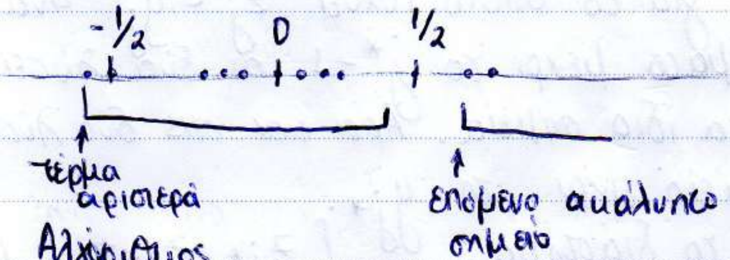
ΑΛΓΟΡΙΘΜΟΙ

• ΦΡΟΝΤΙΣΤΗΡΙΟ 6 •

Άσκηση 1/5 (Άπληστοι Αλγόριθμοι)

Δίνονται σημεία $\{x_1, x_2, \dots x_n\}$ σε μια ευθεία. Να δώσετε άπληστο αλγόριθμο που βρίσκει ένα σύνολο μοναδιαίων διαστημάτων που περιλαμβάνει όλα τα σημεία και χρησιμοποιεί όσο το δυνατόν λιγότερα διαστήματα. Να δείξετε την ορθότητα και να υπολογίσετε την πολυπλοκότητα του αλγορίθμου σας.

Λύση



Αλγόριθμος

```
line_cover (x = [x1, x2, ..., xn])
    x = sort(x)      → ταξινομούμε τα δωσμένα σημεία
    y1 = x1          → θέτουμε ως πρώτο διάστημα εκείνο που
                       αρχίζει από το πρώτο ακάλυπτο σημείο
                       (ξεκινάω να καλύπτω από τέρμα αριστερά).
    i = 1
    j = 1
    while (j<n)          Σε κάθε επανάληψη βρίσκω
      if xj ≤ yi + 1     το επόμενο ακάλυπτο σημείο
          j = j + 1      (το πρώτο σημείο δεξιά αυτού του
      else               διαστήματος), καλύπτω και
          i = i + 1      συνεχίζω.
          yi = xj
          j = j + 1
    return (y1, y2, ... yi)
```

①

Πολυπλοκότητα

sort(x) (ταξινόμηση) της τάξης $O(n\log n)$

while $(j<n)$ → η επανάληψη σταθερού χρόνου η καθεμία $\Rightarrow \Theta(n)$

Άρα συνολική πολυπλοκότητα $O(n\log n)$.

Απόδειξη ορθότητας

Έστω $(y_1, y_2, \dots y_i)$ η λύση του άπληστου αλγορίθμου που ΔΕΝ είναι βέλτιστη. Έστω πως η βέλτιστη λύση είναι η $(z_1, z_2, \dots, z_{i'})$.

Έστω ότι το πρώτο j για το οποίο ισχύει $z_j < y_j$ είναι το j^*. Από τα διαστήματα μέχρι το j^*-1 οι δύο λύσεις καλύπτουν ΑΚΡΙΒΩΣ τα ίδια σημεία. Άρα και στις δύο λύσεις το πρώτο ακάλυπτο σημείο είναι το y_{j^*}.

Άρα αντικαθιστώντας το διάστημα $[z_{j^*}, z_{j^*}+1]$ με το $[y_{j^*}, y_{j^*}+1]$ στη βέλτιστη λύση, δεν αφήνουμε κανένα σημείο πριν το j^* ακάλυπτο.

Επειδή το διάστημα του y_{j^*} καλύπτει και σημεία μετά το $z_{j^*}+1$, σίγουρα η νέα λύση που προκύπτει ΔΕΝ έχει ακάλυπτα σημεία ΧΩΡΙΣ να έχω αυξήσει το πλήθος των διαστημάτων, άρα παραμένει βέλτιστη.

Επαναλαμβάνοντας όσες φορές χρειάζεται, αντικαθιστώ όλα τα z_j με τα αντίστοιχα y_i, και παίρνω τη βέλτιστη λύση $(y_1, y_2, \dots, y_{i'})$.

