Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Τεχνολογίας Πληροφορικής και Υπολογιστών
Αλγόριθμοι και Πολυπλοκότητα
Διδάσκοντες: Δημήτρης Φωτάκης, Δώρα Σούλιου
2η Σειρά Γραπτών Ασκήσεων - Ημ/νία Παράδοσης 14/12/2021

Άσκηση 1: Δίσκοι και Σημεία (1.2 μον.)

Δίνονται n (διαφορετικά μεταξύ τους) σημεία και μία ευθεία l στο επίπεδο. Κάθε ένα από τα n σημεία βρίσκεται σε (κάθετη) απόσταση το πολύ r από την ευθεία l. Θέλουμε να καλύψουμε όλα τα n σημεία με ελάχιστο πλήθος δίσκων ακτίνας r, των οποίων τα κέντρα θα βρίσκονται πάνω στην ευθεία l. Πιο συγκεκριμένα, πρέπει να τοποθετήσουμε δίσκους ακτίνας r, κατά μήκος της ευθείας l, με το κέντρο κάθε δίσκου να ανήκει στην l, ώστε καθένα από τα n σημεία να καλύπτεται από τουλάχιστον έναν τέτοιο δίσκο. Αντικειμενικός μας στόχος είναι να ελαχιστοποιήσουμε το πλήθος των δίσκων που θα χρησιμοποιήσουμε. Να διατυπώσετε έναν αποδοτικό αλγόριθμο για αυτό το πρόβλημα. Να αιτιολογήσετε αναλυτικά την ορθότητα και την υπολογιστική πολυπλοκότητα του αλγορίθμου σας.

Άσκηση 2: Παραλαβή Πακέτων (2.2 μον.)

Η Black Friday και η Cyber Monday ήταν πραγματικά θριαμβευτικές για την επιχείρησή σας. Οι πωλήσεις ξεπέρασαν τις προσδοκίες σας και τώρα είναι η στιγμή που οι πελάτες σας θα παραλάβουν τα δέματά τους. Από το πρωί, έχουν μαζευτεί έξω (λόγω covid) από το κατάστημά σας n πελάτες που πρέπει να παραλάβουν τα δέματά τους. Η προετοιμασία του δέματος κάθε πελάτη i χρειάζεται χρόνο $p_i \in \mathbb{N}^*$, τον οποίο γνωρίζετε με ακρίβεια. Η πελατεία σας είναι σχετικά σταθερή, οπότε για κάθε πελάτη i έχετε εκτιμήσει ένα θετικό βάρος $w_i \in \mathbb{N}^*$, που ποσοτικοποιεί τη σημασία του i για την επιχείρησή σας.

Θέλετε να δρομολογήσετε την παραλαβή των δεμάτων ώστε να ελαχιστοποιηθεί ο συνολικός βεβαρυμένος χρόνος εξυπηρέτησης για τους πελάτες σας. Αν π.χ. έχουμε διαθέσιμο έναν μόνο υπάλληλο και 4 πελάτες, που εξυπηρετούνται με σειρά $(1, 2, 3, 4)$, ο συνολικός βεβαρυμένος χρόνος εξυπηρέτησης είναι:

$$\begin{aligned} w_1p_1+&w_2(p_1+p_2)+w_3(p_1+p_2+p_3)+w_4(p_1+p_2+p_3+p_4) = \\ &p_1(w_1+w_2+w_3+w_4)+p_2(w_2+w_3+w_4)+p_3(w_3+w_4)+p_4w_4\,. \end{aligned}$$

