

HY240: Δομές Δεδομένων

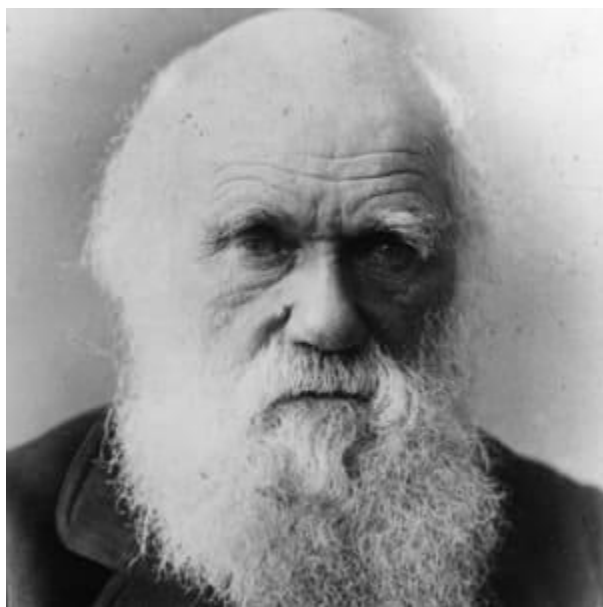
Χειμερινό Εξάμηνο - Ακαδημαϊκό Έτος 2019

Διδάσκουσα: Παναγιώτα Φατούρου

Προγραμματιστική Εργασία - 2ο Μέρος

Ημερομηνία Παράδοσης: Δευτέρα, 20 Δεκεμβρίου 2019, ώρα 23:59

Τρόπος παράδοσης: Χρησιμοποιώντας το πρόγραμμα turnin. Πληροφορίες για το πώς λειτουργεί το πρόγραμμα turnin παρέχονται στην ιστοσελίδα του μαθήματος.



Στη μακρά ιστορία του ανθρώπινου (και του ζωικού) γένους, αυτοί που έμαθαν να συνεργάζονται και να αυτοσχεδιάζουν αποτελεσματικά, έχουν επικρατήσει.

Κάρολος Δαρβίνος

Φωτογραφία:

<https://www.biography.com/scientist/charles-darwin>

Γενική Περιγραφή Εργασίας

Στην εργασία αυτή καλείστε να υλοποιήσετε ένα πρόγραμμα που να προσομοιώνει την εξέλιξη του ανθρώπινου είδους σύμφωνα με την θεωρία του Δαρβίνου και την μετέπειτα έρευνα άλλων επιστημόνων. Αξίζει να σημειωθεί, ότι για τους εκπαιδευτικούς σκοπούς του μαθήματος ορισμένα σημεία της εργασίας ενδέχεται να αποκλίνουν από την πραγματική έκβαση των γεγονότων και της εξέλιξης.

«Η επιστημονική μελέτη της ανθρώπινης εξέλιξης περιλαμβάνει τόσο την πορεία εξέλιξης του γένους Homo, όσο και τη μελέτη άλλων Ανθρωπίδων και Ανθρωπίνων, όπως τον Αυστραλοπίθηκο. Ως "σύγχρονος άνθρωπος" ορίζεται το είδος Homo sapiens, του οποίου το μόνο επιζών υποείδος είναι γνωστό ως Homo sapiens sapiens. Ο Homo sapiens idaltu, το άλλο υποείδος Homo Sapiens έχει εκλείψει. Ο Homo neanderthalensis (Νεάντερταλ), ο οποίος αφανίστηκε 30.000 χρόνια πριν, θεωρείται από μερικούς ως το τρίτο υποείδος Homo Sapiens και ονομάζεται "Homo sapiens neanderthalensis". Οι ανατομικά σύγχρονοι άνθρωποι εμφανίστηκαν στην Αφρική περίπου 195.000 χρόνια πριν, και διάφορες έρευνες υποστηρίζουν πως ο τελευταίος κοινός πρόγονος όλων των σύγχρονων ανθρώπων έζησε περίπου 200.000 χρόνια πριν.» (Wikipedia)

Αναλυτική Περιγραφή Ζητούμενης Υλοποίησης

Τα διάφορα είδη (species) του ανθρώπινου γένους (όπως π.χ., Homo Habilis, H. Georgicus, H. Ergaster, κλπ) περιγράφονται από ένα μοναδικό αναγνωριστικό κι αποθηκεύονται σε ένα δυαδικό μετα-διατεταγμένα ταξινομημένο δένδρο το οποίο ονομάζεται **δένδρο ειδών**. Ένα μετα-διατεταγμένα ταξινομημένο δένδρο είναι ένα δυαδικό δένδρο για το οποίο η **μετα-διατεταγμένη** διάσχιση δίνει τα κλειδιά που είναι αποθηκευμένα στους κόμβους του δένδρου σε **αύξουσα** διάταξη. Το δένδρο ειδών είναι **μετα-διατεταγμένα ταξινομημένο ως προς το αναγνωριστικό του είδους**.

Επιπλέον, κάθε είδος αποτελείται από διάφορους πληθυσμούς (populations) που ανήκουν στο είδος (ή που προϋπήρχαν και ενδεχομένως ανήκαν σε άλλα είδη, αλλά γενετικό τους υλικό έχει ανιχνευθεί στο γονιδίωμα του είδους). Κάθε τέτοιος πληθυσμός περιγράφεται από ένα μοναδικό αναγνωριστικό και θεωρούμε ότι κατοικούσε σε κάποια ήπειρο¹. Οι πληθυσμοί κάθε είδους περιγράφονται με ένα **διπλά** συνδεδεμένο δυαδικό δένδρο αναζήτησης (binary search tree), **ταξινομημένο (σύμφωνα με την ενδοδιατεταγμένη διάταξη) ως προς το αναγνωριστικό του πληθυσμού**, το οποίο ονομάζεται **δένδρο πληθυσμών**.

Κάθε κόμβος του δένδρου ειδών αποτελεί μια εγγραφή τύπου *struct Species* με τα ακόλουθα στοιχεία:

- **sid**: Αναγνωριστικό (τύπου *int*) που χαρακτηρίζει μοναδικά το είδος.
- **population_root**: Δείκτης (τύπου *struct Population*) στη ρίζα του δένδρου πληθυσμών του είδους.
- **lc**: Δείκτης (τύπου *struct Species*) στο αριστερό παιδί του κόμβου.
- **rc**: Δείκτης (τύπου *struct Species*) στο δεξιό παιδί του κόμβου.

