



ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ

Τμήμα Πληροφορικής

ΕΠΛ 232 – Προγραμματιστικές Τεχνικές και Εργαλεία ΑΣΚΗΣΗ 2: Επεξεργασία Χημικών Τύπων με Δυναμικές Δομές Δεδομένων

Διδάσκων: Ανδρέας Αριστείδου
Υπεύθυνος Άσκησης: Πύρρος Μπράτσкас

Ημερομηνία Ανάθεσης: Δευτέρα, 14 Οκτωβρίου 2024
Ημερομηνία Παράδοσης: Τρίτη, 31 Οκτωβρίου 2024, Ώρα 13:00

(ο κώδικας να υποβληθεί μέσω του Moodle)

I. Στόχος Άσκησης

Στην εργασία αυτή θα ασχοληθούμε με στοίβες, ταξινομημένες δομές και δυναμική δέσμευση μνήμης. Το θέμα της άσκησης είναι η υλοποίηση ενός προγράμματος επεξεργασίας χημικών τύπων. Οι λειτουργίες του προγράμματος σας και το αναμενόμενο αποτέλεσμα περιγράφονται αναλυτικότερα στην συνέχεια. Για την υλοποίηση της εργασίας θα πρέπει, πέρα από την χρήση συναρτήσεων και τεχνικών δομημένου προγραμματισμού, να χρησιμοποιήσετε τα ακόλουθα στοιχεία:

1. Διάσπαση του προγράμματος σε πολλαπλά αρχεία .c και .h με χρήση generic **makefile**. Ενδεικτικά αναφέρουμε ότι το πρόγραμμα μπορεί να αποτελείται από μερικά αντικείμενα και αρχεία κεφαλίδας (.c και .h).
2. **Κάθε αντικείμενο** πρέπει να συμπεριλαμβάνει τον **σχετικό οδηγό χρήσης (driver functions)**, δείτε διάλεξη 12) και θα πρέπει να έχετε σχόλια και οδηγό σχολίων με χρήση του **doxygen**.
3. Χρήση του εργαλείου **valgrind** για έλεγχο διαρροών μνήμης.

Θα ξεκινήσουμε με μια εισαγωγή και με την περιγραφή του προγράμματος και στη συνέχεια θα δώσουμε τα ζητούμενα.

II. Εισαγωγή

Χημικός τύπος είναι ο τρόπος αναπαράστασης της αναλογίας των ατόμων ή και κάποιων επιπλέον πληροφοριών για τη μοριακή δομή μιας χημικής ένωσης ή ουσίας. Για την αναπαράσταση αυτή χρησιμοποιούνται τα χημικά σύμβολα των χημικών στοιχείων που μετέχουν, κάποιοι αριθμοί που εκφράζουν τις αναλογίες των ατόμων γενικά ή ειδικά στο μόριο, ονομαζόμενοι ως δείκτες (π.χ. O_3), σύμβολα όπως παρενθέσεις (π.χ. $Ca(OH)_2$), αλλά και κάποια επιπλέον σύμβολα που δεν μας απασχολούν στα πλαίσια της εργασίας αυτής.

Η απλούστερη μορφή χημικού τύπου ονομάζεται *Εμπειρικός Τύπος (ET)*, που χρησιμοποιεί μόνο τα χημικά σύμβολα των στοιχείων που μετέχουν σε ένα χημικό είδος (ή μίγμα αυτών) και αριθμούς αναλογίας συμμετοχής του καθενός από αυτά. Συνήθως ο αριθμός 1 παραλείπεται και εννοείται όταν δεν δηλώνεται άλλος αριθμός αναλογίας. Ο αποκαλούμενος *Μοριακός Τύπος (MT)*, εκφράζει είδος χημικού τύπου που επίσης περιέχει χημικά σύμβολα και αριθμούς. Συχνά, οι όροι χημικός τύπος και μοριακός τύπος συγχέονται και αναφέρεται ο όρος χημικός τύπος εννοώντας ειδικά το μοριακό τύπο ενός χημικού είδους. Μερικές φορές ο εμπειρικός τύπος και ο μοριακός τύπος ενός χημικού είδους ταυτίζονται. Ο μοριακός τύπος ενός χημικού είδους, όμως, μπορεί να περιέχει μεγαλύτερους αριθμούς από τον αντίστοιχο εμπειρικό. Για παράδειγμα ο εμπειρικός τύπος της γλυκόζης είναι CH_2O , ενώ ο μοριακός τύπος της είναι $C_6H_{12}O_6$, δηλαδή προκύπτει από τον εμπειρικό αν πολλαπλασιαστούν όλοι οι συντελεστές αναλογίας επί 6. Σε αυτή την εργασία θα ασχοληθούμε μόνο με μοριακούς τύπους.

