

Ακολουθιακά Αρχεία

Ορισμός

Με τον όρο αρχείο εννοούμε ένα πλήθος πληροφοριών με συγκεκριμένη δομή, που είναι καταχωρημένες σε κάποιο μέσο. Πριν την εμφάνιση των υπολογιστών τα αρχεία ήταν χειρόγραφα και καταχωρούνταν σε χαρτί.

Σήμερα ο όρος όταν χρησιμοποιείται στο χώρο των υπολογιστών αφορά σε κάθε δεδομένο που είναι αποθηκευμένο με μόνιμο τρόπο σε μαγνητικό ή άλλο μέσο.

Ειδικότερα όταν ο όρος αφορά σε αλγορίθμους τότε εννοούνται αρχεία δεδομένων.

Κάθε αρχείο αποτελείται από ένα άγνωστο πλήθος εγγραφών. Κάθε εγγραφή αποτελείται από ένα πλήθος πεδίων ίδιων για κάθε εγγραφή. Κάθε πεδίο περιέχει ένα μέρος της συνολικής πληροφορίας, που αντιπροσωπεύει η εγγραφή.

Για παράδειγμα το αρχείο φοιτητών του τμήματος Πληροφορικής, έχει εγγραφές που κάθε μία αφορά σε ένα φοιτητή/τρια, ενώ κάθε πεδίο της εγγραφής είναι και μία μερική πληροφορία γι' αυτόν, όπως οΑΜΦ, το Επώνυμο, το Όνομα, κ.λπ.

Αλγοριθμικά

Το αρχείο ανάλογα με την χρήση του ονομάζεται, είτε εισόδου, είτε εξόδου, είτε εισόδου – εξόδου

- Εισόδου χαρακτηρίζεται το αρχείο που τα περιεχόμενα του αποτελούν είσοδο για τον αλγόριθμο.
- Εξόδου χαρακτηρίζεται το αρχείο που τα περιεχόμενα του αποτελούν έξοδο για τον αλγόριθμο.
- Εισόδου – Εξόδου χαρακτηρίζεται το αρχείο που τα περιεχόμενα του αποτελούν ταυτόχρονα και είσοδο για τον αλγόριθμο αλλά και έξοδο του.

Τα αρχεία εισόδου-εξόδου έχουν ξεχωριστή οργάνωση (τυχαία) και επιπλέον εντολές για την διαχείριση τους. Το κείμενο που ακολουθεί δεν αφορά σε αυτά τα αρχεία.

Το αρχείο δεδομένων για να χρησιμοποιηθεί από έναν αλγόριθμο πρέπει να συνδεθεί μαζί του. Η εντολή **Άνοιξε όνομα_αρχείου** εκτελεί αυτή την λειτουργία με διαφοροποίηση κατά περίπτωση:

- Άνοιξε (για αρχείο εισόδου): Η εντολή αναζητά το αρχείο και το συνδέει με τον αλγόριθμο, ώστε να είναι δυνατό να διαβαστεί.
- Άνοιξε (για αρχείο εξόδου): Η εντολή αναζητά κενό χώρο στο μέσο που θα γραφτεί το αρχείο και τον δεσμεύει για τον αλγόριθμο, ώστε να είναι δυνατό να γραφτούν τα αποτελέσματα που θα παραχθούν.

Το αρχείο μετά την χρήση του από τον αλγόριθμο πρέπει να αποδεσμευτεί. Η εντολή **Κλείσε όνομα_αρχείου** εκτελεί αυτή την λειτουργία με διαφοροποίηση κατά περίπτωση:

- Κλείσε (για αρχείο εισόδου): Η εντολή αποδεσμεύει το αρχείο.
- Κλείσε (για αρχείο εξόδου): Η εντολή τοποθετεί την ένδειξη τέλος αρχείου (EOF, end of file) και αποδεσμεύει το αρχείο.

Το αρχείο από την οργάνωση του μπορεί να είναι ακολουθιακό ή τυχαίο.

Ένα ακολουθιακό αρχείο έχει την μία εγγραφή μετά την άλλη (ακολουθεί η μία την άλλη). Όταν οι εγγραφές του είναι ταξινομημένες (σε σειρά) τότε το αρχείο λέγεται σειριακό.

Για παράδειγμα, ο χειρόγραφος κατάλογος των παρουσιών του εργαστηρίου είναι ένα ακολουθιακό αρχείο, ενώ τα αποτελέσματα των εξετάσεων που δημοσιεύονται στον πίνακα ανακοινώσεων σειριακό.

Η λογική μεταβλητή eof

Σε κάθε ακολουθιακό αρχείο ανατίθεται μία λογική μεταβλητή eof για να ελέγχεται το τέλος του. Σημειώστε ότι αν επιχειρηθεί να "διαβαστεί" ένα αρχείο που δεν έχει άλλες εγγραφές, θα προκληθεί λάθος στην εκτέλεση του αλγόριθμου.

