

Szalámi taktika

Hajnal Péter

Bolyai Intézet, TTIK, SZTE, Szeged

2023. ősz

Az alapötlet

Ismét a matematikai rekurziót valósítjuk meg algoritmuselméleti környezetben.

A szalámi taktika keresési problémák esetén alkalmazható. A keresett objektum azonosítása helyett „megelégszünk” elemek eldobásával, amelyek esetén meggyőződünk arról, hogy nem a keresett elem.

Célunk minél olcsóbban, minél több elemet eldobni. Az „olcsóság” és „sok” pontos/matematikai megfogalmazástól függ algoritmusunk hatékonysága.

Ismét példákon keresztül ismerhetjük meg ezen algoritmus tervezési elvet.

Bináris keresés alapproblémája

Bináris keresés

Adott $x_1 < x_2 < \dots < x_n$ rendezett szám n -es. Adott egy a szám. Döntsük el, hogy a az x_i számok között van-e. Ha igen, akkor adjuk meg azt a j indexet, amelyre $x_j = a$.

Az algoritmus $n = 2^k - 1$ szám esetén

Egy feltétellel élünk. Feltesszük, hogy $n = 2^k - 1$. Speciálisan a rendezett sorozatunkban van középső elem: $x_{2^{k-1}}$.

Ezt hasonlítsuk össze a -val. Ha egyenlőek, akkor szerencsések vagyunk: meg van a válasz. Ha $x_{2^{k-1}} < a$, akkor az $x_1 < x_2 < \dots < x_{2^{k-1}} < a$ elemek eldobhatók, a keresés redukálható a $x_{2^{k-1}+1} < x_{2^{k-1}+2} < \dots < x_n$ számokra.

Ezzel körülbelül megfeleztük számaink számát (az új „ $n = 2^k - 1$ ” kitevője $k - 1$ lesz). Az $x_{2^{k-1}} > a$ eset hasonlóan kezelhető. Ha $n = 2^2 - 1$, akkor egy összehasonlítás megadja, hogy a megvan-e.

Az algoritmus analízise

Összegezve legfeljebb k összehasonlítás után meg tudjuk válaszolni a kérdést.

Tétel

Algoritmusunk bonyolultsága $\mathcal{O}(\log n)$.

Bináris keresés alapproblémájának másik formája

A feladat gyakran úgy van megfogalmazva, hogy valaki gondolt egy 1 és 100 közötti egészre. Igen/nem kérdésekkel döntsük el, hogy melyik számra gondolt.

Szünet



Medián keresés alapproblémája

Adottak $x_1, x_2, \dots, x_n$ számok. Keressünk egy m indexet, hogy x_m az egyik medián legyen. Azaz

$$|\{i : x_i \leq x_m\}|, |\{i : x_i \geq x_m\}| \geq \left\lceil \frac{n}{2} \right\rceil.$$

teljesüljön.

Medián keresés alapproblémájának élesítése

A medián az egyik rendezett sorrendben a $\lceil n/2 \rceil$ -edik elem.
Feladatunk az egyik medián meghatározása.

A feladat megfogalmazásában rejlő határozatlanság abból ered, hogy számaink között azonos értékűek is lehetnek. Az azonos értékű számok egymás közötti sorrendje tetszőleges lehet a rendezett sorban.

Egy kicsit nehezítjük a feladatunkat: Általánosításunk két dologból ered. Az input számok mellett egy r „rang” is adott. Keressük meg az egyik számot, amely lehet a rendezett sorrend r -edik eleme. Ezen x_m elem megtalálása mellett osszuk a maradék számainkat K és N halmazokba, hogy minden $x_i \in K$ esetén $x_i \leq x_m$ (K elemei „kis” számok) és minden $x_i \in N$ esetén $x_i \geq x_m$ (N elemei „nagy” számok), továbbá $|K| = r - 1$ és így $|N| = n - r$.

Mediánok mediánja

n számainkat egészítsük ki (ha szükséges), hogy öttel osztható páratlan legyen számaink száma. Így n^+ számunk lesz, ahol $n \leq n^+ \leq n + 9$.

Ezeket osszuk ötös csoportokba. Minden ötös blokkban keressük meg a mediánt és a hozzá tartozó K és N halmazokat.

Legyen M az ötösök mediánjai által alkotott halmaz ($|M| = n^+/5$, páratlan szám). Keressük meg a M -beli x_i -k mediánját.

Jelöljük ezt x_μ -vel. Legyen K_M és N_M a M -beli elemek mediánjára vonatkozó kis, illetve nagy elemek halmaza.

$$|N_M| = |K_M| = (|M| - 1)/2.$$

Legyen $\hat{K}$ a K_M -beli elemek és az ezekhez tartozó (az ötösökből eredő) kis elemek halmaza. Hasonlóan legyen $\hat{N}$ a N_M -beli elemek és az ezekhez tartozó nagy elemek halmaza.