Αν π.χ. έχουμε δύο υπαλλήλους, ο πρώτος εξυπηρετεί τους πελάτες 1 και 2 με σειρά $(1, 2)$, και ο δεύτερος εξυπηρετεί τους πελάτες 3 και 4 με σειρά $(4, 3)$, ο συνολικός βεβαρυμένος χρόνος εξυπηρέτησης είναι:

$$w_1p_1 + w_2(p_1 + p_2) + w_4p_4 + w_3(p_4 + p_3)\,.$$

1. Να διατυπώσετε αποδοτικό αλγόριθμο που δρομολογεί την παραλαβή των δεμάτων με στόχο την ελαχιστοποίηση του βεβαρυμένου χρόνου εξυπηρέτησης, αν έχετε μόνο έναν υπάλληλο που προετοιμάζει τα δέματα και εξυπηρετεί τους πελάτες.
2. Να διατυπώσετε αποδοτικό αλγόριθμο για το ίδιο πρόβλημα, αν έχετε δύο υπαλλήλους στην εξυπηρέτηση. Να σκιαγραφήσετε τη γενίκευση του αλγορίθμου σας για την περίπτωση που έχετε $m \geq 3$ υπαλλήλους.

Να αιτιολογήσετε αναλυτικά την ορθότητα και την υπολογιστική πολυπλοκότητα των αλγορίθμων σας.

Άσκηση 3: Τοποθέτηση Στεγάστρων (και Κυρτό Κάλυμμα) (2.3 μον.)

Η διοίκηση του Πανεπιστημίου αποφάσισε να τοποθετήσει στέγαστρα στον Πεζόδρομο μπροστά στη Σχολή, για να μην μας καίει ο ήλιος το καλοκαίρι και να μην βρεχόμαστε το χειμώνα. Ο Πεζόδρομος εκτείνεται σε μία ευθεία από το σημείο $x_0 = 0$ έως το σημείο $x_f = L > 0$. Θέλουμε να καλύψουμε με στέγαστρα n συγκεκριμένα σημεία του Πεζόδρομου με συντεταγμένες $0 < x_1 < x_2 < \cdots < x_n < L$. Αυτό μπορεί να επιτευχθεί με διάφορους τρόπους, όπως π.χ., με ένα μεγάλο στέγαστρο που εκτείνεται από το σημείο x_0 έως το σημείο x_f ή με n "σημειακά" στέγαστρα που καλύπτουν μόνο τα σημεία x_i. Το κόστος για να κατασκευαστεί και να τοποθετηθεί ένα νέο στέγαστρο που εκτείνεται από κάποιο σημείο a έως κάποιο σημείο b δίνεται από την σχέση $(a-b)^2+C$, όπου C κάποια θετική σταθερά. Σημειώνεται ότι μπορεί ένα στέγαστρο να καλύπτει μόνο ένα σημείο, δηλαδή να έχουμε $a = b$, οπότε το κόστος κατασκευής και εγκατάστασής του είναι C.

(α) Να διατυπώσετε αποδοτικό αλγόριθμο που με είσοδο τα σημεία $0 = x_0 < x_1 < x_2 \cdots < x_n < x_f = L$ και τη σταθερά C, υπολογίζει το ελάχιστο κόστος για την κάλυψη των σημείων $x_1, \ldots, x_n$ με στέγαστρα.

(β) Θεωρούμε ευθείες $y = ax + b$ στο $\mathbb{Q}^2$, οι οποίες περιγράφονται από το ζεύγος $(a, b) \in \mathbb{Q}^2$. Να διατυπώσετε αλγόριθμο γραμμικού χρόνου $\Theta(n + k)$, που με είσοδο ένα σύνολο n ευθειών $(a_j, b_j) \in \mathbb{Q}^2$, $j = 1, \cdots, n$, για τις οποίες ισχύει ότι $0 \geq a_1 \geq a_2 \cdots \geq a_n$, και ένα σύνολο k σημείων $x_i \in Q$, με $x_1 \leq x_2 \leq \cdots \leq x_k$, υπολογίζει τις ευθείες $(a_{j(i)}, b_{j(i)})$ που επιτυγχάνουν ελάχιστη τιμή σε καθένα από τα σημεία x_i. Πρέπει, δηλαδή, ο αλγόριθμός σας να επιστρέφει k δείκτες $j(i) = \arg\min_{j=1,\ldots,n}\{a_j x_i + b_j\}$, $i = 1, \ldots, k$. Στη συνέχεια να εξηγήσετε πως μπορούμε να χρησιμοποιήσουμε αυτόν τον αλγόριθμο για να βελτιώσουμε την υπολογιστική πολυπλοκότητα του αλγόριθμου στο (α).