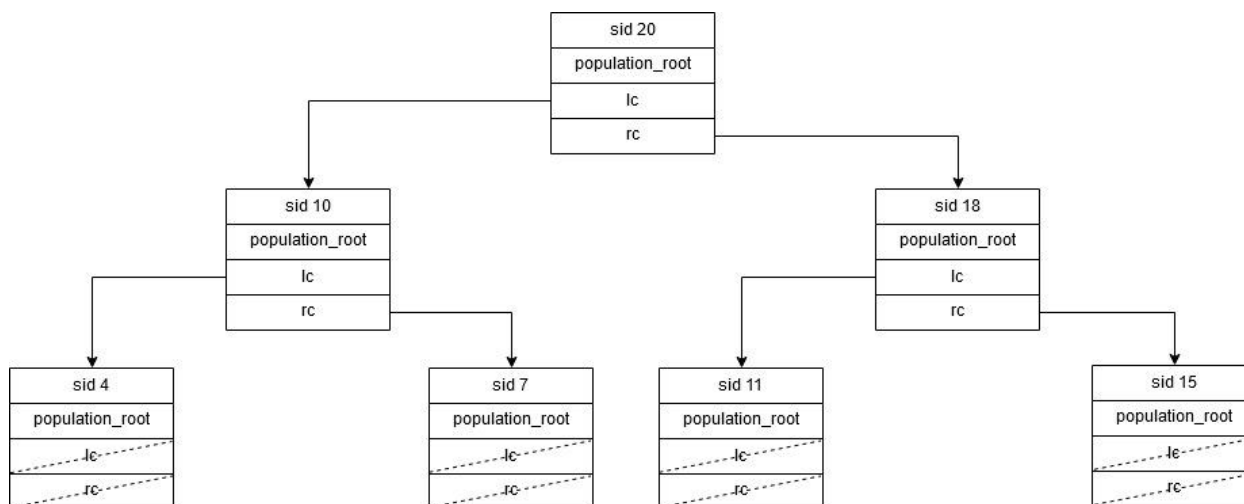
Ο κάθε κόμβος του δένδρου πληθυσμών του κάθε είδους είναι μία εγγραφή τύπου *struct Population* με τα ακόλουθα στοιχεία:

- **gid**: Αναγνωριστικό (τύπου *int*) που χαρακτηρίζει μοναδικά τον πληθυσμό.
- **sid**: Αναγνωριστικό (τύπου *int*) που χαρακτηρίζει μοναδικά το είδος στο οποίο ανήκει ο πληθυσμός².
- **cid**: Αναγνωριστικό (τύπου *int*) που χαρακτηρίζει μοναδικά την ήπειρο στην οποία κατοικεί ο συγκεκριμένος πληθυσμός.
- **parent**: Δείκτης (τύπου *struct Population*) στον γονέα του κόμβου.
- **lc**: Δείκτης (τύπου *struct Population*) στο αριστερό παιδί του κόμβου.
- **rc**: Δείκτης (τύπου *struct Population*) στο δεξιό παιδί του κόμβου.

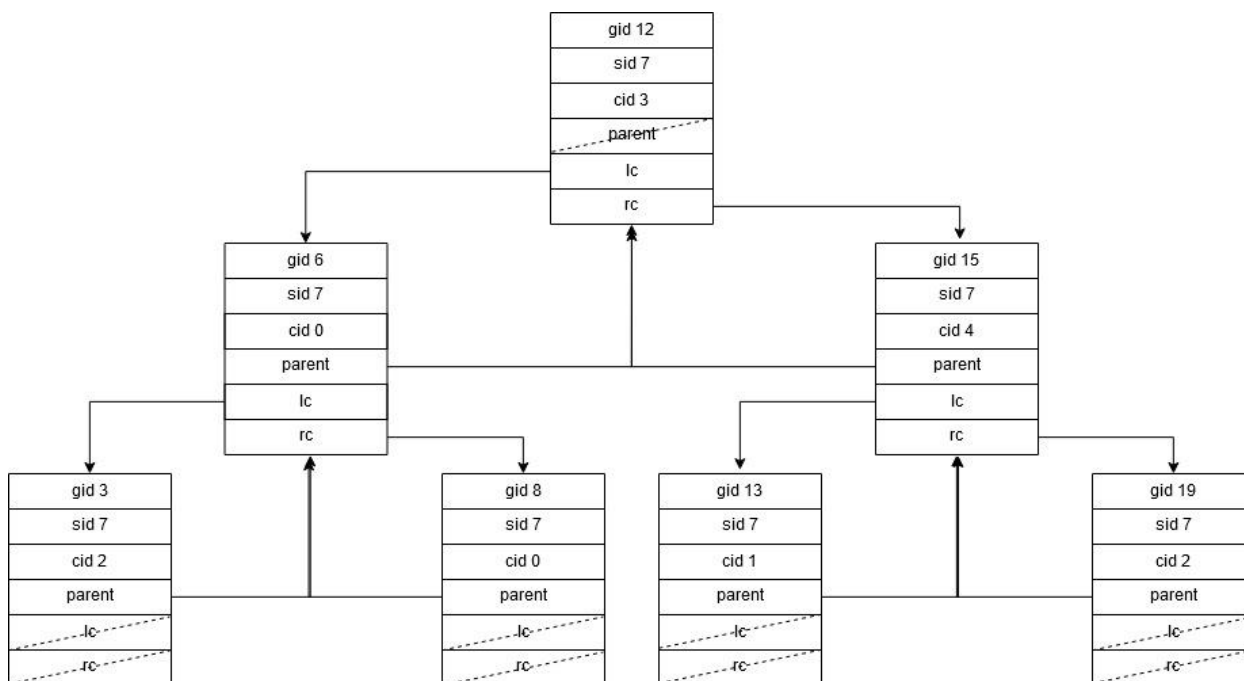
Στο Σχήμα 1, παρουσιάζεται το δένδρο ειδών, ενώ στο Σχήμα 2 το δένδρο πληθυσμών του κάθε είδους.

¹ Παρότι για κάποια είδη, πληθυσμοί βρέθηκαν μόνο στην Αφρικανική ήπειρο, για τις ανάγκες του μαθήματος, θα θεωρήσουμε ότι για όλα τα είδη, υπήρξαν πληθυσμοί σε διάφορες ηπείρους.

² Το πεδίο αυτό θα έχει την ίδια τιμή για όλα τα στοιχεία του δένδρου πληθυσμών ενός είδους. Ωστόσο, το *struct Population* (και το πεδίο *sid*) είναι χρήσιμο και για δομές που περιγράφονται στη συνέχεια.



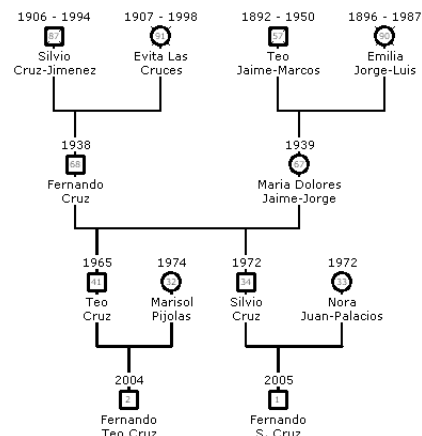
Σχήμα 1: Το μετα-διατεταγμένα ταξινομημένο δένδρο ειδών, ταξινομημένο ως προς το αναγνωριστικό των ειδών.



Σχήμα 2: Το διπλά συνδεδεμένο δυαδικό δένδρο πληθυσμών, ταξινομημένο ως προς το αναγνωριστικό των πληθυσμών. Στο παράδειγμα βλέπουμε το δένδρο πληθυσμών του είδους με αναγνωριστικό 7.

