A-McC na hledání výskytů více slov najednou. Stačí, když hlášení výskytu nahradíme připočtením jedničky k počítadlu. (Na to praktikant Hlaváček:) Dobrý plán, pane rado, ale má jedno háčisko jak na sumce: jelikož se sekvence mohou překrývat, může jich v jednom místě končit až  $k$ , takže jsme opět na  $O(Nk + P)$ , i když tentokrát bez rewindů.

3. *pokus* (Šerlok osobně): Postavíme si vyhledávací automat jako v minulém pokusu, ale místo abychom počítali rovnou výskyt, budeme si pamatovat jen to, kolikrát jsme prošli kterým stavem, a pak z toho výskyt dopočítáme. Well, ale jak?

Pokud máme nějaký stav  $\alpha$  (o kterém víme, že je prefixem některého z vyhledávaných slov, takže mimo jiné mezi stavy najdeme všechny sekvence stop, které počítáme) a chceme zjistit, kolikrát se slovo  $\alpha$  v textu vyskytlo, stačí sečíst počet průchodů tímto stavem a všemi dalšími stavy, které končí na  $\alpha$ , což jsou přesně ty, ze kterých se do  $\alpha$  lze dostat pomocí zpětné funkce (případně zavolané vícekrát).

Stačí tedy projít automat v opačném pořadí, než ve kterém jsme vytvářeli zpětnou funkci (nejlepší bude si během konstrukce automatu toto pořadí zapamatovat, třeba v poli, v němž jsme měli uloženu frontu). Pro každý stav  $\alpha$  pak přičteme počítadlo odpovídající tomuto stavu k počítadlu stavu, do něž vede z  $\alpha$  zpětná funkce. (To se pak přičte podle další zpětné funkce atd., takže počítadlo stavu  $\alpha$  se opravdu postupně popřičítá ke všem rozšířením stavu  $\alpha$ .)

To vše zvládneme v čase  $O(P + N + P)$  (konstrukce automatu + průchod textem + dopočítání), čili  $O(P + N)$ , a v paměti  $O(P + N)$ , bez jediného zavolání rewindu.

Program obnáší připsání cca čtyř řádků ke zdrojáku z kuchařky. Abychom z toho nevybrusili tak snadno, ukážeme si trochu jinou implementaci v Cěčku.

It's a lemon tree, my dear Watson!

Martin Mareš

P.S.: V kuchařce si, nečekán, nezávan, opět zařadil šotek a trochu pomíchal vypisování nalezených slov. Opravenou verzi naleznete na webu, rdíciho se kuchaře opodál.

### 18-5-5 Do vysokých kruhů

Nejprve bylo potřeba oblastí převést na objekty, se kterými umíme manipulovat rozumněji než s obecnými množinami bodů v rovině. Velmi užitečné je představit si protínající se kružnice jako graf s průsečíky a dotyky kružnic jako vrcholy. Hrany budou oblouky mezi sousedními vrcholy. Tento graf je vlastně multigraf, což je graf, ve kterém může mezi dvěma vrcholy vést více než jedna hrana a z jednoho vrcholu do toho samého může vést více než jedna smyčka. Takový graf je určité jednoznačně zadán polohami a poloměry kružnic a je rovinný (původní rozmístění kružnic je jeho rovinné nakreslení). Bohužel se nám do něj nijak nepromítnou izolované kružnice, ty je třeba ošetřit jinak.

Nyní se nám z na první pohled neuchopitelného problému stal problém mnohem jednodušší – spočítat stěny rovinného grafu. K tomu se ideálně hodí Eulerova věta:

$$V + F = K + E + 1$$

Toto je vztah mezi počtem vrcholů ( $V$ ), stěn (včetně té vnější) ( $F$ ), komponent souvislosti ( $K$ ) a hran ( $E$ ). Tato