Άσκηση 4: Δρομολόγια Λεωφορείων στην Εποχή του Κορωνοϊού (1.6 μον.)

(α) Στη στάση της Κατεχάκη υπάρχουν n φοιτητές, οι οποίοι έχουν σχηματίσει μία ουρά και περιμένουν k λεωφορεία τα οποία θα τους μεταφέρουν στη Σχολή. Οι πρώτοι q_1 φοιτητές θα μπουν στο 1ο λεωφορείο, οι επόμενοι q_2 στο 2ο κοκ., μέχρι τους τελευταίους q_k φοιτητές οι οποίοι θα μπουν στο k-οστό λεωφορείο (πρέπει να ισχύει ότι $q_\ell > 0$, για κάθε $\ell \in \{1, \ldots, k\}$, και ότι $\sum_{\ell=1}^{k} q_\ell = n$). Είναι γνωστό ότι λόγω covid, πρέπει να τηρούνται αποστάσεις μεταξύ των επιβατών στα λεωφορεία. Με βάση την κατάσταση της υγείας, τη δική τους και των οικείων τους, υπάρχει ένας δείκτης ευαισθησίας A_{ij} για κάθε ζεύγος φοιτητών (i, j) στο ίδιο λεωφορείο (ισχύει ότι $A_{ii} = 0$ και ότι $A_{ij} = A_{ji} > 0$, για κάθε $i \neq j$). Ο δείκτης ευαισθησίας ενός λεωφορείου είναι το άθροισμα των δεικτών ευαισθησίας για κάθε ζεύγος φοιτητών στο συγκεκριμένο λεωφορείο, ενώ ο συνολικός δείκτης ευαισθησίας μιας ανάθεσης των n φοιτητών στα k λεωφορεία δίνεται από το άθροισμα των δεικτών ευαισθησίας για όλα τα λεωφορεία. Δοθέντων των n, k και A_{ij}, θέλουμε να υπολογίσουμε μια ανάθεση $(q_1, \ldots, q_k)$ των n φοιτητών στα k λεωφορεία που ελαχιστοποιεί τον συνολικό δείκτη ευαισθησίας. Να διατυπώσετε αποδοτικό αλγόριθμο για αυτό το πρόβλημα και να αιτιολογήσετε την ορθότητα και την υπολογιστική πολυπλοκότητα του αλγορίθμου σας.

Πως θα προσεγγίζατε το πρόβλημα (και ποια θα ήταν η υπολογιστική πολυπλοκότητα του αντίστοιχου αλγόριθμου), αν το ζητούμενο είναι μια ανάθεση $(q'_1, \ldots, q'_k)$ των n φοιτητών στα k λεωφορεία που ελαχιστοποιεί τον μέγιστο δείκτη ευαισθησίας των λεωφορείων;

(β – bonus) Έστω $\mathrm{Last}(i, \ell) \in \{1, \ldots, i-1\}$, για κάθε $i \in \{2, \ldots, n\}$ και κάθε $\ell \in \{2, \ldots, k\}$, ο τελευταίος φοιτητής που μπαίνει στο λεωφορείο $\ell - 1$ στη βέλτιστη λύση για τους πρώτους i φοιτητές και ℓ λεωφορεία. Να δείξετε ότι για κάθε για κάθε $i \in \{2, \ldots, n\}$ και κάθε $\ell \in \{2, \ldots, k\}$,

$$\mathrm{Last}(i-1, \ell) \leq \mathrm{Last}(i, \ell)\,.$$

Με βάση αυτήν την ιδιότητα, να τροποποιήσετε τον αλγόριθμο που υπολογίζει μια ανάθεση με ελάχιστο συνολικό δείκτη ευαισθησίας ώστε να έχει χρόνο εκτέλεσης $O(n^2 + nk \log n)$.