«Το γενεαλογικό ή οικογενειακό δένδρο αποτελεί περιγραφική ή σχηματική αναφορά επί του συνόλου των γνωστών προγόνων ή απογόνων κάποιου ανθρώπου, καθώς και το διάγραμμα που τους εμφανίζει σε δενδρική δομή μέχρι κάποιο όριο, είτε στο παρόν, είτε σε ιστορικό παρελθόν.

Ο σκοπός του γενεαλογικού δέντρου είναι να αναδείξει συσχετισμούς και βαθμούς οικογενειακών ή συγγένειας ανθρώπων.» (Wikipedia)



Παράδειγμα δενδρικής διάταξης
«γενεαλογικού δένδρου προγόνων»
(Wikipedia)

Όπως προαναφέρθηκε, κάθε πληθυσμός ενός είδους κατοικεί σε μία συγκεκριμένη ήπειρο η οποία περιγράφεται από το αναγνωριστικό **cid** που περιέχεται στη δομή του πληθυσμού. Για την προσομοίωση της γεωγραφικής κατανομής των πληθυσμών, θα χρησιμοποιήσετε έναν στατικό πίνακα πέντε θέσεων, που ονομάζεται πίνακας ηπείρων, όπου κάθε στοιχείο του αποθηκεύει ένα δείκτη, ο οποίος δείχνει στη ρίζα ενός ταξινομημένου, **απλά** συνδεδεμένου δυαδικού δένδρου αναζήτησης (binary search tree) **με κόμβο φρουρό**.

Ο κάθε κόμβος του δένδρου πληθυσμών του πίνακα ηπείρων είναι μία εγγραφή τύπου *struct ContinentPopulation* με τα ακόλουθα στοιχεία:

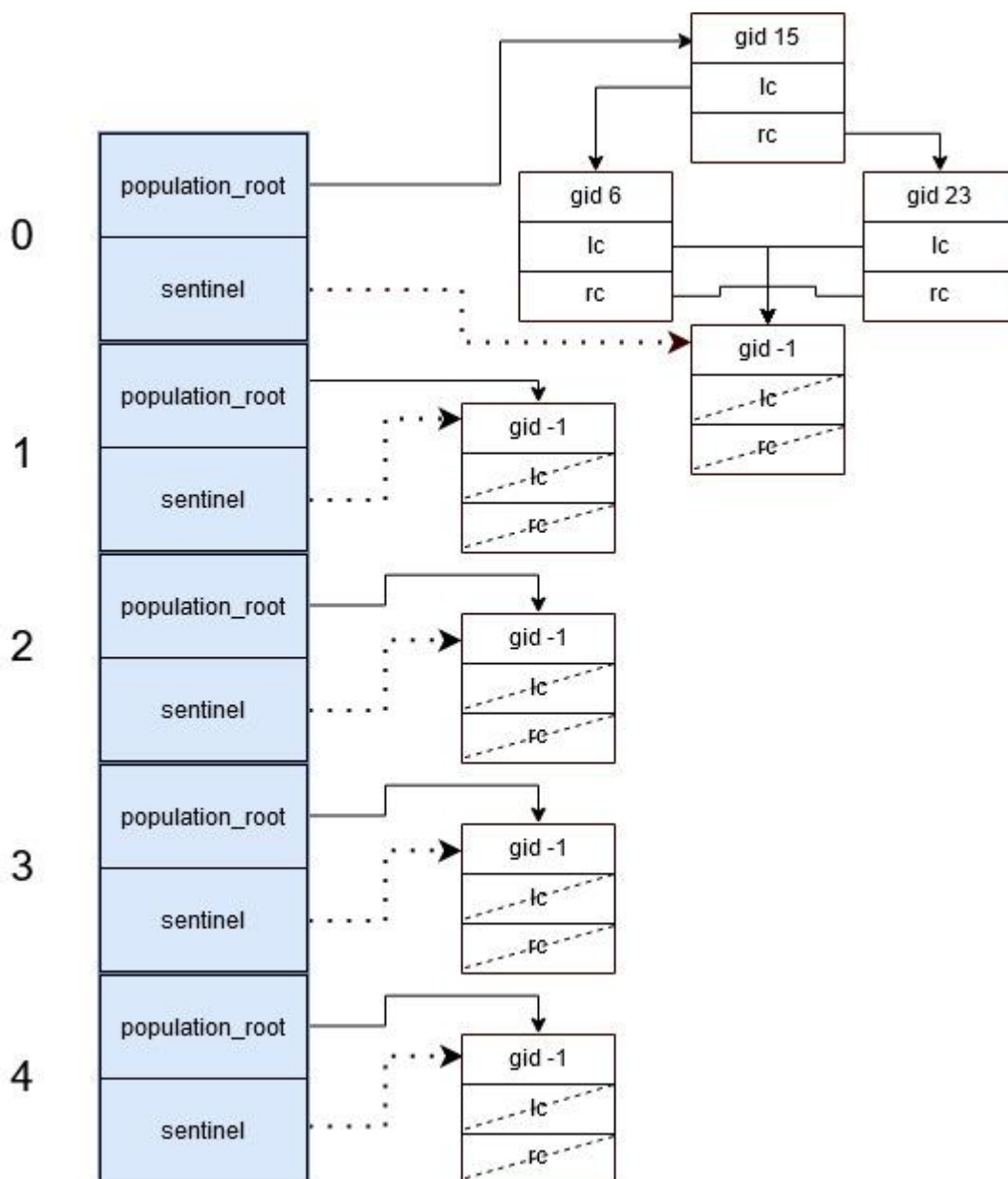
- **gid**: Αναγνωριστικό (τύπου *int*) που χαρακτηρίζει μοναδικά τον πληθυσμό.
- **lc**: Δείκτης (τύπου *struct ContinentPopulation*) στο αριστερό παιδί του κόμβου.
- **rc**: Δείκτης (τύπου *struct ContinentPopulation*) στο δεξί παιδί του κόμβου.

Ακόμη, κάθε στοιχείο του πίνακα θα είναι τύπου *struct ContinentTree* με τα ακόλουθα στοιχεία:

- **population_root**: Δείκτης (τύπου *struct ContinentPopulation*) στη ρίζα του δένδρου πληθυσμών της ηπείρου.
- **sentinel**: Δείκτης (τύπου *struct ContinentPopulation*) στον κόμβο φρουρό του δένδρου.

Ο κόμβος φρουρός του κάθε δένδρου θα έχει την τιμή -1 στο αναγνωριστικό *gid* και την τιμή *NULL* στους δείκτες *lc* και *rc*.

Στο Σχήμα 3, παρουσιάζεται ο πίνακας ηπείρων, καθώς και τα δένδρα πληθυσμών της κάθε ηπείρου.



Σχήμα 3: Ο πίνακας ηπείρων και τα δένδρα πληθυσμών της κάθε ηπείρου. Στο παράδειγμα μόνο η πρώτη ήπειρος περιέχει πληθυσμούς. Τα δένδρα πληθυσμών είναι ταξινομημένα με κόμβο φρουρό.

Με το πέρασμα των αιώνων, η διαδικασία της εξέλιξης οδηγεί στο τελευταίο είδος που επικράτησε, το **Homo Sapiens**. Για εκπαιδευτικούς λόγους θεωρήστε ότι το γονιδίωμα του Homo Sapiens περιέχει γενετικό υλικό από όλα τα εναπομείναντα είδη.