III. Περιγραφή

Ο αριθμός των πρωτονίων του πυρήνα ενός ατόμου ονομάζεται *Ατομικός Αριθμός*. Είναι χαρακτηριστικό κάθε χημικού στοιχείου, μιας και είναι αυτό που καθορίζει την ταυτότητά του. Το σύμβολο και το όνομα του στοιχείου, εξαρτώνται από τον ατομικό αριθμό, αφού είναι αυτός που το καθορίζει. Για παράδειγμα ένας πυρήνας που έχει ατομικό αριθμό 6, συμβολίζεται με το γράμμα C και ονομάζεται άνθρακας. Χάρη σε αυτό το ατομικό αριθμό καθορίζεται η θέση κάθε στοιχείου στον περιοδικό πίνακα, όπως φαίνεται και στην παρακάτω εικόνα. Οι χημικές ιδιότητες των στοιχείων είναι περιοδικές συναρτήσεις του ατομικού τους αριθμού.

The image shows a standard periodic table of elements. The columns are labeled Group 1 through 18, and the rows are labeled Period 1 through 7. The elements are arranged in their standard positions, with color-coding for different groups: Group 1 (red), Group 2 (orange), Groups 3-10 (green), Groups 11-12 (yellow), Groups 13-16 (blue), Group 17 (purple), Group 18 (pink), and the lanthanide and actinide series (grey) are shown below the main table.

Εικόνα 1: Τυπική μορφή του περιοδικού πίνακα

Σε αυτήν την εργασία θα ασχοληθούμε με μοριακούς τύπους οι οποίοι περιέχουν χημικά σύμβολα, αριθμούς και παρενθέσεις. Ένας μοριακός τύπος μπορεί να αποτελείται από ένα μόνο άτομο (π.χ. H) ή ένα μόνο άτομο που επαναλαμβάνεται (π.χ. O₃). Μέρος του χημικού τύπου μπορεί να επαναληφθεί περικλείοντας το σε παρένθεση που ακολουθείται από έναν ακέραιο αριθμό μεταξύ 2 και 99 (π.χ. Ca(OH)₂). Μπορεί ακόμη να αποτελείται από πολλαπλά επίπεδα παρενθέσεων (π.χ. Co₃(Fe(CN)₆)₂). Οι παρενθέσεις παραλείπονται, αν το επαναλαμβανόμενο συστατικό είναι ένα μόνο άτομο (π.χ. H₂O, C₆H₁₂O₆, Na₂SO₄).

IV. Ζητούμενα Άσκησης

Ζητούμενο αυτής της άσκησης είναι η κατασκευή ενός προγράμματος στη γλώσσα προγραμματισμού C, το οποίο υπολογίζει το συνολικό αριθμό πρωτονίων για ένα έγκυρο μοριακό χημικό τύπο. Συγκεκριμένα, το πρόγραμμα σας πρέπει:

- Να φορτώσει στη μνήμη, σε μια ταξινομημένη δομή, το περιοδικό πίνακα. Η δομή μπορεί να είναι ένας πίνακας ο οποίος πρέπει να ταξινομηθεί χρησιμοποιώντας έναν από τους αλγόριθμους που έχετε ήδη διδαχτεί. Τα στοιχεία του περιοδικού πίνακα βρίσκονται σε ένα αρχείο που πρέπει να δίνεται στην γραμμή εντολής ως πρώτη παράμετρος του προγράμματος. Κάθε γραμμή αυτού του αρχείου περιέχει το σύμβολο ενός χημικού στοιχείου και δίπλα του, τον ατομικό αριθμό (είναι η ηλεκτρονική μορφή του περιοδικού πίνακα τις εικόνας 1).

periodicTable.txt

H	1
He	2
Li	3
Be	4
...	

- Να διαβάσει από ένα αρχείο το οποίο δίνεται στην γραμμή εντολής ως παράμετρος του προγράμματος, τους χημικούς τύπους. Ο κάθε χημικός τύπος μπορεί να έχει οποιοδήποτε αριθμό (φωλιασμένων ή μη) παρενθέσεων. Για κάθε έναν από τους χημικούς τύπους του αρχείου εισόδου, το πρόγραμμα πρέπει να δημιουργεί την εκτεταμένη μορφή του, και να το γράψει σε ένα αρχείο εξόδου.

