Δημήτριος Κυριακίδης
el17077

Αλγόριθμοι και Πολυπλοκότητα

1º Φυλλάδιο Ασκήσεων

Άσκηση 1:

Master Theorem

Let $a \geq 1$ and $b > 1$ be constants, let $f(n)$ be a function, and let $T(n)$ be defined on the non-negative integers by the recurrence

$T(n) = aT(n/b) + f(n)$

where we interpret n/b to mean either $\lfloor n/b \rfloor$ or $\lceil n/b \rceil$. Then $T(n)$ has the following asymptotic bounds:

1. If $f(n) = O\left(n^{\log_b a - \varepsilon}\right)$ for some constant $\varepsilon > 0$, then $T(n) = \Theta\left(n^{\log_b a}\right)$.
2. If $f(n) = \Theta\left(n^{\log_b a}\right)$, then $T(n) = \Theta\left(n^{\log_b a} \log_2 n\right)$
3. If $f(n) = \Omega\left(n^{\log_b a + \varepsilon}\right)$ for some constant $\varepsilon > 0$, and if $af(n/b) \leq cf(n)$ for some constant $c < 1$ and all sufficiently large n, then $T(n) = \Theta\left(f(n)\right)$.

1. $T(n) = 4T(n/2) + \Theta(n^2 \log n)$
$f(n) = \Theta(n^2 \log n)$
$\log_b a = \log_2 4 = 2$, το n^2 δεν είναι πολυωνυμικά διαχωρίσιμο (μικρότερο ή μεγαλύτερο ασυμπτωματικά κατά παράγοντα n^{ε}) από το $n^2 \log n$. Επομένως δεν μπορεί να χρησιμοποιηθεί το master theorem. Χρησιμοποιώντας μέθοδο επανάληψης στην αναδρομική σχέση:
$T(n) = 4T(n/2) + \Theta(n^2 \log n) = 4\left(4(T(n/2^2) + (n/2)^2 \log(n/2))\right) + n^2 \log n =$
$$= \cdots = 4^{i+1} T\left(n/2^{i+1}\right) + n^2 \log n \sum_{k=0}^{i} \left(\frac{4}{2}\right)^k - \sum_{k=0}^{i} 4^k \log 2^k = \Theta(n^2 \log^2 n)$$

Άσκηση 2:

α.)

- Βρίσκουμε το μεγαλύτερο στοιχείο του Α.
- Με μία προθεματική περιστροφή, το φέρνουμε στην αριστερότερη θέση.
- Με μία ακόμη προθεματική περιστροφή, το μετακινούμε στην δεξιότερη θέση.
- Συνεχίζουμε αναδρομικά στον υποπίνακα $A[1..(n - 1)]$.

Συνολικά 2 περιστροφές για κάθε στοιχείο του πίνακα, άρα $2n$.

β.) Στον αλγόριθμο του προηγούμενου ερωτήματος, προσθέτουμε το πολύ μία προθεματική περιστροφή του εκάστοτε μεγαλύτερου στοιχείου όταν αυτό βρίσκεται στην αριστερότερη θέση, ώστε να αντιστραφεί το πρόσημό του.

γ.)

1. Θεωρούμε ότι ο τρέχων πίνακας δεν έχει συμβατά ζεύγη, γιατί αν είχε κάποιο, θα μπορούσαμε να το αντικαταστήσουμε με ένα στοιχείο (του οποίου το πρόσημο ταυτίζεται με το κοινό πρόσημο των στοιχείων του ζεύγους) και να καταλήξουμε σε ένα "μικρότερο" ισοδύναμο πρόβλημα.
Διακρίνουμε δύο περιπτώσεις.
Αν υπάρχει στοιχείο με θετικό πρόσημο:

- Βρίσκουμε το μεγαλύτερο (σε απόλυτη τιμή) θετικό στοιχείο, έστω x.
- Αν $x = n$, με δύο κινήσεις (βλ. ερώτημα (α)) δημιουργούμε το τετριμμένο συμβατό ζεύγος.
- Αλλιώς, το $x + 1$ έχει αρνητικό πρόσημο και έτσι μπορούμε να δημιουργήσουμε ένα νέο συμβατό ζεύγος με δύο κινήσεις το πολύ:
  - Αν το $-(x+1)$ βρίσκεται δεξιότερα του x, τότε κάνουμε μία προσημασμένη προθεματική περιστροφή που περιλαμβάνει μέχρι και το $-(x+1)$. Έτσι, έχουμε έναν πίνακα της μορφής: $[(x + 1), \dots, -x, \dots]$. Με μία κίνηση, σχηματίζουμε το νέο συμβατό ζεύγος $(-(x + 1), -x)$.
  - Διαφορετικά, η πρώτη μας κίνηση περιλαμβάνει μέχρι και το x (προκύπτει πίνακας της μορφής $[-x, \dots, (x + 1), \dots]$) και η δεύτερη σχηματίζει το νέο συμβατό ζεύγος $(x, x + 1)$.