Ένα εμπλουτισμένο φυλλοπροσανατολισμένο δένδρο δυαδικής αναζήτησης, ταξινομημένο (σύμφωνα με την ενδοδιατεταγμένη) ως προς το αναγνωριστικό των ειδών, που θα ονομάζεται **δένδρο homo Sapiens**, θα περιέχει όλα τα είδη των οποίων γενετικό υλικό έχει ανιχνευθεί στο γονιδίωμα του homo Sapiens.

Τα φυλλο-προσανατολισμένα δένδρα δυαδικής αναζήτησης (leaf-oriented binary search trees) αποτελούν μια εναλλακτική υλοποίηση του αφηρημένου τύπου δεδομένων του λεξικού. Ορίζονται ως εξής:

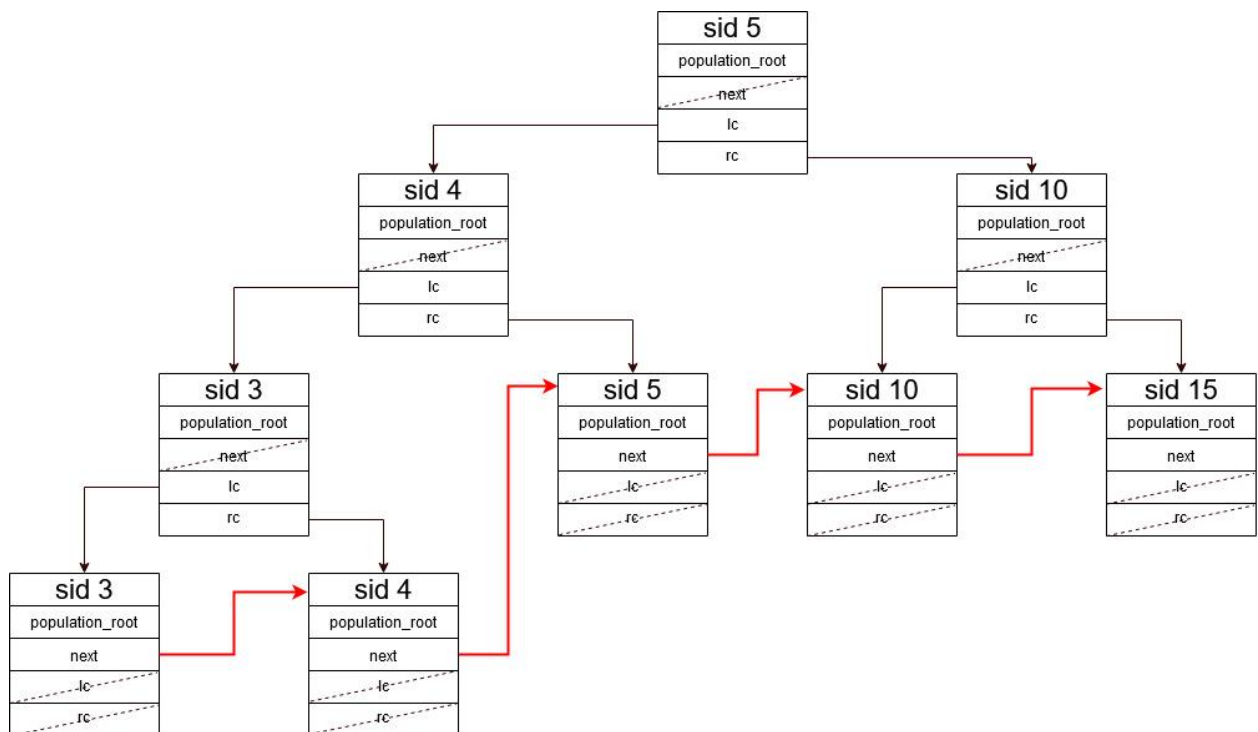
- Όλα τα κλειδιά του λεξικού αποθηκεύονται στα φύλλα του δένδρου, από αριστερά προς τα δεξιά κατά μη φθίνουσα τιμή κλειδιού.
- Οι εσωτερικοί κόμβοι αποθηκεύουν κλειδιά (που δεν αντιστοιχούν απαραίτητα σε κλειδιά του λεξικού), έτσι ώστε να ισχύει η κάτωθι αμετάβλητη συνθήκη σε κάθε κόμβο v :

Το κλειδί του αριστερού παιδιού του v είναι μικρότερο από αυτό του v , ενώ το δεξιό παιδί του v διαθέτει κλειδί μεγαλύτερο ή ίσο από εκείνο του v

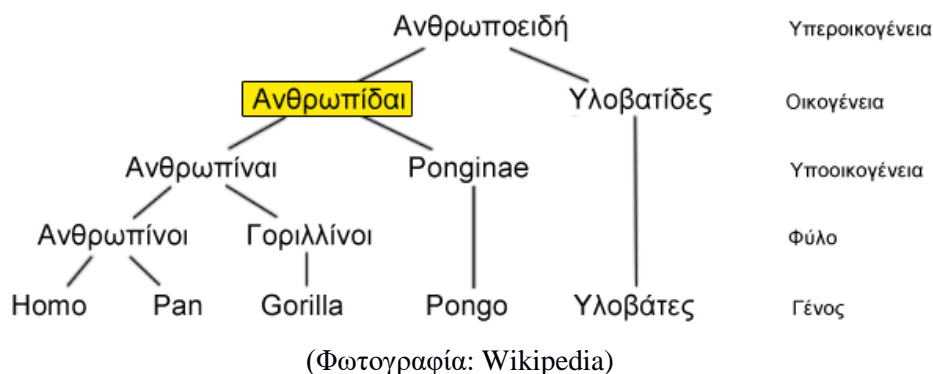
Ένα εμπλουτισμένο φυλλοπροσανατολισμένο δένδρο είναι ένα δένδρο του οποίου όλα τα φύλλα είναι συνδεδεμένα ώστε να σχηματίζουν μια ταξινομημένη απλά συνδεδεμένη λίστα. Επομένως, υπάρχει ένας επιπρόσθετος δείκτης *next* σε κάθε κόμβο. Ο δείκτης αυτός είναι NULL αν ο κόμβος είναι εσωτερικός, ενώ δείχνει στο φύλλο με το αμέσως μεγαλύτερο κλειδί διαφορετικά.

Ο κάθε κόμβος του δένδρου *homo Sapiens* είναι μία εγγραφή τύπου *struct HomoSapiens* με τα ακόλουθα στοιχεία:

- **sid**: Αναγνωριστικό (τύπου *int*) που χαρακτηρίζει μοναδικά το είδος.
- **population_root**: Δείκτης (τύπου *struct Population*) στη ρίζα του δένδρου πληθυσμών του είδους.
- **next**: Δείκτης (τύπου *struct HomoSapiens*) στο επόμενο φύλλο του δένδρου.
- **lc**: Δείκτης (τύπου *struct HomoSapiens*) στο αριστερό παιδί του κόμβου.
- **rc**: Δείκτης (τύπου *struct HomoSapiens*) στο δεξί παιδί του κόμβου.



