Πρώτη εργασία στο μάθημα “Αλγόριθμοι και Πολυπλοκότητα”

Το πρόβλημα της μέγιστης υποακολουθίας, θεωρητική και εμπειρική μελέτη

Γκόγκος Χρήστος
Τμήμα Πληροφορικής και Τηλεπικοινωνιών, Πανεπιστημίου Ιωαννίνων

Αρτα, Οκτώβριος 2021

1 Εισαγωγή

Το πρόβλημα της μέγιστης υποακολουθίας (maximum subarray) είναι ένα πρόβλημα στο οποίο δίνεται μια ακολουθία τιμών (αρνητικές, θετικές ή μηδέν) και ζητείται να βρεθεί η συνεχόμενη σειρά τιμών της ακολουθίας που δίνει το μεγαλύτερο δυνατό άθροισμα. Στην παρούσα εργασία ζητείται η αλγοριθμική διερεύνηση του προβλήματος. Επιπλέον, ζητείται η υλοποίηση αλγορίθμων επίλυσης του προβλήματος με διαφορετική υπολογιστική πολυπλοκότητα, η καταγραφή και η παρουσίαση των χρόνων εκτέλεσης.

2 Περιγραφή προβλήματος

Εστω μια ακολουθία A με n ακέραιους αριθμούς. Στο πρόβλημα μέγιστης υποακολουθίας ζητείται ο εντοπισμός των δεικτών i και j με $1 \leq i \leq j \leq n$ για τους οποίους μεγιστοποιείται το άθροισμα $\sum_{x=i}^{j} A[x]$.

-2	1	-3	4	-1	2	1	-5	4
1	2	3	4	5	6	7	8	9

Πίνακας 1: Μια ακολουθία τιμών

Για την ακολουθία του πίνακα 1 η μέγιστη υποακολουθία είναι η 4, -1, 2, 1 με άθροισμα 6, που ξεκινά στη θέση 4 και τερματίζει στη θέση 6 της ακολουθίας (έναρξη μέτρησης από το 1).

2.1 Απλοϊκός τρόπος επίλυσης

Ενας απλοϊκός τρόπος υπολογισμού της μέγιστης υποακολουθίας είναι χρησιμοποιώντας 3 εμφωλευμένους βρόχους επανάληψης και έναν αθροιστή.

2.2 Βελτιωμένος αλγόριθμος μέγιστης υποακολουθίας

Ο προηγούμενος τρόπος “χάνει” χρόνο καθώς υπολογίζει ξανά όλα τα αθροίσματα υποακολουθίας. Ο αλγόριθμος μπορεί να βελτιωθεί υπολογίζοντας εκ των προτέρων όλα τα προθεματικά αθροίσματα της ακολουθίας, που είναι όλα τα αθροίσματα των t ακεραίων του A για $t = 1, 2, ..., n$. Διαθέτοντας τα προθεματικά αθροίσματα μπορούμε να υπολογίσουμε το άθροισμα οποιασδήποτε υποακολουθίας πραγματοποιώντας αφαίρεση ανάμεσα σε δύο ήδη υπολογισμένες τιμές προθεματικών αθροισμάτων.

Τα προθεματικά αθροίσματα για την ακολουθία τιμών του πίνακα 1 εμφανίζονται στον πίνακα 2.

0	-2	-1	-4	0	-1	1	2	-3	1
0	1	2	3	4	5	6	7	8	9

Πίνακας 2: Προθεματικά αθροίσματα

2.3 Αλγόριθμος του Kadane

Η βασική ιδέα του αλγορίθμου του Kadane είναι ότι υπολογίζει μέγιστα επιθεματικά αθροίσματα, δηλαδή δημιουργεί μια ακολουθία που κάθε στοιχείο της στη θέση t περιέχει το μέγιστο υποάθροισμα ανάμεσα σε όλες τις υποακολουθίες που τελειώνουν στο t. Σε περίπτωση που η τιμή που υπολογίζεται είναι αρνητική τότε αντικαθίσταται από το μηδέν.

Τα επιθεματικά μέγιστα αθροίσματα για την ακολουθία τιμών του πίνακα 1 εμφανίζονται στον πίνακα 3.

0	0	1	0	4	3	5	6	1	5
0	1	2	3	4	5	6	7	8	9

Πίνακας 3: Επιθεματικά μέγιστα αθροίσματα

3 Παραδοτέα εργασίας

Τα παραδοτέα της εργασίας θα είναι τα ακόλουθα:

1. Κώδικας που να υλοποιεί όλους τους αλγορίθμους επίλυσης του προβλήματος που αναφέρθηκαν. Κάθε αλγόριθμος να επιστρέφει το άθροισμα της μέγιστης υποακολουθίας, τη θέση έναρξής της και τη θέση τερματισμού της.

2. Unit Tests ελέγχου της ορθότητας των αλγορίθμων.

3. Οδηγίες εκτέλεσης του κώδικα.

4. Τεχνική αναφορά για την εργασία, στα πρότυπα σύντομου επιστημονικού άρθρου. Η αναφορά θα πρέπει να είναι το πολύ 2 σελίδες και να περιέχει:

   (α) Το υλικό που χρησιμοποιήθηκε για τα πειράματα (CPU, RAM, ...).

   (β) Το λογισμικό που χρησιμοποιήθηκε (π.χ. Python 3.8, GCC, VsCode, Excel, gnuplot, matplotlib, ...).

   (γ) Τη θεωρητική πολυπλοκότητα των αλγορίθμων.

   (δ) Τις μετρήσεις χρόνου εκτέλεσης του αλγορίθμου (π.χ. χρησιμοποιώντας τυχαία δεδομένα εισόδου 100, 1.000, 10.000, 50.000 ακεραίων τιμών σε εύρος τιμών από -100 έως και 100).

   (ε) Γραφική σύγκριση μέσω διαγραμμάτων των αποτελεσμάτων που παράγουν οι αλγόριθμοι.

4 Παρατηρήσεις

- Η υλοποίηση του κώδικα θα πρέπει να γίνει κατά προτίμηση στη γλώσσα προγραμματισμού Python. Εναλλακτικά μπορεί να χρησιμοποιηθεί C, C++ ή Java.

- Η εργασία είναι ατομική.