

# ΑΝΑΦΟΡΑ ΕΡΓΑΣΙΑΣ 2 - ΧΩΡΙΚΑ ΔΕΔΟΜΕΝΑ

Δημήτριος Παγώνης

1 Μαΐου 2025

## Μέρος 1 Κατασκευή R-Tree μέσω Bulk Loading

### Τρόπος σκέψης για την επίλυση της άσκησης

Αρχικά, υπολόγισα τον Morton Code (Z-order curve) χρησιμοποιώντας τη συνάρτηση `interleave_latlng` από τον σύνδεσμο που μας προτάθηκε. Στη συνέχεια, δημιούργησα τη συνάρτηση `create_sortedListOfMBRs`, η οποία διαβάζει τα δεδομένα από τα αρχεία εισόδου, υπολογίζει τα *Minimum Bounding Rectangles (MBRs)* για κάθε αντικείμενο και ταξινομεί τα δεδομένα σε μια λίστα από λεξικά με βάση τον *Morton Code*. Ακολούθως, πραγματοποιήσα τη δημιουργία των φύλλων του *R-tree* μέσω της συνάρτησης `build_leaf_nodes`, όπου τα δεδομένα χωρίστηκαν σε ομάδες των 20 και υπολογίστηκαν τα *MBRs* που περιγράφουν τα όρια σε κάθε κόμβο. Στη συνέχεια, η συνάρτηση `build_internal_nodes` αναδρομικά δημιούργησε τους εσωτερικούς κόμβους του *R-tree*, συνδυάζοντας τα *MBRs* των παιδιών τους μέχρι να φτάσει στη ρίζα. Η δομή του *R-tree* αποθηκεύτηκε στο αρχείο *Rtree.txt* μέσω της συνάρτησης `write_rtree_to_file`, η οποία πρώτα γράφει τα φύλλα, έπειτα τους εσωτερικούς κόμβους και τέλος τη ρίζα, διατηρώντας τη σωστή ιεραρχία. Τέλος, ο κώδικας εκτυπώνει τον αριθμό των κόμβων σε κάθε επίπεδο του *R-tree*, ξεκινώντας από τα φύλλα και φτάνοντας στη ρίζα, παρέχοντας μια συνολική εικόνα της δομής του δέντρου.

### Εντολή εκτέλεσης κώδικα

```
$ python3 R_tree_bulk.py offsets.txt coords.txt
```

## Μέρος 2 Ερωτήσεις Εύρους(Range queries)

### Τρόπος σκέψης για την επίλυση της άσκησης

Ακολουθώντας τον ίδιο ανάποδο τρόπο σκέψης με την πρώτη άσκηση, προκειμένου να πάρω τις πληροφορίες που χρειάζομαι από το *Rtree.txt* ώστε να σχηματίσω ένα δέντρο και να το διατρέξω μέσω των ερωτημάτων που θα τεθούν από την *Rqueries.txt*

Αρχικά, ανέλυσα τη δομή του *R-tree*, όπου τα φύλλα περιέχουν τα αντικείμενα με τα *MBRs*, ενώ οι εσωτερικοί κόμβοι περιέχουν τα *MBRs* των παιδιών τους. Στη συνέχεια, σχεδίασα τη διαδικασία φόρτωσης των δεδομένων από το αρχείο *Rtree.txt*, με διαχωρισμό των φύλλων, των εσωτερικών κόμβων και της ρίζας.

Έπειτα, προκειμένου να μπορέσω να διατρέξω καλύτερα και αποδοτικότερα, δημιούργησα κάποιες βοηθητικές δομές, έναν πίνακα όλων των κόμβων (*all\_nodes*) και έναν πίνακα που συνδέει τους γονείς με τα φύλλα τους (*leaves\_by\_parent*). Η αναδρομική συνάρτηση *range\_query* υλοποιήθηκε ώστε να ελέγχει αν τα *MBRs* τέμνονται με το παράθυρο αναζήτησης και να επιστρέφει τα *IDs* των αντικειμένων που ικανοποιούν το ερώτημα.

Τέλος, η συνάρτηση *process\_query\_file* διαβάζει τα ερωτήματα από το αρχείο *Rqueries.txt*, εκτελεί την αναζήτηση και εκτυπώνει τα αποτελέσματα στη μορφή που ζητήθηκε.

---

## Εντολή εκτέλεσης κώδικα

```
$ python3 range_queries.py Rtree.txt Rqueries.txt
```