Σχήμα 4: Το εμπλουτισμένο φυλλοπροσανατολισμένο δένδρο *Homo Sapiens*. Το πεδίο *next* του κάθε φύλλου δείχνει στο φύλλο με το αμέσως μεγαλύτερο κλειδί.



Το φυλογενετικό δέντρο των σωζόμενων Ανθρωποειδών: άνθρωπος (γένος Homo), χιμπαντζήδες και μπονόμπο (γένος Pan), γορίλες (γένος Gorilla), ουρακοτάγκοι (γένος Pongo) και γίββωνες (τέσσερα γένη της οικογένειας Υλοβατίδες: Υλοβάτες, Hoolock, Nomascus και Symphalangus). (Wikipedia)

Τρόπος Λειτουργίας Προγράμματος

Το πρόγραμμα που θα δημιουργηθεί θα πρέπει να εκτελείται καλώντας την ακόλουθη εντολή:

<executable> <input_file>

Όπου <executable> είναι το όνομα του εκτελέσιμου αρχείου του προγράμματος (π.χ., a.out) και <input_file> είναι το όνομα ενός αρχείου εισόδου (π.χ., testfile) το οποίο περιέχει γεγονότα των ακόλουθων μορφών:

S <sid>

Γεγονός τύπου insert species, που υποδηλώνει την εισαγωγή ενός νέου είδους με αναγνωριστικό <sid>. Κατά το γεγονός αυτό, θα γίνεται εισαγωγή ενός νέου κόμβου τύπου *struct Species* στο δένδρο ειδών. Μετά από κάθε εισαγωγή, το δένδρο θα πρέπει να εξακολουθεί να είναι μετα-διατεταγμένα ταξινομημένο. Μετά το πέρας της εκτέλεσης ενός τέτοιου γεγονότος, το πρόγραμμα θα πρέπει να τυπώνει την ακόλουθη πληροφορία:

```
S <sid>
  <sid1> ... <sidN>
DONE
```

όπου N είναι ο αριθμός των κόμβων στο δένδρο ειδών και για κάθε $i \in \{1, \dots, N\}$, <sid_i> είναι το αναγνωριστικό του είδους που αντιστοιχεί στον i-οστό κόμβο του δένδρου αυτού, σύμφωνα με την μετα-διατεταγμένη διάταξη. Δηλαδή, για κάθε $i, j \in \{1, \dots, N\}$, $\text{sid}_i < \text{sid}_j$

G <gid> <sid> <cid>

Γεγονός τύπου *insert population*, που υποδηλώνει την εισαγωγή ενός νέου πληθυσμού, με αναγνωριστικό <gid>, που ανήκει στο είδος με αναγνωριστικό <sid> και βρέθηκε στην ήπειρο <cid>. Κατά το γεγονός αυτό, θα πρέπει να εντοπίζεται το αναγνωριστικό <sid> στο δένδρο ειδών σε χρόνο $O(h)$, όπου h το ύψος του δένδρου ειδών και να πραγματοποιείται εισαγωγή ενός νέου κόμβου τύπου *struct Population* στο δένδρο πληθυσμών του είδους. Μετά από κάθε εισαγωγή, το δένδρο πληθυσμών θα πρέπει να παραμένει ταξινομημένο (binary search tree). Μετά το πέρας της εκτέλεσης ενός τέτοιου γεγονότος, το πρόγραμμα θα πρέπει να τυπώνει την ακόλουθη πληροφορία:

```
G <gid> <sid> <cid>
  <sid>
    <gid1, cid1>, <gid2, cid2> .... <gidM, cidM>
DONE
```

όπου M είναι το πλήθος των πληθυσμών που ανήκουν στο είδος με αναγνωριστικό <sid> και για κάθε $i \in \{1, \dots, M\}$, <gid _{i} > είναι το αναγνωριστικό του πληθυσμού που αντιστοιχεί στον i -οστό κόμβο του δένδρου πληθυσμών του είδους (σύμφωνα με την ενδο-διατεταγμένη διάταξη), ενώ <cid _{i} > είναι το αναγνωριστικό της ηπείρου όπου κατοικεί ο συγκεκριμένος πληθυσμός.

L <sid> <gid₁> <gid₂>

Γεγονός τύπου *lowest common ancestor*, που υποδηλώνει την εύρεση του τελευταίου (νεότερου) πληθυσμού που αποτελεί κοινό πρόγονο των πληθυσμών με αναγνωριστικά <gid₁> και <gid₂>, του είδους με αναγνωριστικό <sid>. Κατά το γεγονός αυτό, αρχικά θα πρέπει να εντοπίζεται το είδος με αναγνωριστικό <sid> στο δένδρο ειδών σε χρόνο $O(h)$, όπου h το ύψος του δένδρου ειδών. Στη συνέχεια, θα πρέπει να εντοπίζονται οι πληθυσμοί με αναγνωριστικά <gid₁> και <gid₂>. Στη συνέχεια θα πρέπει να εκτελείται ο αλγόριθμος υπολογισμού του κοινού προγόνου τους στο δένδρο που βρίσκεται στο μεγαλύτερο βάθος.

Η χρονική πολυπλοκότητα εκτέλεσης του γεγονότος αυτού πρέπει να είναι $O(h_1 + h_2)$, όπου h_1 είναι το ύψος του δένδρου ειδών και h_2 είναι το ύψος του δένδρου πληθυσμών αυτού του είδους.

Μετά το πέρας της εκτέλεσης ενός τέτοιου γεγονότος, το πρόγραμμα θα πρέπει να τυπώνει την ακόλουθη πληροφορία:

```
L <sid> <gid1> <gid2>
  Lowest Common Ancestor: <gid>
DONE
```

Όπου <gid> είναι το αναγνωριστικό του πληθυσμού που αποτελεί τον τελευταίο κοινό πρόγονο των πληθυσμών με αναγνωριστικά <gid₁> και <gid₂>.

M <sid₁> <sid₂> <sid₃>

Γεγονός τύπου *merge species*, που υποδηλώνει τη δημιουργία ενός νέου είδους με αναγνωριστικό <sid₃> στο οποίο έχει ανιχνευθεί γενετικό υλικό των ειδών με αναγνωριστικά <sid₁> και <sid₂>. Κατά το γεγονός αυτό, αρχικά θα πρέπει να εισάγεται ένα νέο είδος στο δένδρο ειδών. Στη συνέχεια, θα πρέπει να διατρέχεται το δένδρο ειδών και να εντοπίζονται τα είδη με αναγνωριστικά <sid₁> και <sid₂>, αντίστοιχα.