Άσκηση 5: Το Σύνολο των Συνδετικών Δέντρων (2.7 μον.)

Θεωρούμε ένα συνεκτικό μη κατευθυνόμενο γράφημα $G(V, E)$ με n κορυφές και m ακμές.

(α) Έστω T_1 και T_2 δύο διαφορετικά συνδετικά δέντρα του G. Να δείξετε ότι για κάθε ακμή $e \in T_1 \setminus T_2$, υπάρχει ακμή $e' \in T_2 \setminus T_1$, τέτοια ώστε το $(T_1 \setminus \{e\}) \cup \{e'\}$ είναι συνδετικό δέντρο. Να διατυπώσετε αποδοτικό αλγόριθμο που με δεδομένα τα T_1, T_2 και e, υπολογίζει μια τέτοια ακμή e'.

(β) Σχηματίζουμε γράφημα H που κάθε κορυφή του αντιστοιχεί σε ένα διαφορετικό συνδετικό δέντρο του G. Δύο συνδετικά δέντρα T_1 και T_2 του G (κορυφές του H) συνδέονται με ακμή στο H αν διαφέρουν κατά μία μόνο ακμή, δηλ. αν $|T_1 \setminus T_2| = |T_2 \setminus T_1| = 1$. Να δείξετε ότι το H είναι συνεκτικό και ότι η απόσταση (στο H) μεταξύ δύο συνδετικών δέντρων T_1 και T_2 του G είναι ίση με $|T_1 \setminus T_2|$. Να εξηγήσετε πως θα χρησιμοποιήσουμε τον αλγόριθμο του (α) για να υπολογίσουμε ένα συντομότερο μονοπάτι (στο H) μεταξύ των T_1 και T_2.

(γ) Θεωρούμε τώρα ότι το γράφημα G έχει βάρη στις ακμές του. Θεωρούμε, δηλαδή, ένα γράφημα $G(V, E, w)$, με n κορυφές, m ακμές και θετικό βάρος $w(e)$ σε κάθε ακμή $e \in E$. Να διατυπώσετε αποδοτικό αλγόριθμο, που για κάθε ακμή $e \in E$, υπολογίζει το ελάχιστο βάρος ενός συνδετικού δέντρου T_e του G που περιέχει την ακμή e. Ο αλγόριθμός σας θα υπολογίζει το ελάχιστο βάρος ενός τέτοιου συνδετικού δέντρου T_e, για **όλες τις ακμές** $e \in E$, και θα επιστρέφει m ζεύγη της μορφής $(e_1, w(T_{e_1})), \dots, (e_m, w(T_{e_m}))$, όπου $w(T_{e_j})$ είναι το συνολικό βάρος ενός ελάχιστου συνδετικού δέντρου T_{e_j} του G που περιέχει την ακμή $e_j \in E$. Να αιτιολογήσετε αναλυτικά την ορθότητα και την υπολογιστική πολυπλοκότητα του αλγορίθμου σας. Θα υπάρχει επιπλέον βαθμολογία (bonus) για απαντήσεις με χρονική πολυπλοκότητα $\mathrm{O}(|E| \log |E|)$.

Άσκηση 6: Υποπίνακας με Μέγιστο Άθροισμα Στοιχείων (bonus)

(α) Δίνεται πίνακας $A[1 \dots n]$ με n ακέραιους (θετικούς ή αρνητικούς) αριθμούς. Θέλουμε να υπολογίσουμε τον υποπίνακα με το μεγαλύτερο άθροισμα στοιχείων. Αν το κάνουμε εξαντλητικά, πόσους υποπίνακες θα χρειαστεί να ελέγξουμε; Να διατυπώσετε αλγόριθμο γραμμικού χρόνου για την εύρεση του υποπίνακα με το μέγιστο άθροισμα στοιχείων.