Αν όλα τα στοιχεία έχουν αρνητικό πρόσημο:

- Από υπόθεση, δεν έχουμε τον πίνακα $[-1, -2, \dots, -n]$.
- Άρα, για κάποιο στοιχείο $-x$ $(x < n)$, το $-(x + 1)$ είναι αριστερότερα.
- Με δύο κινήσεις ($[\dots, -(x + 1), \dots, -x, \dots] \to [(x + 1), \dots, -x, \dots] \to [\dots, -(x + 1), -x, \dots]$), δημιουργούμε το νέο συμβατό ζεύγος $(-(x + 1), -x)$.

2. Ο συνολικός μας αλγόριθμος έχει τα ακόλουθα βήματα:

1. Απαλείφουμε όλα τα συμβατά ζεύγη αναδρομικά, αντικαθιστώντας τα με ένα προσημασμένο στοιχείο. Προκύπτει ισοδύναμο πρόβλημα, με μικρότερο πλήθος στοιχείων, έστω n'. Αν $n' = 0$, ο πίνακας έχει ταξινομηθεί.

2. Αν ο πίνακας είναι ο $[-1, -2, \ldots, -n']$, τότε:
   Επαναλαμβάνουμε για n' φορές:
   - Εκτέλεσε μία προσημασμένη προθεματική περιστροφή όλων των στοιχείων.
   - Εκτέλεσε μία προσημασμένη προθεματική περιστροφή των $n' - 1$ αριστερότερων στοιχείων.

   Ο πίνακας έχει ταξινομηθεί.
3. Αλλιώς, με βάση το ερώτημα 1, κατασκευάζουμε ένα νέο συμβατό ζεύγος και πηγαίνουμε στο βήμα 1 του αλγορίθμου.

Αρκεί να κατασκευάσουμε n συμβατά ζεύγη. Αν k το πλήθος επαναλήψεων του συνολικού αλγορίθμου, τότε το συνολικό πλήθος των κινήσεών μας είναι το πολύ: $2k + 2(n-k) = 2n$, γιατί σε κάθε επανάληψη (εκτός ίσως της τελευταίας) κάνουμε το πολύ δύο κινήσεις και κάθε μία από αυτές μειώνει το πλήθος των ζευγών που μας μένουν τουλάχιστον κατά 1. Στην τελευταία επανάληψη, πιθανώς να χρειαστεί να κάνουμε $2n' \le 2(n - k)$ κινήσεις.

<u>Άσκηση 6:</u>

Το b αποτελεί φυσική γενίκευση του a, άρα απαντάμε απ' ευθείας στο b.
$S \subset \mathbb{N}^+ \colon |S| = n$ (κλειδιά)
Έστω πίνακας A μεγέθους m, αρχικοποιημένος με 0, και $k \geq 1$ συναρτήσεις κατακερματισμού h_i:
$\mathbb{N}^+ \to [m] \colon \mathbb{P}[h_i(x) = j] = \dfrac{1}{m} \forall x \in \mathbb{N}^+, j \in [m]$
Η αρχικοποίηση της δομής γίνεται ως εξής: Για κάθε κλειδί $x \in S$, υπολογίζουμε το διάνυσμα:
$x \to (h_1(x), \ldots, h_k(x))$ (1)
και υλοποιούμε τη μέθοδο $Insert(x)$: Εισάγουμε 1 στο πίνακα A στις k θέσεις $h_i(x)$.
$Query(x)$: Για να ελέγξουμε αν το δοθέν κλειδί x ανήκει στο S, ελέγχουμε αν $A[h_1(x)] \wedge \ldots \wedge A[h_k(x)] = 1$. Απαντάμε «ναι» ή «όχι» αναλόγως.
Η δομή αυτή απαντά πάντα ορθά αν το κλειδί ανήκει στο σύνολο. Υπάρχει όμως και το ενδεχόμενο να απαντήσει «ναι», ακόμα και αν $x \notin S$ (false positive).
Για να τεθεί ένα τυχόν bit $i \in [m]$ ίσο με 1, αρκεί ένα από τα n κλειδιά να γίνει map στο bit αυτό.
Αυτό συμβαίνει με πιθανότητα: $\mathbb{P}[A[i] = 1] = 1 - \left(1 - \frac{1}{m}\right)^{nk} \simeq 1 - e^{-\frac{nk}{m}}$ (2)
$\mathbb{P}[\{A[h_1(x) = 1, \ldots, A[h_k(x) = 1]\}] = \prod_{i=1}^{k} \mathbb{P}[A[h_i(x) = 1]] \simeq \left(1 - e^{-\frac{nk}{m}}\right)^k$

Βελτιστοποιούμε την συνάρτηση ως προς k (είναι κυρτή), θέτοντας $m = 8n$:

$\frac{d}{dk}\left(1 - e^{-\frac{k}{8}}\right)^k = 0 \Rightarrow k \simeq 5.54$ (δομή "bloom filter").