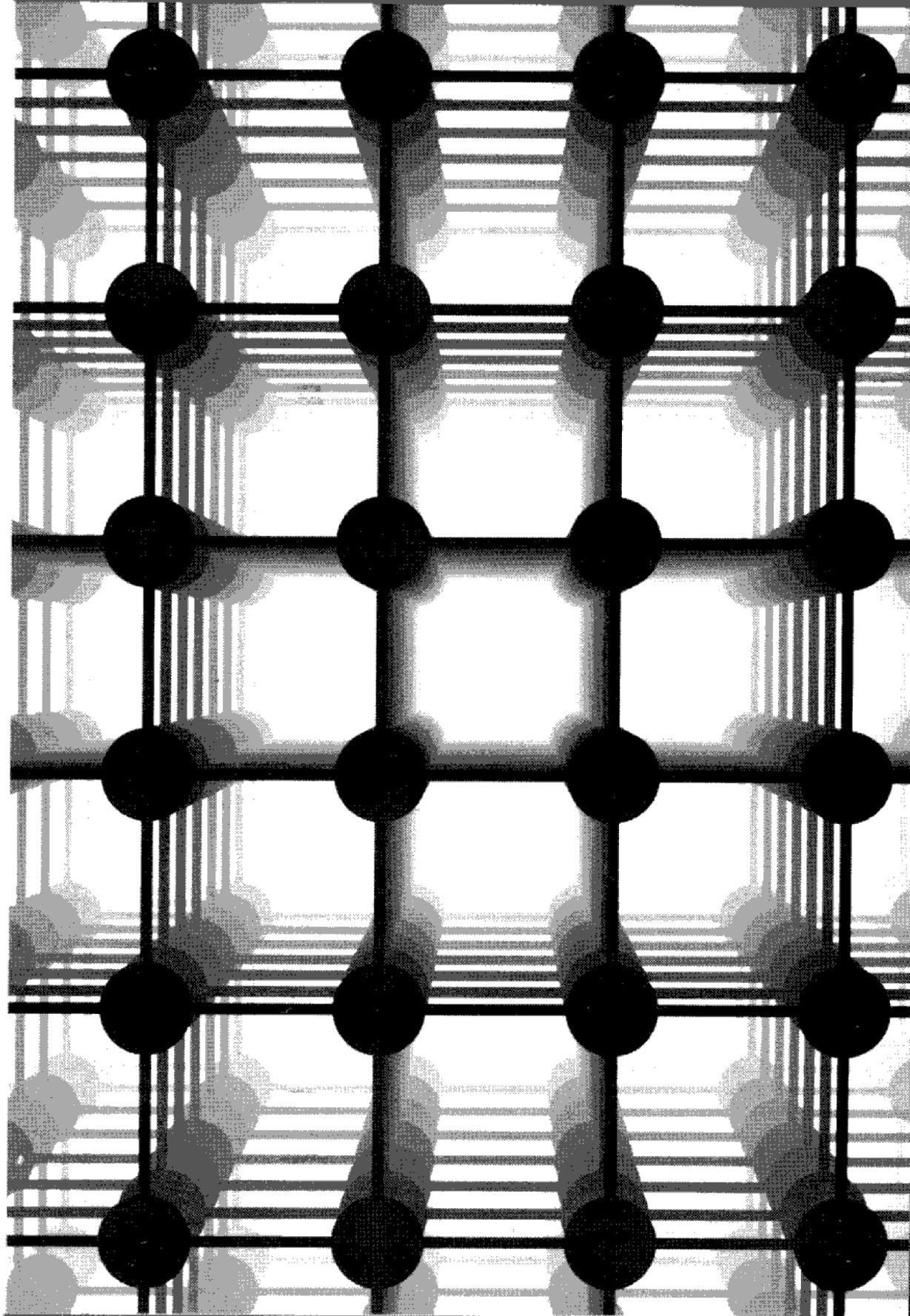


Χρήστου Κοίλια

ΔΙΟΜΕΤΕΣ ΔΕΛΤΑΟΜΕΤΩΝ



ΚΑΙ ΟΡΓΑΝΩΣΕΙΣ
ΑΡΧΕΙΩΝ

000000
101

Χρ. Κοίλιας

Δ Ω Ρ Ε Α

Τ.Ε.Ι. ΑΘΗΝΑΣ
ΒΙΒΛΙΟΘΗΚΗ
Αρ. εισ. 82026

ΔΟΜΕΣ ΔΕΔΟΜΕΝΩΝ
ΚΑΙ
ΟΡΓΑΝΩΣΕΙΣ ΑΡΧΕΙΩΝ

Αθήνα

Κάθε γνήσιο αντίγραφο έχει την υπογραφή του συγγραφέα

Έκδοση 1^η, Copyright © 1993.
Ανατυπώσεις: 1995, 1998, 2000
ISBN: 960-7235-44-4



ΕΚΔΟΣΕΙΣ ΝΕΩΝ ΤΕΧΝΟΛΟΓΙΩΝ
Στουρνάρη 49^Α, 106 82, Αθήνα
Τηλ. 3845594, Fax: 3808009

***Καλλιτεχνική επιμέλεια, Σχεδίαση εξωφύλλου,
Ηλεκτρονική σελιδοποίηση:***

Laser DTP Services, Β. Ουγκώ 45, 104 37 Αθήνα, τηλ. 5221949

Απαγορεύεται η με οποιονδήποτε τρόπο ανατύπωση, καταχώρηση σε σύστημα αποθήκευσης και επανάκτησης ή μετάδοση με κάθε μορφή και μέσο (ηλεκτρονικό, μηχανικό, φωτοαντιγραφικό κ.λπ.) του συνόλου ή μέρους του βιβλίου αυτού, χωρίς την έγγραφη άδεια του εκδότη.