(β) Θέλουμε να λύσουμε το ίδιο πρόβλημα, αλλά σε διδιάστατο πίνακα $B[1 \dots n, 1 \dots n]$ με n^2 ακέραιους αριθμούς. Συγκεκριμένα, μας ενδιαφέρει να υπολογίσουμε τον δισδιάστατο (όχι απαραίτητα τετραγωνικό) υποπίνακα του B με μέγιστο άθροισμα στοιχείων. Πόσους υποπίνακες χρειάζεται να ελέγξουμε, αν προσπαθήσουμε να λύσουμε το πρόβλημα εξαντλητικά; Χρησιμοποιώντας το (α), να διατυπώσετε αλγόριθμο με χρονική πολυπλοκότητα $O(n^3)$ για αυτό το πρόβλημα και να αιτιολογήσετε την ορθότητα του αλγορίθμου σας.

Άσκηση 7: Ένας Παράξενος Περίπατος (bonus)

Δίνεται κατευθυνόμενο γράφημα $G(V, E, c)$ με θετικά ακέραια βάρη c στις κορυφές του, όπου $c(v)$ είναι το βάρος μιας κορυφής $v \in V$. Ορίζουμε ως συνάρτηση κόστους ενός περιπάτου στο G τον *μέγιστο κοινό διαιρέτη* του βάρους των κορυφών που ανήκουν στον περίπατο. Π.χ., το κόστος ενός περιπάτου $w = (v_1, v_2, \dots, v_k)$ είναι $c(w) = \gcd(c(v_1), c(v_2), \dots, c(v_k))$, όπου gcd είναι ο μέγιστος κοινός διαιρέτης ενός συνόλου θετικών ακεραίων. Να διατυπώσετε αποδοτικό αλγόριθμο που υπολογίζει το ελάχιστο κόστος ενός περιπάτου στο γράφημα $G(V, E, c)$. Να αιτιολογήσετε την ορθότητα και την υπολογιστική πολυπλοκότητα του αλγορίθμου σας.

Άσκηση 8: Προσεγγίσεις του Προβλήματος Σακιδίου (bonus)

Θεωρούμε στιγμιότυπο του (Διακριτού) Προβλήματος Σακιδίου (Discrete Knapsack) με σακίδιο μεγέθους B και n αντικείμενα, όπου κάθε αντικείμενο i έχει μέγεθος $s_i \in \mathbb{N}^*$, με $s_i \leq B$, και αξία $p_i \in \mathbb{N}^*$. Συμβολίζουμε με $p_{\max} = \max_i\{p_i\}$ τη μέγιστη αξία ενός αντικειμένου.

(α) Θεωρούμε τον αλγόριθμο που επιλέγει την καλύτερη από τις παρακάτω λύσεις: (i) ένα αντικείμενο με αξία $p_{\max}$, και (ii) την άπληστη λύση, όπου εξετάζουμε τα αντικείμενα σε φθίνουσα σειρά p_i/s_i και συμπεριλαμβάνουμε στο σακίδιο κάθε αντικείμενο με μέγεθος που δεν ξεπερνά την υπολειπόμενη χωρητικότητα του σακιδίου. Να δείξετε ότι αυτός ο αλγόριθμος επιτυγχάνει τουλάχιστον το $\frac{1}{2}$ της αξίας της βέλτιστης λύσης.