Έπειτα, το δένδρο πληθυσμών του είδους με αναγνωριστικό $\langle \text{sid}_1 \rangle$ θα πρέπει να συνενώνεται με το δένδρο πληθυσμών του είδους $\langle \text{sid}_2 \rangle$. Το νέο δένδρο πληθυσμών που θα προκύπτει θα αποτελεί το δένδρο πληθυσμών του είδους με αναγνωριστικό $\langle \text{sid}_3 \rangle$ και θα πρέπει να παραμένει ταξινομημένο ως προς το αναγνωριστικό των πληθυσμών (βάσει της ενδοδιατεταγμένης διάσχισης). Επιπρόσθετα, η χρονική πολυπλοκότητα της διαδικασίας συνένωσης θα πρέπει να είναι $O(n+m)$, όπου n είναι το μέγεθος του δένδρου πληθυσμών του είδους $\langle \text{sid}_1 \rangle$ και m το μέγεθος του δένδρου πληθυσμών του είδους $\langle \text{sid}_2 \rangle$, αντίστοιχα. Τέλος, τα είδη με αναγνωριστικά $\langle \text{sid}_1 \rangle$ και $\langle \text{sid}_2 \rangle$ θα πρέπει να διαγράφονται από το δένδρο ειδών. Μετά το πέρας της εκτέλεσης ενός τέτοιου γεγονότος, το πρόγραμμα θα πρέπει να τυπώνει την ακόλουθη πληροφορία:

```
M <sid1> <sid2> <sid3>
  <sid1>
    <gid1,1>, <gid1,2> .... <gid1,M1>
  ...
  <sidN>
    <gidN,1>, <gidN,2> .... <gidN,MN>
DONE
```

όπου N είναι ο αριθμός των κόμβων στο δένδρο ειδών και για κάθε $i \in \{1, \dots, N\}$, $\langle \text{sid}_i \rangle$ είναι το αναγνωριστικό του είδους που αντιστοιχεί στον i -οστό κόμβο του δένδρου αυτού, σύμφωνα με την μετα-διατεταγμένη διάσχιση. Ακόμη, για κάθε $i \in \{1, \dots, N\}$, το M_i είναι το πλήθος των πληθυσμών που ανήκουν στο είδος με αναγνωριστικό $\langle \text{sid}_i \rangle$ και για κάθε $j \in \{1, \dots, M_i\}$, $\langle \text{gid}_{i,j} \rangle$ είναι το αναγνωριστικό του πληθυσμού που αντιστοιχεί στον j -οστό κόμβο του δένδρου πληθυσμών του είδους (σύμφωνα με την ενδο-διατεταγμένη διάταξη).

D

Γεγονός τύπου *distribute species*, το οποίο υποδηλώνει τη γεωγραφική κατανομή των διαφόρων πληθυσμών στις πέντε ηπείρους. Κατά το γεγονός αυτό, θα διατρέχεται τόσο το δένδρο ειδών όσο και το δένδρο πληθυσμών κάθε είδους με **προ-διατεταγμένη** διάσχιση και για κάθε πληθυσμό θα εισάγεται ένας **νέος κόμβος** τύπου *struct ContinentPopulation* στο δένδρο της κατάλληλης θέσης του πίνακα ηπείρων, χρησιμοποιώντας το πεδίο *cid* του πληθυσμού. Η χρονική πολυπλοκότητα εκτέλεσης του γεγονότος αυτού πρέπει να είναι $O(n)$, όπου n είναι το συνολικό πλήθος στοιχείων στα δένδρα πληθυσμών όλων των ειδών. **Προσοχή:** Είναι σημαντικό να διασχίσετε με προ-διατεταγμένη διάσχιση τα δένδρα πληθυσμών, διαφορετικά τα δένδρα που θα σχηματιστούν στις ηπείρους θα έχουν εκφυλιστεί σε λίστες λόγω της ταξινόμησης.

Μετά το πέρας της εκτέλεσης ενός τέτοιου γεγονότος, το πρόγραμμα θα πρέπει να τυπώνει την ακόλουθη πληροφορία:

```
D
Continent 0 : <gid0,1>, <gid0,2> .... <gid0,M0>
Continent 1 : <gid1,1>, <gid1,2> .... <gid1,M1>
...
Continent 4 : <gid4,1>, <gid4,2> .... <gid4,M4>
DONE
```

όπου $0, \dots, 4$ είναι τα αναγνωριστικά των 5 ηπείρων και για κάθε $i \in \{0, \dots, 4\}$, το M_i είναι το πλήθος των πληθυσμών που κατοικούν στην ήπειρο που αντιστοιχεί στη i -οστή θέση του πίνακα ηπείρων και για κάθε $j \in \{1, \dots, M_i\}$, $\langle \text{gid}_{i,j} \rangle$ είναι το αναγνωριστικό του πληθυσμού που αντιστοιχεί στον j -οστό κόμβο του δένδρου πληθυσμών της ηπείρου.

K $\langle \text{gid} \rangle \langle \text{sid} \rangle$

Γεγονός τύπου *delete population*, το οποίο δηλώνει τη μη επιβίωση του πληθυσμού με αναγνωριστικό $\langle \text{gid} \rangle$ ο οποίος ανήκει στο είδος με αναγνωριστικό $\langle \text{sid} \rangle$. Κατά το γεγονός αυτό, θα πρέπει να εντοπίζεται το είδος με αναγνωριστικό $\langle \text{sid} \rangle$ στο δένδρο ειδών σε χρόνο $O(h)$, όπου h το ύψος του δένδρου ειδών και στη συνέχεια να διαγράφεται ο πληθυσμός με αναγνωριστικό $\langle \text{gid} \rangle$ από το δένδρο πληθυσμών του είδους. Ακόμη, ο πληθυσμός θα πρέπει να διαγράφεται κι από το κατάλληλο δένδρο πληθυσμών στον πίνακα ηπείρων. Μετά το πέρας της εκτέλεσης ενός τέτοιου γεγονότος, το πρόγραμμα θα πρέπει να τυπώνει την ακόλουθη πληροφορία:

```
K <gid> <sid>
SPECIES
  <sid>
    <gids,1s,2s,Mscid,1cid,2cid,Mc

```

όπου για την εκτύπωση του δένδρου ειδών, $\langle \text{sid} \rangle$ είναι το αναγνωριστικό του είδους όπου ανήκει ο πληθυσμός με αναγνωριστικό $\langle \text{gid} \rangle$ και S είναι η θέση του είδους στο δένδρο ειδών. Ακόμη, το M_s είναι το πλήθος των πληθυσμών που ανήκουν στο είδος με αναγνωριστικό $\langle \text{sid} \rangle$ και για κάθε $j \in \{1, \dots, M_s\}$, $\langle \text{gid}_{i,j} \rangle$ είναι το αναγνωριστικό του πληθυσμού που αντιστοιχεί στον j -οστό κόμβο του δένδρου πληθυσμών του είδους. Για την εκτύπωση του πίνακα ηπείρων, $\langle \text{cid} \rangle$ είναι το αναγνωριστικό (και η θέση στον πίνακα ηπείρων) της ηπείρου στην οποία κατοικεί ο πληθυσμός, M_c είναι το πλήθος των πληθυσμών που κατοικούν στην ήπειρο και για κάθε $j \in \{1, \dots, M_c\}$, $\langle \text{gid}_{cid,j} \rangle$ είναι το αναγνωριστικό του πληθυσμού που αντιστοιχεί στον j -οστό κόμβο του δένδρου πληθυσμών της ηπείρου.