*Στο Γιάννη Κόλλια
καθηγητή Ε.Μ.Π.
που έφυγε πρόωρα*

Περιεχόμενα

Πρόλογος	9
1. Πληροφορίες και δεδομένα	13
2. Αλγόριθμοι	17
2.1 Παράσταση αλγορίθμων	20
2.2 Θεωρία αλγορίθμων	24
2.3 Ανάλυση αλγορίθμων	26
2.4 Είδη αλγορίθμων	29
3. Γραμμικές Δομές Δεδομένων	33
3.1 Πίνακες	34
Ειδικές μορφές πινάκων	36
Συμβολοσειρές	37
Οι πίνακες στις γλώσσες προγραμματισμού	38
Ασκήσεις	43
3.2 Στοίβες	45
Ποληωνικός συμβολισμός	49
Ασκήσεις	53
3.3 Ουρές	56
Δακτύλιοι	58
Ουρές με δύο άκρα	58
FIFO και LIFO	59
Ασκήσεις	61
3.4 Λίστες	63
3.4.1 Εισαγωγή και διαγραφή κόμβου	64
3.4.2 Αποκομιδή αχρήστων δεδομένων	67
3.4.3 Διαχείριση ακυρωμένων κόμβων	68
Αντιστροφή λίστας	70
Συμμετρικές λίστες	70
Κυκλικές λίστες	70