Στον αλγόριθμο βρήκαμε ότι το $y_{i'+1}$ ήταν ακάλυπτο και για αυτό πήραμε τα παραπάνω διαστήματα. Δηλαδή η λύση $(y_1, y_2, \dots, y_{i'})$ δεν καλύπτει όλα τα σημεία και άρα δεν είναι βέλτιστη). ΑΤΟΠΟ!

Άρα ο αλγόριθμος βρίσκει μία βέλτιστη λύση.

Άσκηση 2/5 (Άπληστοι Αλγόριθμοι).

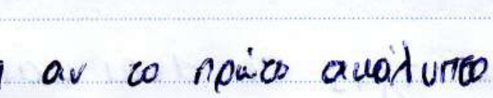
Δίνονται n σημεία σε μία ευθεία (πχ σπίτια κατά μήκος ενός δρόμου). Να σχεδιάσετε άπληστο αλγόριθμο που υπολογίζει το ελάχιστο πλήθος τηλεφωνικών θαλάμων που μπορώ να τοποθετήσω ώστε κανένα σπίτι να μην απέχει πάνω από 5 km από τον πλησιέστερο θάλαμο. Τα σημεία της ευθείας δίνονται ταξινομημένα. Να δείξετε την ορθότητα και να υπολογίσετε την πολυπλοκότητα του αλγορίθμου σας.

Λύση

Αλγόριθμος

Όσο υπάρχουν ακάλυπτα σημεία:
Βρες το πρώτο ακάλυπτο, έστω x_j, και
βάλε ένα θάλαμο στο σημείο $x_j + 5$ και
σβήσε όλα τα σημεία μέχρι το $x_j + 10$.

Έτσι αν το πρώτο ακάλυπτο είναι το x_j:



βάζοντας ένα θάλαμο στο $x_j + 5$ (δηλαδή σε απόσταση 5km), όλα τα σπίτια από x_j έως $x_j + 10$ θα καλύπτονται από τον θάλαμο στο $x_j + 5$, για αυτό μπορούμε να τα αγνοήσουμε/διαγράψουμε τα σημεία αυτά.

Πολυπλοκότητα: $O(n)$, ελέγχω τα σημεία από μία φορά, σε σταθερό χρόνο.

Απόδειξη ορθότητας

Έστω ότι ο άπληστος αλγόριθμος βάζει θαλάμους στα σημεία $t_1 < t_2 < \dots < t_k$ και η βέλτιστη λύση είναι τα σημεία $s_1 < s_2 < \dots < s_m$. Επειδή ο αλγόριθμός μας καλύπτει όλα

τα x_i, αρκεί να δείξουμε ότι $k=m$, δηλαδή ότι το πλήθος των θαλάμων που χρησιμοποιεί ο αλγόριθμος μας είναι ίδιο με τη βέλτιστη λύση.
Θα δείξουμε επαγωγικά ότι για κάθε j από 1 μέχρι k, τα x_i που καλύπτει η λύση $(t_i, \ldots, t_k)$ είναι τουλάχιστον όσα καλύπτει η λύση $(s_i, \ldots, s_m)$.

Βάση επαγωγής
Το s_1 και το t_1 είναι το πολύ ίσα με $x_1 + 5$ αφού καλύπτουν το σημείο x_1. Άρα το t_1 καλύπτει μέχρι $t_1 + 5 = x_1 + 0 \geq s_1 + 5$, δηλαδή πράγματι το t_1 καλύπτει τουλάχιστον τόσα στοιχεία όσα το s_1.