Η τιμή της eof αλλάζει κάθε φορά που εκτελείται μία εντολή για αρχείο, όπως άνοιξε, κλείσε και διάβασε.

Συγκεκριμένα η eof είναι ψευδής όταν υπάρχουν και άλλες εγγραφές στο αρχείο και γίνεται αληθής όταν τελειώσει το αρχείο. Η τιμή της συνδέεται με τον εσωτερικό δείκτη του αρχείου.

Ο εσωτερικός δείκτης δείχνει την εγγραφή που θα διαβαστεί ή την εγγραφή που διαβάζεται. Η διαφορετική ένδειξη αφορά στην διαφορετική χρήση του, από τις γλώσσες προγραμματισμού. Σε άλλες γλώσσες προγραμματισμού δείχνει την εγγραφή που διαβάστηκε, ενώ σε άλλες την εγγραφή που θα διαβαστεί.

Έτσι η λογική συνάρτηση eof είναι αληθής σε διαφορετικές χρονικές στιγμές, που εξαρτώνται από τον τρόπο που λειτουργεί ο εσωτερικός δείκτης του αρχείου στην συγκεκριμένη γλώσσα.

Όπως όμως και αν γίνεται είναι φανερό ότι η λειτουργίες **Διάβασε** και **Αν eof** είναι δύο λειτουργίες που συνδέονται στενά. Στην μία περίπτωση προηγείται το **Διάβασε** και ακολουθεί το **Αν eof**, στην άλλη προηγείται ο **Αν eof** και ακολουθεί το **Διάβασε**. Ο στόχος είναι να μην επιχειρηθεί να διαβαστεί ένα αρχείο που δεν έχει άλλες εγγραφές.

Για τις ανάγκες του μαθήματος θα θεωρούμε ότι τα αρχεία εισόδου λειτουργούν με τον τρόπο (**Διάβασε**, **Αν eof**), και επομένως μετά από κάθε διάβασμα πρέπει να ακολουθεί ο έλεγχος της τιμής της eof.

Παρακάτω δίνονται δύο παραδείγματα για τις δύο διαφορετικές περιπτώσεις.

Παραδείγματα:

(Διάβασε και Αν eof)	(Διάβασε και Αν eof)
...	...
Άνοιξε αρχείο	Άνοιξε αρχείο
...	...
Διάβασε εγγραφή	Αρχή επανάληψης
Όσο NOT eof επανάλαβε	Διάβασε εγγραφή
Εντολές	Εντολές
Διάβασε εγγραφή	Μέχρι eof επανάλαβε
Τέλος επανάληψης	...
...	Κλείσε αρχείο
Κλείσε αρχείο	...
...	...

Σειριακά αρχεία

Ένα ακολουθιακό αρχείο που είναι ταξινομημένο λέγεται σειριακό.

Η ταξινόμηση γίνεται βάση κάποιου ή κάποιων πεδίων της εγγραφής.

Το κλειδί ταξινόμησης διακρίνεται σε πρωτεύον και δευτερεύοντα όταν είναι παραπάνω από ένα. Πρωτεύον κλειδί είναι το πεδίο βάση του οποίου είναι ταξινομημένο όλο το αρχείο για παράδειγμα το επώνυμο σε ένα κατάλογο. Δευτερεύον κλειδί ταξι-

νόμησης είναι το πεδίο βάση του οποίου γίνεται δεύτερη ταξινόμηση μέσα στην πρώτη, για παράδειγμα το όνομα στην προηγούμενη ταξινόμηση. Είναι φανερό ότι πρωτεύον κλειδί μπορεί να είναι μόνο ένα, ενώ δευτερεύοντα κλειδιά πολλά. Ακόμα είναι φανερό ότι κάθε κλειδί μπορεί να έχει διπλότυπα.

Τα σειριακά αρχεία έχουν συγκεκριμένου τύπου αλγορίθμους με τα οποία επεξεργάζονται. Τα προβλήματα είναι προβλήματα επιπέδων ελέγχου. (βλέπε σχετικές διαφάνειες).

Το γενικό μοντέλο ενός αλγόριθμου που χειρίζεται ένα ακολουθιακό αρχείο μπορεί να είναι το παρακάτω:

```
...
Ανοιξε αρχείο (είσοδος / έξοδος)
Διάβασε εγγραφή
Αν NOT eof
    τότε εντολής /ές ! αρχικοποίηση μεταβλητών
    ! εργασίες που γίνονται μόνο στην αρχή
    αλλιώς εντολή /ές ! προετοιμασία τερματισμού
Τέλος Αν
Όσο NOT eof επανάλαβε
    εντολή / ές
    Διάβασε εγγραφή
Τέλος επανάληψης
εντολή /ές
Κλείσε αρχείο
...
```