Υλοποιήσεις στις γλώσσες προγραμματισμού.....	71
Ασκήσεις.....	79
4. Μη γραμμικές Δομές Δεδομένων.....	81
4.1 Δένδρα	82
Ορισμοί.....	82
Ποσοτικά στοιχεία	84
Αναπαράσταση δένδρων στη μνήμη του υπολογιστή	85
4.1.1 Δυαδικά δένδρα αναζήτησης.....	87
4.1.2 Διάσχιση δένδρων.....	90
Μετατροπή παραστάσεων	92
Νηματοποιημένα δένδρα	93
Ασκήσεις.....	95
4.2 Γράφοι.....	98
Ιστορικό	99
Εσωτερική παράσταση γράφων	99
Βασικοί ορισμοί και προτάσεις	100
Θεωρία γράφων.....	103
Αλγόριθμοι	104
5. Αναζήτηση	107
5.1 Σειριακή αναζήτηση.....	108
5.2 Αναζήτηση κατά ομάδες	109
5.3 Δυαδική αναζήτηση	111
5.4 Αναζήτηση Fibonacci.....	112
5.5 Αναζήτηση παρεμβολής	115
Ασκήσεις.....	117
6. Ταξινόμηση - σύζευξη	119
6.1 Ταξινόμηση με παρεμβολή.....	120
6.2 Ταξινόμηση με επιλογή	123
6.3 Ταξινόμηση με αντιμετάθεση.....	124
6.4 Ταξινόμηση παρεμβολής με φθίνοντα διαστήματα	128
6.5 Ταξινόμηση με διαμερισμό	129
6.6 Σύζευξη.....	133
Ασκήσεις.....	134
7. Αρχεία.....	137
7.1 Οι βοηθητικές μνήμες.....	138

Μαγνητικές ταινίες	138
Μαγνητικοί δίσκοι.....	139
Οπτικοί δίσκοι.....	142
7.2 Χαρακτηριστικά των αρχείων - ορισμοί.....	144
7.3 Είδη αρχείων	145
7.4 Επεξεργασία αρχείων	146
7.5 Κατηγορίες λογικών εγγραφών.....	151
7.6 Φυσική αποθήκευση αρχείων.....	153
7.7 Περιορισμοί στη χρήση αρχείων.....	156
7.8 Χαρακτηριστικά χρήσης των αρχείων	157
7.9 Οργανώσεις αρχείων.....	158
7.10 Προσπέλασης μέθοδοι	158
Ασκήσεις.....	160
8. Σειριακή οργάνωση	163
Επεξεργασία σειριακών αρχείων	164
Υλοποίηση των σειριακών αρχείων στις γλώσσες προγραμματισμού.....	166
Ασκήσεις.....	170
9. Σειριακή με δείκτες οργάνωση.....	173
9.1 Μέθοδος προσπέλασης ISAM	174
9.2 Αρχεία VSAM	180
Υλοποιήσεις στις γλώσσες προγραμματισμού.....	182
10. Απ' ευθείας οργάνωση.....	185
10.1 Άμεσα αρχεία	186
Υλοποιήσεις άμεσων αρχείων στις γλώσσες προγραμματισμού.....	188
Ασκήσεις.....	192
10.2 Με τη βοήθεια λεξικού	194
Αναζήτηση.....	195
Δημιουργία νέων εγγραφών	195
Μεταβολή εγγραφής	196
Ακύρωση εγγραφής	197
Αναδιοργάνωση	198
Δεικτοποίηση	198
Ασκήσεις.....	205
10.3 Με μετασχηματισμό κλειδιού	208
Ορισμοί	210
Συναρτήσεις κατακερματισμού	211

Διαχείριση συγκρούσεων	213
Ασκήσεις.....	220
11. Β-Δένδρα	223
Αναζήτηση κλειδιού	226
Εισαγωγή κλειδιού	227
Διαγραφή κλειδιού	230
Τα αρχεία δεικτών του Clipper και της dBASE	234
Ασκήσεις	237
12. Οργάνωση αρχείων σύμφωνα με δευτερεύοντα κλειδιά	239
12.1 Αντεστραμμένα λεξικά	241
12.2 Πολλοπληθές λίστες	243
Ασκήσεις.....	252
13. Ταξινόμηση αρχείων.....	253
13.1 Παράμετροι ταξινόμησης	254
13.2 Αλγόριθμος ταξινόμησης	256
13.3 Ταξινόμηση με επικάλυψη εισόδου-εξόδου.....	258
13.4 Ταξινόμηση σειριακών αρχείων	264
13.4.1 Απ'ευθείας σύζευξη	264
13.4.2 Φυσική σύζευξη	265
13.4.3 Ισοζυγισμένη σύζευξη	266
Ασκήσεις.....	267
14. Μερικά ειδικά θέματα	269
14.1 Συμπύεση δεδομένων	270
14.2 Ασφάλεια δεδομένων.....	277
14.3 Κριτήρια σχεδιασμού αρχείων	287
15. Επίλογος.....	291
Βιβλιογραφία	295
Ευρετήριο όρων	297