F

Γεγονός τύπου *species extinction*, το οποίο υποδηλώνει την εξαφάνιση του παλαιότερου είδους. Κατά το γεγονός αυτό, θα πρέπει να εντοπίζεται το είδος με το μικρότερο $\langle \text{sid} \rangle$ στο δένδρο ειδών σε χρόνο $O(h)$, όπου h το ύψος του δένδρου ειδών, στη συνέχεια να διαγράφονται όλοι οι κόμβοι από το δένδρο πληθυσμών του είδους αυτού και τέλος να διαγράφεται και το struct που αντιστοιχεί στο είδος αυτό (με αναγνωριστικό $\langle \text{sid} \rangle$) από το δένδρο πληθυσμών. Μετά τη διαγραφή το δένδρο των ειδών θα πρέπει να παραμένει ταξινομημένο (ως προς τη μεταδιατεταγμένη διάταξη). Επιπλέον, κάθε πληθυσμός που διαγράφεται από την λίστα πληθυσμών του είδους, θα πρέπει να διαγράφεται κι από τον πίνακα ηπείρων χρησιμοποιώντας το πεδίο cid του πληθυσμού. Μετά το πέρας της εκτέλεσης ενός τέτοιου γεγονότος, το πρόγραμμα θα πρέπει να τυπώνει την ακόλουθη πληροφορία:

```

F
SPECIES
  <sid1>
    <gid1,1>,<gid1,2> .... <gid1,M1>
  ...
  <sidN>
    <gidN,1>,<gidN,2> .... <gidN,MN>
CONTINENTS
  Continent 0 : <gid0,1>,<gid0,2> .... <gid0,M1>
  Continent 1 : <gid1,1>,<gid1,2> .... <gid1,M2>
  ...
  Continent 4 : <gid4,1>,<gid4,2> .... <gid4,M4>
DONE

```

όπου για την εκτύπωση της λίστας ειδών, N είναι ο αριθμός των κόμβων στο δένδρο ειδών και για κάθε $i \in \{1, \dots, N\}$, $\langle \text{sid}_i \rangle$ είναι το αναγνωριστικό του είδους που αντιστοιχεί στον i -οστό κόμβο του δένδρου αυτού, σύμφωνα με τη μετα-διατεταγμένη διάσχιση. Ακόμη, για κάθε $i \in \{1, \dots, N\}$, το M_i είναι το πλήθος των πληθυσμών που ανήκουν στο είδος με αναγνωριστικό $\langle \text{sid}_i \rangle$ και για κάθε $j \in \{1, \dots, M_i\}$, $\langle \text{gid}_{i,j} \rangle$ είναι το αναγνωριστικό του πληθυσμού που αντιστοιχεί στον j -οστό κόμβο του δένδρου πληθυσμών του είδους. Για την εκτύπωση του πίνακα ηπείρων, για κάθε $i \in \{0, \dots, 4\}$, το M_i είναι το πλήθος των πληθυσμών που κατοικούν στην ήπειρο που αντιστοιχεί στην i -οστή θέση του πίνακα ηπείρων και για κάθε $j \in \{1, \dots, M_i\}$, $\langle \text{gid}_{i,j} \rangle$ είναι το αναγνωριστικό του πληθυσμού που αντιστοιχεί στον j -οστό κόμβο του δένδρου πληθυσμών της ηπείρου.

E

Γεγονός τύπου evolution, που υποδηλώνει τη δημιουργία του εμπλουτισμένου φυλλοπροσανατολισμένου δένδρου homo Sapiens. Για τους εκπαιδευτικούς σκοπούς του μαθήματος, θεωρούμε ότι το γονιδίωμα του Homo Sapiens περιέχει γενετικό υλικό από όλα τα εναπομείναντα είδη. Κατά το γεγονός αυτό, θα πρέπει να διατρέχεται το δένδρο ειδών με **προ-διατεταγμένη διάσχιση**, κάθε είδος να εισάγεται στο δένδρο ειδών homo Sapiens και για κάθε είδος, ο δείκτης population_root της εγγραφής του είδους θα πρέπει να αρχικοποιηθεί κατάλληλα ώστε να δείχνει στη ρίζα του δένδρου πληθυσμών του είδους.

Μετά το πέρας του γεγονότος αυτού, το δένδρο ειδών πρέπει να είναι κενό.

Προσοχή: Είναι σημαντικό να διασχίσετε με προ-διατεταγμένη διάσχιση το δένδρο ειδών, διαφορετικά το δένδρο homo Sapiens που θα σχηματιστεί θα έχει εκφυλιστεί σε λίστα λόγω της ταξινόμησης.

Μετά το πέρας της εκτέλεσης ενός τέτοιου γεγονότος, το πρόγραμμα θα πρέπει να τυπώνει την ακόλουθη πληροφορία:

```

E
Homo Sapiens: [<gid1,1, sid1>, ...<gid1,M1, sidM1>] .... [<gidN,1, sid1>,...<gidN,MN, sidMN>]
DONE

```

όπου N είναι ο αριθμός των ειδών, για κάθε $i \in \{1, \dots, N\}$, M_i είναι το πλήθος των πληθυσμών που ανήκουν στο είδος που το δένδρο του ξεκινά στο i -οστό φύλλο του δένδρου homo Sapiens, ενώ για

κάθε $j \in \{1, \dots, M_i\}$, $\langle \text{gid}_{i,j} \rangle$ είναι το αναγνωριστικό του πληθυσμού που αντιστοιχεί στον j -οστό κόμβο του δένδρου πληθυσμών του είδους αυτού.

N

Γεγονός τύπου *species statistics*, που υποδηλώνει την εξαγωγή στατιστικών για όλα τα είδη που συνέβαλαν στη δημιουργία της αλυσίδας του δένδρου *homo Sapiens*. Συγκεκριμένα, κατά το γεγονός αυτό θα πρέπει να διατρέχεται η αλυσίδα του δένδρου *homo Sapiens* και να καταγράφεται το πόσα διαφορετικά είδη περιέχονται σ' αυτή. Το γεγονός αυτό θα πρέπει να εκτελείται με χρονική πολυπλοκότητα $O(n)$, όπου n είναι το πλήθος των διαφορετικών ειδών που περιέχονται στην αλυσίδα των φύλλων του δένδρου *homo Sapiens*. (Για να το πετύχετε αυτό, θα πρέπει να χρησιμοποιήσετε κατάλληλα τους δείκτες *next_species* και **να μην διατρέξετε** ολόκληρο το δένδρο *homo Sapiens*.) Μετά το πέρας της εκτέλεσης ενός τέτοιου γεγονότος, το πρόγραμμα θα πρέπει να τυπώνει την ακόλουθη πληροφορία:

N

Homo Sapiens: $\langle \text{sid}_1 \rangle, \langle \text{sid}_2 \rangle \dots \langle \text{sid}_N \rangle$

Homo Sapiens species: $\langle \text{NUM} \rangle$

DONE

όπου N είναι ο αριθμός των ειδών, για κάθε $i \in \{1, \dots, N\}$, $\langle \text{sid}_i \rangle$ είναι το αναγνωριστικό του είδους που αντιστοιχεί στο i -οστό φύλλο του δένδρου *homo Sapiens* και $\langle \text{NUM} \rangle$ είναι το πλήθος των μοναδικών ειδών στο δένδρο *homo Sapiens*.