### Úloha 18-5-5 – Do vysokých kruhů – program

```
#include <stdio.h>
#include <math.h>

#define SQR(x) ((x)*(x)) // Druhá mocnina
#define EPSILON (4.2e-10) // Podovnění s tolerancí
#define EQ(x,y) ((x)+EPSILON>(y))&&((x)-EPSILON<(y))

#define MAX_N 10000
int N; // Počet kružnic

double x[N], y[N], r[N]; // Kružnice
int byl[N]; // Navštíveno?
int v=0; // Vrcholy (průsečíky)
int k=0; // Komponenty

int pruseciku(int a, int b) // Kolik je mezi a a b průsečíků?
{
    double d=sqrt(SQR(x[a]-x[b])+SQR(y[a]-y[b])); // Vzdálenost středů
    if (d>r[a]+r[b]) return 0; // Úplně mimo
    if ((d+r[a]<r[b])|| (d+r[b]<r[a])) return 0; // Jedna ve druhé
    if (EQ(d,r[a]+r[b])) return 1; // Vnější dotyk
    if ((EQ(d+r[a],r[b]))|| (EQ(d+r[b],r[a]))) return 1; // Vnitřní dotyk
    return 2; // Jinak mají právě dva průsečíky
}

void navstiv(int koho) // Rekurse pro průchod komponent
{
    int i,p;

    byl[koho]=1;

    for (i=0;i<N;i++) { // Teď projdi všechny sousedy
        p=pruseciku(i,koho);
        if ((i!=koho) && (p>0)) {
            if (i<koho) v+=p; // Připočti průsečíky (ale ne dvakrát)
            if (!byl[i]) navstiv(i); // Navštiv
        }
    }
}

int main()
{
    int i;
    scanf("%d",&N); // Načteme
    for (i=0;i<N;i++) {
        scanf("%lf %lf %lf",&x[i],&y[i],&r[i]);
        byl[i]=0;
    }
    for (i=0;i<N;i++)
        if (!byl[i]) { // V této komponentě jsme ještě nebyli
            k++;
            navstiv(i);
        }

    printf("Oblastí: %d\n",k+v+1);
    return 0;
}
```

```

void count(void) // propagování počtů po zpětných hranách
{
    for (struct state *s=tail; s != &root; s=s->qprev) {
        s->back->count += s->count;
        if (s->out) // ... a vypisování četnosti slov
            printf("%6d %s\n", s->count, s->out);
    }
}

int main(int argc, char **argv) // main sweet main :)
{
    for (int i=1; i<argc; i++)
        insert(argv[i]);
    build();
    run();
    count();
    return 0;
}

```

věta platí pro rovinné grafy a platí i pro multigrafy, pokud si zvolím, že mezi „rovnoběžnými“ násobnými hranami jsou také stěny a že smyčka přidává jednu stěnu. Toto rozšíření přesně odpovídá naší představě toho, jak kružnice dělí rovinu na oblasti.

Věta se dokazuje indukcí podle složitosti grafu. Pro prázdný graf určitě platí  $(0 + 1 = 0 + 0 + 1)$ . Přidáme-li nový vrchol, stoupnou  $V$  i  $K$  o jedna a rovnost zůstane zachována. Přidáme-li hranu a zvýšíme tak  $E$  o jedna, pak jsme buď spojili dvě komponenty souvislosti a snížili  $K$  o jedna, nebo přidali jednu stěnu rozdělením nějaké existující na právě dvě. V obou případech zůstane rovnost zachována a druhý případ navíc zahrnuje přidávání násobných hran a smyček. Každý rovinný (multi)graf lze postavit z prázdného přidáváním vrcholů a hran, takže pro něj věta musí platit.

Stačilo by tedy spočítat počet komponent, hran a průsečíků. Víme, že na každé kružnici je stejně vrcholů a hran. Vrchol je ale sdílen mezi dvěma kružnicemi, zatímco hrana patří právě jedné. Jinak řečeno je stupeň každého vrcholu 4. Z toho plyne, že  $E = 2V$ . Tedy:

$$F = K + 2V + 1 - V = K + V + 1.$$

Tento vzorec nám navíc zahrne i izolované kružnice, počítáme-li je jako jednu komponentu bez průsečíků. To nám trochu zjednoduší algoritmus.