Πρόλογος

Ο προγραμματισμός είναι μέχρι σήμερα η κυριότερη μέθοδος, με την οποία ο άνθρωπος καθοδηγεί τον υπολογιστή στην επίλυση προβλημάτων. Πρόκειται για μια από τις συνθετότερες νοητικές διαδικασίες, η οποία αντιστέκεται ακόμη στις διάφορες προσπάθειες συστηματοποίησής της που γίνονται συνεχώς από τους ερευνητές. Βασικά συστατικά κάθε προγράμματος είναι οι αλγόριθμοι και οι δομές δεδομένων. Για τη συγγραφή επιτυχημένων προγραμμάτων είναι απαραίτητες οι τεχνικές δόμησης και αποσύνθεσης ενός προβλήματος σε μικρότερα τμήματα κατάλληλου μεγέθους και πολυπλοκότητας, ώστε να είναι δυνατή η υλοποίησή τους.

Δεν είναι λοιπόν αδικαιολόγητο το γεγονός, ότι το αντικείμενο Δομές Δεδομένων αποτελεί από πολλά χρόνια αυτόνομο μάθημα υποδομής σε όλα τα Ιδρύματα Τριτοβάθμιας Εκπαίδευσης που διδάσκεται η Πληροφορική.

Στο βιβλίο αυτό παρουσιάζονται οι σπουδαιότερες δομές, που χρησιμοποιούνται τόσο στην κύρια μνήμη όσο και στη βοηθητική. Απευθύνεται σε σπουδαστές και φοιτητές τμημάτων πληροφορικής των ΑΕΙ και ΤΕΙ, καθώς επίσης και σε μαθητές ΤΕΛ, ΕΠΛ και ΙΕΚ αντίστοιχων ειδικοτήτων. Στο βιβλίο αυτό σε μεγάλο βαθμό αντανakλάται η πολυετής πείρα του συγγραφέα στη διδασκαλία του ομώνυμου μαθήματος στο Τμήμα Πληροφορικής του ΤΕΙ Αθήνας.

Σε κάθε αναφερόμενη δομή παρουσιάζονται μέθοδοι (αλγόριθμοι) για την αποτελεσματική επεξεργασία της, ώστε να αντλούνται οι ζητούμενες κάθε φορά πληροφορίες. Για την παρουσίαση των σχετικών αλγορίθμων προτιμήσαμε τη χρήση της ψευδογλώσσας, προκειμένου να υπάρχει ανεξαρτησία από τις γνώσεις προγραμματισμού, παρόλο που τα περισσότερα σχετικά βιβλία χρησιμοποιούν την Pascal για το σκοπό αυτό. Πιστεύουμε ότι η χρήση της ψευδογλώσσας παρουσιάζει πλεονεκτήματα, γιατί παρέχει μεγαλύτερες ελευθερίες για τον ορισμό της δομής και των λειτουργιών επ' αυτής. Αλήθεια η υλοποίηση σε οποιαδήποτε αλγοριθμική γλώσσα είναι εξαιρετικά ευχερής. Ωστόσο σε μερικά κρίσιμα ζητήματα, όπως οι πίνακες, οι δείκτες και τα αρχεία παρουσιάζεται στα αντίστοιχα κεφάλαια, πώς αντιμετωπίζονται από τις διάφορες γλώσσες προγραμματισμού.

Η επεξήγηση των σχετικών εννοιών διευκολύνεται σημαντικά με την παράθεση πλήθους σχημάτων. Σε πολλές μάλιστα περιπτώσεις κρίθηκε επιβεβλημένη και η χρήση χρώματος λόγω αυξημένης πολυπλοκότητας.