(β) Για σταθερό $k \geq 2$, θεωρούμε τον αλγόριθμο που επιλέγει την καλύτερη από τις παρακάτω λύσεις: (i) τη μέγιστη αξία που μπορούμε να επιτύχουμε συμπεριλαμβάνοντας το πολύ k αντικείμενα στο σακίδιο (μπορεί να υπολογισθεί σε χρόνο $O(kn^k)$ με εξαντλητική αναζήτηση), και (ii) την άπληστη λύση (όπως στο (α)). Να δείξετε ότι αυτός ο αλγόριθμος επιτυγχάνει τουλάχιστον το $\frac{k-1}{k}$ της αξίας της βέλτιστης λύσης.

(γ) Να δείξετε πως (με χρήση δυναμικού προγραμματισμού) μπορούμε να υπολογίσουμε το ελάχιστο μέγεθος σακιδίου που χρειαζόμαστε για συνολική αξία τουλάχιστον P από τα πρώτα i αντικείμενα (ας συμβολίσουμε αυτή την ποσότητα με $B(i, P)$). Ποια είναι η υπολογιστική πολυπλοκότητα για τον υπολογισμό της βέλτιστης λύσης σε αυτή την περίπτωση;

(δ) Έστω σταθερά $\epsilon > 0$ και $d = \epsilon\, p_{\max}/n$. Θέτουμε $\hat{p}_i = \lfloor p_i/d \rfloor$, και εφαρμόζουμε τον αλγόριθμου του (γ) για μέγεθος σακιδίου B, μεγέθη αντικειμένων s_i και αξίες αντικειμένων $\hat{p}_i$. Ποια είναι η υπολογιστική πολυπλοκότητα του τροποποιημένου αλγόριθμου; Να δείξετε ότι η συνολική αξία (με βάση τα αρχικά p_i) της λύσης που υπολογίζει ο τροποποιημένος αλγόριθμος είναι τουλάχιστον $(1-\epsilon)P^*$, όπου P^* η συνολική αξία μιας βέλτιστης λύσης για το αρχικό στιγμιότυπο.

Άσκηση 9: Προσέγγιση του Προβλήματος Πλανόδιου Πωλητή (bonus)

Θεωρούμε το Πρόβλημα του Πλανόδιου Πωλητή (Travelling Salesman Problem, TSP) με σύνολο πόλεων $V = \{v_1, \ldots, v_n\}$ και αποστάσεις $d : V \times V \to \mathbb{N}^*$ μεταξύ των πόλεων που είναι μη-αρνητικές, συμμετρικές (δηλ. $d(v_i, v_j) = d(v_j, v_i)$ για κάθε $v_i, v_j \in V$) και ικανοποιούν την τριγωνική ανισότητα (δηλ. $d(v_i, v_j) + d(v_j, v_\ell) \leq d(v_i, v_\ell)$ για κάθε $v_i, v_j, v_\ell \in V$). Συμβολίζουμε με D^* το μήκος μιας βέλτιστης περιοδείας στο V.

(α) Αντιμετωπίζουμε το στιγμιότυπο του TSP ως πλήρες γράφημα G με σύνολο κορυφών V και μήκος $d(v_i, v_j)$ για κάθε ακμή $\{v_i, v_j\}$, και υπολογίζουμε ένα Ελάχιστο Συνδετικό Δέντρο T^* του G. Να δείξετε ότι $d(T^*) \leq (1 - 1/n)D^*$, όπου $d(T^*)$ είναι το συνολικό μήκος των ακμών που έχουν συμπεριληφθεί στο T^*.

(β) Θεωρήστε ένα Euler tour στο T^* για να υπολογίσετε (με χρήση και της τριγωνικής ανισότητας) μια περιοδεία στο V με μήκος που δεν ξεπερνά το $2d(T^*) \leq 2(1-1/n)D^*$. Ποια ιδιότητα χρησιμοποιήσατε για το Euler tour; Να δείξετε ότι υπάρχει βελτιωμένο Euler tour που οδηγεί σε περιοδεία μήκους το πολύ $D^*/2 + d(T^*)$.