Stačí tedy spočítat počet komponent a průsečíků, obojí zvládneme v čase  $\mathcal{O}(N^2)$  průchodem do hloubky (s hledáním sousedů vyzkoušením všech) a vyzkoušením všech dvojic. Zkoušení dvojic navíc zahrneme do toho průchodu. V programu je průchod do hloubky realizován rekurzivní funkcí `navstiv()`. Ta bude pro každý vrchol spuštěna určitě právě jednou, určitě projde celou komponentu a zároveň správně napočítá počet průsečíků. Jen je si třeba dát pozor, abychom nezapočítávali průsečíky dvakrát (za páry kružnic  $(k_i, k_j)$  a  $(k_j, k_i)$ ).

Toto řešení má časovou složitost  $\mathcal{O}(N^2)$ , paměťovou  $\mathcal{O}(N)$ . Existuje ještě jiné o dost složitější řešení používající *zámětnací čáru* k dosažení složitosti  $\mathcal{O}((N+V) \log N)$ , což je lepší než naše  $\mathcal{O}(N^2)$ , pokud je počet průsečíků  $V < N^2 / \log N$ , tedy pro dost „řídké“ konfigurace kružnic. Pro  $V = \mathcal{O}(N^2)$  má ale časovou složitost až  $\mathcal{O}(N^2 \log N)$ . Paměťová složitost tohoto algoritmu je  $\mathcal{O}(N)$ . Jeho popis by ale byl dost komplikovaný a proto ho neuvádím.

Tomáš Gavenčíak

## 18-5-6 Propoflování

*Úloha 1:* Buď  $G$  graf, který vznikne z CFG tak, že neme na orientaci hran, přidáme hranu mezi vstupní a výstupní blokem CFG, a zahodíme hrany, na kterých není žádný čítač. Graf  $G$  nemůže obsahovat cyklus – počty provedení hran na odpovídajícím cyklu (nebo cestě) v CFG nemohou být libovolně zvyšovat a snižovat bez toho, že bychom libovolný čítač, tedy by nebylo možné pouze z čítacího profilu  $G$  je tedy les, a má nejvýše  $N - 1$  hran, počet bloků v programu. Proto v původním CFG vynechat čítač na nejvýše  $N - 2$  hranách.

Naopak, pokud vynecháme čítač na libovolných  $N - 2$  hranách takových, aby  $G$  byl strom, dokážeme počty provedení hran na zbývajících hran dopočítat: protože  $G$  je strom, má každý vrchol stupně jedna (vede do něj jen jedna hrana  $e$ ). To znamená, že pro každý blok  $b$ , u něž neznáme počty provedení hran, které do něj vstupují nebo z něj vystupují. Součet počtů provedení hran vstupujících do bloku je roven součtu počtů provedení hran, které z něj vystupují, tedy dokážeme počítat počet provedení hrany  $e$ . Hranu  $e$  vyhodíme z rovnice a tento postup opakujeme, dokud neurčíme počty provedení všech hran. Časová i paměťová složitost tohoto algoritmu je lineární.

*Úloha 2:* Zjevně stačí spočítat počty provedení bloků. Pokud hrana vychází z bloku, který se provede  $n$ -krát, provedeme ji s pravděpodobností  $p$ , pak výpočet touto cestou projde  $pn$ -krát. Mějme blok  $b$ , do něž vstupují hrany  $e_1, e_2, \dots, e_k$ , které vycházejí z bloků  $b_1, b_2, \dots, b_k$  s pravděpodobnostmi  $p_1, p_2, \dots, p_k$ . Počet provedení bloku  $b$  je součtem počtů provedení hran, které do něj vstupují. Jestliže je blok  $b$  proveden  $n$ -krát a bloky  $b_i$  jsou provedeny  $n_i$ -krát, pak platí