Τέλος σημαντική καινοτομία αποτελεί το περιεχόμενο της δισκέτας που συνοδεύει το βιβλίο. Σε αυτή συμπεριλαμβάνονται σχεδόν όλοι οι αλγόριθμοι του βιβλίου αναπτυγμένοι στις γλώσσες προγραμματισμού Pascal, C, GWBASIC, QuickBASIC και Clipper.

Για όλους τους παραπάνω λόγους πιστεύουμε, ότι το βιβλίο αυτό θα αντιμετωπισθεί ευνοϊκά από τη σπουδάζουσα νεολαία στον τομέα της Πληροφορικής.

Η ύλη του βιβλίου

Το βιβλίο λογικά χωρίζεται σε τρεις ενότητες. Η πρώτη περιλαμβάνει τα κεφάλαια 1 και 2, τα οποία είναι εισαγωγικά στις δομές και τους αλγορίθμους. Η δεύτερη περιλαμβάνει τα κεφάλαια 3 έως 6 και ασχολείται με τις δομές που χρησιμοποιούνται στην κύρια μνήμη. Τα υπόλοιπα κεφάλαια συγκροτούν την τρίτη ενότητα, η οποία περιλαμβάνει τις δομές που χρησιμοποιούνται σε βοηθητική μνήμη.

Στο **κεφ. 1** γίνεται μια εισαγωγή στις έννοιες πληροφορία και δεδομένο.

Το **κεφ. 2** ασχολείται γενικά με τους αλγορίθμους. Γίνεται η θεμελίωσή τους, παρουσιάζονται οι τρόποι παράστασής τους, οι διάφορες κατηγορίες και διατυπώνονται βασικά στοιχεία ανάλυσης και θεωρίας αλγορίθμων. Στο κεφάλαιο αυτό παρουσιάζονται ακόμη όλη τα στοιχεία και σχήματα που απαιτούνται για την παράσταση αλγορίθμων με χρήση ψευδογλώσσας, τα οποία χρησιμοποιούνται στη συνέχεια του βιβλίου.

Το **κεφ. 3** είναι αφιερωμένο στις γραμμικές δομές δεδομένων, δηλαδή στις δομές των οποίων τα στοιχεία παρουσιάζουν μια φυσική ή λογική γειτονικότητα. Στην πρώτη ενότητα παρουσιάζεται η βασική δομή του πίνακα. Αναφέρεται ο τρόπος αποθήκευσης αριθμητικών και αλφαριθμητικών πινάκων στη μνήμη και πώς χρησιμοποιούνται στις διάφορες γλώσσες προγραμματισμού. Στην επόμενη ενότητα αναλύεται η δομή της στοίβας και η χρήση της στη μετατροπή και αποτίμηση παραστάσεων. Οι ουρές και οι παραλληλαγές τους είναι το αντικείμενο της τρίτης ενότητας του κεφαλαίου αυτού. Τέλος στην τέταρτη ενότητα παρουσιάζονται οι λίστες. Αναλύεται πώς γίνεται η εισαγωγή και η διαγραφή κόμβου σε μια λίστα, είτε με τη χρήση πινάκων είτε με τη χρήση δεικτών. Μεγάλο μέρος της ενότητας αυτής είναι αφιερωμένο στην επεξήγηση των ληπτών μηχανισμών της γλώσσας Pascal στη διαχείριση δεικτών.

Στο **κεφ. 4** παρουσιάζονται οι μη γραμμικές δομές δεδομένων. Στην πρώτη ενότητα του κεφαλαίου αναπτύσσεται η πολύ σημαντική δομή του δένδρου δίνοντας έμφαση στα δυαδικά δένδρα αναζήτησης. Παρουσιάζονται οι τρόποι παράστασης στη μνήμη, η αναζήτηση κόμβου, οι μέθοδοι διάσχισης κ.α. Στη δεύτερη ενότητα του κεφαλαίου δίδονται βασικά στοιχεία από τη δομή του γράφου.