$\hat{K}$ mindegyik eleme legfeljebb x_μ , $\hat{N}$ mindegyik eleme legalább x_μ , továbbá $\hat{K}$ és $\hat{N}$ diszjunktak, közös elemszámuk $3(n^+ - 5)/10$.

A rekurzió

Hasonlítsuk össze x_μ -t az összes többi $\hat{K}$ és $\hat{N}$ -be nem tartozó számmal.

A x_μ -nél kisebb elemek és $\hat{K}$ alkossák a $\tilde{K}$ halmazt. A x_μ -nél nem kisebb elemek és $\hat{N}$ alkossák a $\tilde{N}$ halmazt.

Számaink rendezhetők úgy, hogy $\tilde{K}$ elemeivel kezdünk, majd következik x_μ , végül jönnek $\tilde{N}$ elemei.

Ha nagyon nagy szerencsénk van, akkor x_μ az r -edik elem és készen vagyunk. A realisabb feltevés, hogy meg tudjuk, hogy az r -edik elem $\tilde{K}$ elemei között vagy $\tilde{N}$ elemei között van. Sőt ezen halmazban hanyadik elemnek kell lennie a globális r -ediknek.

Egy kisebb feladatunk van. Biztos, hogy vagy $\hat{K}$, vagy $\hat{N}$ elemeit kizártuk. Tehát a kisebb feladatban szereplő számok száma legfeljebb $n^+ - 3(n^+ - 5)/10 \leq 7n/10$.

A rekurzió analízise

Legyen $M(n)$ az általánosított feladathoz szükséges összehasonlítások maximális száma n szám esetén. Összegezve eddigi lépéseinket

Tétel

$$M(n) \leq M(n/5) + M(7n/10) + O(n).$$

Következmény

$$M(n) = O(n).$$

Szünet



Konvex burok keresése

A konvex burok keresése (adott n pont a koordináta síkon, határozzuk meg felső burkolójukat) egy alapfeladat.

Az adott pontok x -koordinátáinak mediánját megkeresve a ponthalmazt egy függőleges egyenessel két azonos elemszámú ponthalmazra oszthatjuk. A két félre megoldva a problémát "csak" a két fél ponthalmaz felső burkolójának egyesítését kell megoldanunk (oszd meg és uralkodj).

A híd-keresés problémája

Probléma kitűzés

Adott $2n$ pont a síkon ($2n$ darab $P_i(x_i, y_i)$ pontpár). Adott még egy f függőleges egyenes, amely pontosan felezi ponthalmazunkat egy B (bal oldali) és egy J (jobb oldali) részre. Keressünk egy $e : ax + by + c = 0$ egyenest (feltesszük a $b > 0$), amely áthalad B -beli (P_B) és J -beli (P_J) ponton, továbbá teljes ponthalmazunk a $F : ax + by + c \leq 0$ zárt félsíkba esik.

A $P_B P_J$ szakaszt a két oldal közötti *hídnak* nevezzük. Ennek meghatározását kéri a feladat.

Az algoritmus eleje

B és J elemeit párosítsuk. Így n pontpárunk/szakaszunk lesz.

Számoljuk ki az n meredekséget.

Határozzuk meg m -et, a meredekségek mediánját.

Osszuk szakaszainkat $\mathcal{S}_{\leq m}$ és $\mathcal{S}_{\geq m}$ halmazba úgy, hogy $\sigma \in \mathcal{S}_{\leq m}$ szakasz meredeksége legfeljebb m legyen, míg $\sigma \in \mathcal{S}_{\geq m}$ szakasz meredeksége legalább m legyen, továbbá $|\mathcal{S}_{\leq m}| - |\mathcal{S}_{\geq m}| \leq 1$.

Vegyünk egy m meredekségű, felső támaszegyenesét pontthalmazunknak: t . Ez egy min/max probléma (miért?). Ez könnyen megoldható.

t lehet a keredett híd egyenese. Reálisabb feltevés, hogy pontthalmazunk vagy a bal vagy a jobb oldalon támasztja meg az egyenest.

A szalámizás

Tegyük fel, hogy a bal oldalon támasztja meg ponthalmazunk t -t.
(A másik lehetőség hasonlóan/duálisan tárgyalható.)

Ekkor "billentésekkel" kapjuk meg a hídat: Baloldali pontok körül forgatjuk t -t míg a jobb oldal is megtámasztja.

Ezek a forgatások csökkentik a meredekséget.

A korábbi pontpárok szakaszai közül a nagyobb meredekségűek olyanok, hogy a bal végpontjuk nem szerepelhet a hídon (miért?).

Ezek a végpontok "leszalámizhatóak".

Lineáris költséggel megszabadultunk lineáris sok ponttól.

Rekurzióval összeállt egy algoritmus.

Az analízis

Tétel

A híd-keresés költsége $\mathcal{O}(n)$.

Tétel

A konvex burok meghatározásának költsége $\mathcal{O}(n \log n)$.

Vége van!

Köszönöm a figyelmet!