$$n = \sum_{i=1}^k p_i n_i.$$

Pokud navíc položíme počet provedení vstupních bloků rovný 1, dostáváme soustavu lineárních rovnic, jejímž řešením je hledaný profil (samozřejmě, kdybychom chtěli profil, který není la přesní, je třeba ještě dokázat, že tato soustava má jednoznačné řešení). Časová složitost závisí na tom, jak rychle dostaneme soustavu budeme řešit. Pokud použijeme Gaussovu metodu, dostáváme kubickou časovou složitost, což je příliš mnoho. Proto se většinou používají algoritmy, které využívají toho, že CFG má speciální strukturu, která se nepoužije příkaz skoku `goto`, každý cyklus má jen jeden vstup. V případě, že `goto` použijeme a tuto podmínku neporušíme, tyto algoritmy vrátí pouze přibližné řešení, které však fungují v lineárním čase.

Zdeněk

```

program Permutace;

procedure swap(var a, b : char);
var
  p : char;
begin
  p := a; a := b; b := p;
end;

var
  perm : string;
  N, i, j : integer;
begin
  readln(perm);
  N := length(perm);

  i := N;
  while (i > 1) and (perm[i-1] > perm[i]) do
    dec(i); { Hledáme zbytek }
  if i > 1 then begin
    for j := 0 to (N-i-1) div 2 do
      swap(perm[j+i], perm[N-j]); { Setřídíme zbytek }

      j := i;
      dec(i);
      while (perm[i] > perm[j]) do
        inc(j); { Najdeme nejbližší vyšší prvek než prvek přilehlý }
      swap(perm[i], perm[j]); { Zaměníme je }
      writeln(perm);
    end else { Pokud takový neexistuje, permutace byla poslední }
      writeln('Dopermutovali jsme');
  end.

```

## Úloha 18-5-3 – Číňanské volby – program

```

program Cina;

const N = 100;
var hlasy : array[1..N] of integer;
    i, x, k, c : integer;

begin
  k := 0; { Najdeme kandidáta. }
  c := 0;
  for i := 1 to N do
    begin
      readln(x);
      hlasy[i] := x;
      if c = 0 then k := x;
      if k = x then inc(c) else dec(c);
    end;

  c := 0; { Ověříme, zda vyhrál. }
  for i := 1 to N do
    begin
      x := hlasy[i];
      if k = x then inc(c);
    end;

  if 2 * c > N then
    writeln('Vyhrál kandidát číslo ', k, '.')
  else
    writeln('Nikdo nevyhrál.').
end.

```

```

/* Counting words. MM scribebat me per III Id. Mai. MMDCLIX AUC. */

#include <stdio.h>
#include <stdlib.h>

struct state {
  struct state *fwd[256], *back; // jeden stav automatu
  struct state *qnext, *qprev; // hrany dopředné a zpětná
  char *out; // předchůdce a následník ve frontě
  int count; // které slovo v tomto stavu končí
  // počet průchodů stavem
};

struct state root; // počáteční stav (kořen stromu)
struct state *head, *tail; // první a poslední ve frontě

struct state *step(struct state *s, int c) // jeden krok automatu
{ // (včetně vracení se po zpětných hranách)
  while (s) {
    if (s->fwd[c]) // hrr na ně!
      return s->fwd[c];
    s = s->back; // a když to nejde, tož cónem
  }
  return &root;
}

void insert(char *x) // vložení jednoho slova do stromu
{
  struct state *s = &root;
  int c;
  for (char *y = x; c=*y++; ) {
    if (!s->fwd[c])
      s->fwd[c] = calloc(1, sizeof(struct state));
    s = s->fwd[c];
  }
  s->out = x;
}

void build(void) // konstrukce zpětných hran přesně podle kuchář
{
  struct state *s;
  head = tail = &root;
  while (head) {
    while (head) {
      for (int c=0; c<256; c++)
        if (s = head->fwd[c]) {
          s->back = step(head->back, c);
          tail->qnext = s;
          s->qprev = tail;
          tail = s;
        }
      head = head->qnext;
    }
  }
}

void run(void) // projití celého vstupu s počítáním průchodů s
{
  struct state *s = &root;
  for (int c; (c = getchar()) != EOF;) {
    s = step(s, c);
    s->count++;
  }
}

```