Στο **κεφ. 5** αναπτύσσονται οι πέντε πιο βασικοί αλγόριθμοι αναζήτησης σε πίνακες, ενώ στο **κεφ. 6** παρουσιάζονται αντίστοιχα οι πιο σπουδαίοι αλγόριθμοι εσωτερικής ταξινόμησης.

Στο **κεφ. 7** γίνεται μια γενική παρουσίαση των αρχείων δεδομένων. Προηγούνται μερικές υπενθυμίσεις από την τεχνολογία των βοηθητικών μνημών και ακολουθούν στοιχεία για τα χαρακτηριστικά των αρχείων, τα είδη αυτών, τις λογικές και φυσικές εγγραφές, καθώς και τις δυνατές επεξεργασίες που μπορούν να γίνουν επί αυτών. Το κεφάλαιο κλείνει με μια εισαγωγή στις μεθόδους οργάνωσης και προσπέλασης αρχείων.

Στα επόμενα τρία κεφάλαια παρουσιάζονται οι βασικές οργανώσεις αρχείων. Πιο συγκεκριμένα, στο **κεφ. 8** αναπτύσσεται η σειριακή οργάνωση, ενώ στο **κεφ. 9** η σειριακή με δείκτες οργάνωση, όπου παρουσιάζονται οι μέθοδοι ISAM και VSAM με έμφαση στην πρώτη.

Στο **κεφ. 10** παρουσιάζονται τα απ' ευθείας οργανωμένα αρχεία. Το κεφάλαιο αυτό χωρίζεται σε τρεις μεγάλες ενότητες που θα μπορούσαν να αποτελούν αυτόνομα κεφάλαια. Κάθε μια αφιερώνεται σε μια οικογένεια μεθόδων απ' ευθείας οργάνωσης, που είναι: τα άμεσα αρχεία, τα αρχεία με χρήση λεξικού και τα αρχεία στα οποία χρησιμοποιείται μετασχηματισμός κλειδιού.

Στο **κεφ. 11** παρουσιάζεται μια προωθημένη και ιδιαίτερα αποδοτική οργάνωση αρχείων με χρήση B-δένδρων.

Στο **κεφ. 12** αναφέρονται μέθοδοι οργάνωσης των αρχείων σύμφωνα με δευτερεύοντα κλειδιά, ενώ στο **κεφ. 13** δίδονται μερικοί αλγόριθμοι ταξινόμησης αρχείων.

Τέλος στο **κεφ. 14** παρουσιάζονται μερικά θέματα που μπορούν να χαρακτηρισθούν ως ειδικά. Τα θέματα αυτά είναι η συμπίεση δεδομένων, η ασφάλεια αρχείων, τα κριτήρια επιλογής αρχείων και μερικές σκέψεις για το μέλλον τους.

Το βιβλίο κλείνει με έναν επίλογο που παρουσιάζεται στο **κεφ. 15**.

Σημείωση προς τον αναγνώστη

Η πλειοψηφία των κεφαλαίων του βιβλίου συμπληρώνεται με ένα αριθμό υποδειγματικών ασκήσεων. Συνιστάται στον αναγνώστη, παράλληλα με τη μελέτη του βιβλίου να δουλεύει πάνω στις ασκήσεις αυτές. Οι περισσότερες δε από αυτές ζητούν τη δημιουργία κάποιου προγράμματος ή ρουτίνας, με το οποίο υλοποιείται μια δομή. Δημιουργήστε αυτά τα προγράμματα στη γλώσσα που γνωρίζετε καλύτερα ή στη γλώσσα που διδάσκεσθε ταυτόχρονα. Με τον τρόπο αυτό βλέποντας στην πράξη τις λειτουργίες των δομών θα κατανοηθεί πλήρως η θεωρία. Επιπλέον

ον, προοδευτικά θα δημιουργείται ένα σύνολο ρουτινών, το οποίο θα είναι σίγουρα πολύτιμο στο μέλλον για την ανάπτυξη ολοκληρωμένων εφαρμογών. Στο θέμα αυτό ο αναγνώστης θα βρει σημαντική βοήθεια από τις έτοιμες ρουτίνες της δισκέτας που συνοδεύει το βιβλίο.