J $\langle \text{sid} \rangle$

Γεγονός τύπου *population statistics*, που υποδηλώνει την εξαγωγή στατιστικών για τους πληθυσμούς που ανήκουν στο είδος με αναγνωριστικό $\langle \text{sid} \rangle$ και συνέβαλαν στο δένδρο *homo Sapiens*. Κατά το γεγονός αυτό, θα πρέπει αρχικά να εντοπίζεται γρήγορα το φύλλο που αποθηκεύει το είδος με αναγνωριστικό $\langle \text{sid} \rangle$ και στη συνέχεια θα πρέπει να καταγράφεται το πλήθος όλων των πληθυσμών που ανήκουν στο είδος.

Το γεγονός αυτό θα πρέπει να εκτελείται με χρονική πολυπλοκότητα $O(h + m)$, όπου h είναι το ύψος του δένδρου *homo Sapiens* και m είναι το πλήθος των πληθυσμών του είδους με αναγνωριστικό $\langle \text{sid} \rangle$ που υπάρχουν στο δένδρο *homo Sapiens*. Μετά το πέρας της εκτέλεσης ενός τέτοιου γεγονότος, το πρόγραμμα θα πρέπει να τυπώνει την ακόλουθη πληροφορία:

J $\langle \text{sid} \rangle$

Homo Sapiens populations: $\langle \text{NUM} \rangle$

DONE

όπου $\langle \text{NUM} \rangle$ είναι ο αριθμός των πληθυσμών που ανήκουν στο είδος με αναγνωριστικό $\langle \text{sid} \rangle$ από το δένδρο *homo Sapiens*.

P

Γεγονός τύπου *print species*, κατά το οποίο θα εκτυπώνεται το δένδρο ειδών, χωρίς τους επιμέρους πληθυσμούς κάθε είδους. Μετά το πέρας της εκτέλεσης ενός τέτοιου γεγονότος, το πρόγραμμα θα πρέπει να τυπώνει την ακόλουθη πληροφορία:

```

P
  <sid1> ... <sidN>
DONE

```

όπου N είναι ο αριθμός των κόμβων στο δένδρο ειδών και για κάθε $i \in \{1, \dots, N\}$, $\langle \text{sid}_i \rangle$ είναι το αναγνωριστικό του είδους που αντιστοιχεί στον i -οστό κόμβο του δένδρου αυτού, σύμφωνα με τη μετα-διατεταγμένη διάταξη.

X

Γεγονός τύπου *print populations*, κατά το οποίο θα εκτυπώνεται το δένδρο ειδών, αλλά και το δένδρο πληθυσμών του κάθε είδους. Μετά το πέρας της εκτέλεσης ενός τέτοιου γεγονότος, το πρόγραμμα θα πρέπει να τυπώνει την ακόλουθη πληροφορία:

```

X
  <sid1>
    <gid1,1, cid1,1>, <gid1,2, cid1,2> .... <gid1,M1, cid1,M1>
  ...
  <sidN>
    <gidN,1, cidN,1>, <gidN,2, cidN,2> .... <gidN,MN, cidN,MN>
DONE

```

όπου N είναι ο αριθμός των κόμβων στο δένδρο ειδών και για κάθε $i \in \{1, \dots, N\}$, $\langle \text{sid}_i \rangle$ είναι το αναγνωριστικό του είδους που αντιστοιχεί στον i -οστό κόμβο του δένδρου αυτού, σύμφωνα με τη μετα-διατεταγμένη διάταξη. Ακόμη, για κάθε $i \in \{1, \dots, N\}$, το M_i είναι το πλήθος των πληθυσμών που ανήκουν στο είδος με αναγνωριστικό $\langle \text{sid}_i \rangle$ και για κάθε $j \in \{1, \dots, M_i\}$, $\langle \text{gid}_{i,j} \rangle$ είναι το αναγνωριστικό του πληθυσμού που αντιστοιχεί στον j -οστό κόμβο του δένδρου πληθυσμών του είδους (σύμφωνα με την ενδο-διατεταγμένη διάταξη), ενώ $\langle \text{cid}_{i,j} \rangle$ είναι το αναγνωριστικό της ηπείρου όπου κατοικεί ο συγκεκριμένος πληθυσμός.

C

Γεγονός τύπου *print continents*, κατά το οποίο θα εκτυπώνεται ο πίνακας ηπείρων και το δένδρο πληθυσμών της κάθε ηπείρου. Μετά το πέρας της εκτέλεσης ενός τέτοιου γεγονότος, το πρόγραμμα θα πρέπει να τυπώνει την ακόλουθη πληροφορία:

```

Continent 0 : <gid0,1>, <gid0,2> .... <gid0,M0>
Continent 1 : <gid1,1>, <gid1,2> .... <gid1,M1>
...
Continent 4 : <gid4,1>, <gid4,2> .... <gid4,M4>
DONE

```

όπου 5 είναι το πλήθος των ηπείρων και για κάθε $i \in \{0, \dots, 4\}$ το M_i είναι το πλήθος των πληθυσμών που κατοικούν στην ήπειρο που αντιστοιχεί στην i -οστή θέση του πίνακα ηπείρων και για κάθε $j \in \{1, \dots, M_i\}$, $\langle \text{gid}_{i,j} \rangle$ είναι το αναγνωριστικό του πληθυσμού που αντιστοιχεί στον j -οστό κόμβο του δένδρου πληθυσμών της ηπείρου.

W

Γεγονός τύπου print evolution, κατά το οποίο θα εκτυπώνεται η αλυσίδα φύλλων του δένδρου homo Sapiens. Το γεγονός αυτό θα πρέπει να εκτελείται σε χρόνο $O(h+m+n)$, όπου h είναι το ύψος του δένδρου homo Sapiens, m είναι το πλήθος των φύλλων του και n είναι ο μέγιστος αριθμός πληθυσμών σε οποιοδήποτε από τα δένδρα πληθυσμών που δεικτοδοτούνται από τα φύλλα του δένδρου homo Sapiens. Μετά το πέρας της εκτέλεσης ενός τέτοιου γεγονότος, το πρόγραμμα θα πρέπει να τυπώνει την ακόλουθη πληροφορία:

W

Homo Sapiens: [$\langle \text{gid}_{1,1}, \text{sid}_{1,1} \rangle, \dots, \langle \text{gid}_{1,M_1}, \text{sid}_{1,M_1} \rangle$] [$\langle \text{gid}_{N,1}, \text{sid}_{N,1} \rangle, \dots, \langle \text{gid}_{N,M_N}, \text{sid}_{N,M_N} \rangle$]
DONE

όπου N είναι ο αριθμός των ειδών, για κάθε $i \in \{1, \dots, N\}$, M_i είναι το πλήθος των πληθυσμών που ανήκουν στο είδος που το δένδρο του ξεκινά στο i -οστό φύλλο του δένδρου homo Sapiens, ενώ για κάθε $j \in \{1, \dots, M_i\}$, $\langle \text{gid}_{i,j} \rangle$ είναι το αναγνωριστικό του πληθυσμού που αντιστοιχεί στον j -οστό κόμβο του δένδρου πληθυσμών του είδους αυτού.