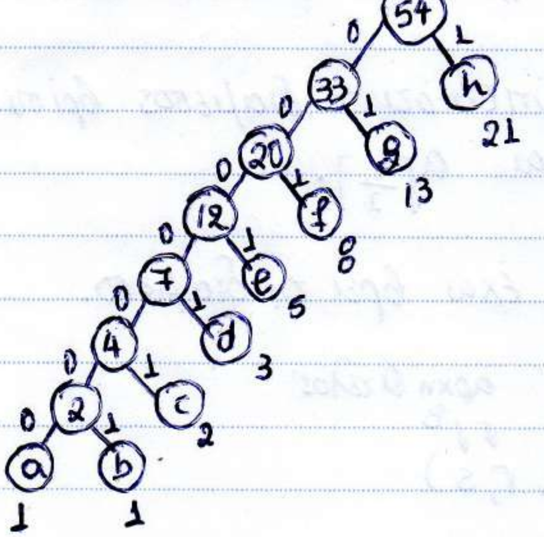
Επαγωγικό βήμα
Το πρώτο που καλύπτει το t_{j+1} θα είναι κάποιο σημείο x_i που είναι δεξιότερα από το πρώτο που καλύπτει το s_{j+1}, έστω το $x_{i'}$.
Άρα η λύση του Greedy καλύπτει μέχρι το ...
$t_{j+1} + 5 = x_i + 10 \geq x_{i'} + 10 \geq s_{j+1} + 5$.
(αφού το s_{j+1} έχει μπει το πολύ στη θέση $x_{i'} + 5$)

Άρα το $(s_1, \ldots, s_k)$ καλύπτει όλα τα σημεία και άρα ταυτίζεται με το $(s_1, \ldots, s_m)$, δηλαδή $k=m$ και συνεπώς ο αλγόριθμος είναι βέλτιστος.

Άσκηση 3/5 (Άπληστοι Αλγόριθμοι)

Ποια είναι μια βέλτιστη κωδικοποίηση Huffman για την ακολουθία συχνοτήτων εμφάνισης που δίνεται από τους πρώτους 8 αριθμούς Fibonacci, δηλαδή a:1, b:1, c:2, d:3, e:5, f:8, g:13, h:21; Γενικεύστε τη λύση σας για τους n πρώτους αριθμούς Fibonacci.

Λύση



a: 0000000
b: 0000001
c: 000001
d: 00001
e: 0001
f: 001
g: 01
h: 1

Έστω n=40. Το σύμβολο s με συχνότητα ίση με του 37ο αριθμό Fibonacci έχει κωδικοποίηση της μορφής 00...01.
Αναλυτικά στο παράδειγμα έχουμε:

40ος → 1 σύμβολο/bits (1)
39ος → 2 -//- (01)
38ος → 3 -//- (001)
37ος → 4 -//- (0001)
Άρα 0001
⋮
2ος → 39 bits
1ος → 39 bits όλα 0.

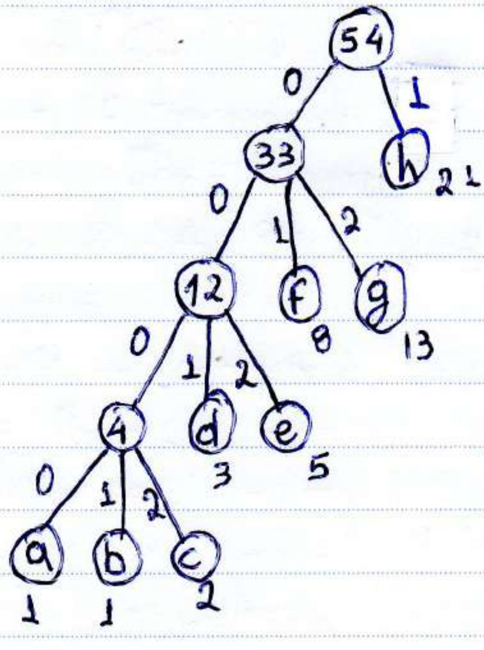
Άρα η γενικευμένη λύση είναι το 1ο στοιχείο να έχει n-1 στοιχεία (όλα 0), το 2ο n-1 (το τελευταίο 1) κι από αυτό το σημείο ανά "επίπεδο" αφαιρείται ένα 0.

Άσκηση 4/5 (Άπληστοι Αλγόριθμοι)

Γενικεύστε τον αλγόριθμο Huffman για τρία ψηφία κωδικοποίησης (κωδικοποίηση με τα ψηφία 0, 1 και 2). Αποδείξτε ότι παράγει βέλτιστες κωδικοποιήσεις.

Λύση

0, 1, 2



a: 0000
b: 0001
c: 0002
d: 001
e: 002
f: 01
g: 02
h: 1

8 σύμβολα → 4 επαναλήψεις
Στην άσκηση 3 για 8 σύμβολα κάναμε 7 επαναλήψεις.