Σημείωση προς το διδάσκοντα

Η διδασκαλία του αντικειμένου απαιτεί τουλάχιστον ένα εξάμηνο σπουδών με 4 ώρες ανά εβδομάδα. Αν στο ίδιο εξάμηνο υπάρχει διδασκαλία γλώσσας προγραμματισμού, τότε πρέπει να γίνεται συνεργασία των εισηγητών, ώστε οι σπουδαστές να υλοποιούν ζητήματα που έχουν διδαχθεί στις Δομές Δεδομένων. Σε κάθε περίπτωση ο διδάσκων πρέπει να ενθαρύνει τους σπουδαστές να επεξεργάζονται τις σχετικές ασκήσεις.

Ανάλογα με το διατιθέμενο χρόνο και τα άλλα μαθήματα ο εισηγητής μπορεί, αν το κρίνει, να προτάξει τα θέματα των αρχείων (κεφ. 7-10). Καλό είναι για τα ίδια θέματα, να δημιουργήσει ορισμένες διαφάνειες από τα σχήματα του βιβλίου, τις οποίες να παρουσιάζει στην τάξη κατά την παράδοση.

Συμβάσεις

Για την παρουσίαση των αλγορίθμων σε ψευδογλώσσα έχει χρησιμοποιηθεί διαφορετική γραμματοσειρά. Όσες λέξεις έχουν γραφεί με έντονα γράμματα (μαύρα), αυτές αποτελούν λέξεις-κλειδιά και ανήκουν στο ρεπερτόριο λέξεων που χρησιμοποιούνται για την περιγραφή των αλγορίθμων. Οι υπόλοιπες είναι μεταβλητές ή σταθερές που ορίζει ο χρήστης.

Το σύμβολο του λογαρίθμου \log , αν δεν αναφέρεται διαφορετικά, θα εννοείται με βάση το 2.

Κάθε αναφορά στη βιβλιογραφία συμβολίζεται με [η], όπου η ο αριθμός του αντίστοιχου βιβλίου. Ο κατάλογος των βιβλιογραφικών πηγών αναφέρεται στο τέλος του βιβλίου.

Χρ. Κοίλιας

Το βιβλίο αυτό φιλοδοξεί να γίνει το κλασικό εγχειρίδιο στις Δομές Δεδομένων και τις Οργανώσεις Αρχείων. Απευθύνεται σε σπουδαστές και φοιτητές τμημάτων πληροφορικής ΑΕΙ, ΤΕΙ, ΙΕΚ, Τεχνικών και Επαγγελματικών Λυκείων και ιδιωτικών σχολών αντίστοιχου επιπέδου, καθώς και σε κάθε νέο επαγγελματία του κλάδου.

Στο βιβλίο παρουσιάζονται οι δομές που χρησιμοποιούνται τόσο στην κύρια όσο και σε βοηθητικές μνήμες και αναπτύσσονται οι σχετικοί αλγόριθμοι και τεχνικές για τη διαχείρισή τους.

Περιλαμβάνονται:

- Πληροφορίες και Δεδομένα
- Αλγόριθμοι
- Πίνακες, στοίβες, ουρές, λίστες, δένδρα και γράφοι
- Αναζήτηση και ταξινόμηση
- Αρχεία δεδομένων
- Σειριακά, σειριακά με δείκτες και απ' ευθείας οργανωμένα αρχεία
- Δενδρικοί κατάλογοι και αντεστραμμένα αρχεία
- Ταξινόμηση αρχείων.
- Ειδικά θέματα

Το βιβλίο συνοδεύεται με δισκέτα στην οποία περιέχονται όλοι οι αλγόριθμοι του βιβλίου αναπτυγμένοι στις γλώσσες προγραμματισμού Pascal, C, GWBASIC και QuickBASIC, dBASE και Clipper.