Παράδειγμα:

Για το H η εκτεταμένη μορφή του είναι: H

Για το KBr η εκτεταμένη μορφή του είναι: K Br

Για το $\text{Co}_3(\text{Fe}(\text{CN})_6)_2$ η εκτεταμένη μορφή του είναι:

Co Co Co Fe C N C N C N C N C N C N Fe C N C N C N C N C N C N

- Να υπολογίζει το συνολικό αριθμό πρωτονίων. Για να κάνει αυτό πρέπει να βρει από το περιοδικό πίνακα το ατομικό αριθμό κάθε στοιχείου του χημικού τύπου και να κάνει τις κατάλληλες πράξεις λαμβάνοντας υπόψη την ύπαρξη ή μη παρενθέσεων και δεικτών στο χημικό τύπο.

Παράδειγμα:

Για το H ο συνολικός αριθμός πρωτονίων είναι: 1

Για το KBr ο συνολικός αριθμός πρωτονίων είναι: $19+35=54$ όπου 19 είναι ο ατομικός αριθμός του στοιχείου K και 35 ο ατομικός αριθμός του στοιχείου Br.

Για το $\text{Co}_3(\text{Fe}(\text{CN})_6)_2$ ο συνολικός αριθμός πρωτονίων είναι: $3*27+2*(26+6*(6+7))=289$ όπου 27 είναι ο ατομικός αριθμός του στοιχείου Co, 26 ο ατομικός αριθμός του στοιχείου Fe, 6 είναι ο ατομικός αριθμός του στοιχείου C, και 7 ο ατομικός αριθμός του στοιχείου N. Όπως φαίνεται από τους υπολογισμούς, ο ατομικός αριθμός του στοιχείου Co πολλαπλασιάζεται με 3 καθώς το άτομο Co βρίσκεται 3 φορές στο μόριο, ενώ το άθροισμα των ατομικών αριθμών των ατόμων C και N πολλαπλασιάζεται με 6. Όπως μπορείτε να καταλάβετε, οι παρενθέσεις καθορίζουν και την προτεραιότητα των πράξεων.

- Όλοι οι υπολογισμοί των ζητούμενων της άσκησης πρέπει να γίνουν χρησιμοποιώντας δυναμικές στοίβες. Για παιδαγωγικούς λόγους, απαγορεύεται στην εργασία αυτή να γίνει χρήση αναδρομικών συναρτήσεων.

V. Ορίσματα Προγράμματος

α) Επικύρωση ορθότητας παρενθέσεων (validation)

Αυτή η λειτουργία επικυρώνει την ορθότητα (ισοζυγισμός) των παρενθέσεων, δεικτών, που υπάρχουν σε όλους τους χημικούς τύπους του αρχείου εισόδου. Η λειτουργία αυτή θα ενεργοποιείται όταν δίνεται η παράμετρος `-v` στη γραμμή εντολής. Το πρόγραμμα, στην περίπτωση που ανιχνευθεί ότι ένα αρχείο εισόδου περιέχει έστω και ένα χημικό τύπο που δεν πληροί τον έλεγχο επικύρωσης ορθότητας, πρέπει να εκτυπώσει ένα μήνυμα για κάθε γραμμή του αρχείου όπου υπάρχει λάθος (να εκτυπώσει και τον αριθμό της γραμμής π.χ. `Parentheses NOT balanced in line: 3`). Σε περίπτωση ορθότητας παρενθέσεων: `Parentheses are balanced for all chemical formuals`. Για αυτήν την περίπτωση η κλήση του προγράμματος είναι η ακόλουθη:

```
./parseFormula periodicTable.txt -v <input.txt>
```

β) Εκτύπωση σε εκτεταμένη μορφή σε αρχείο

Αυτή η λειτουργία μετατρέπει τους χημικούς τύπους του αρχείου εισόδου, σε εκτεταμένη μορφή, όπως περιγράφεται στα ζητούμενα της άσκησης. Η λειτουργία θα ενεργοποιείται όταν δίνεται η παράμετρος `-ext` στη γραμμή εντολής. Τα αποτελέσματα θα γραφτούν σε ένα αρχείο εξόδου το όνομα του οποίου δίνεται στη γραμμή εντολής. Για αυτήν την περίπτωση η κλήση του προγράμματος είναι η ακόλουθη:

```
./parseFormula periodicTable.txt -ext <input.txt> <output.txt>
```

γ) Υπολογισμός συνολικού αριθμού πρωτονίων

Αυτή η λειτουργία υπολογίζει το συνολικό αριθμό πρωτονίων για κάθε χημικό τύπο του αρχείου εισόδου. Η λειτουργία θα ενεργοποιείται όταν δίνεται η παράμετρος `-pn` στη γραμμή εντολών. Τα

αποτελέσματα θα γραφτούν σε ένα αρχείο εξόδου το όνομα του οποίου δίνεται στη γραμμή εντολής. Για αυτήν την περίπτωση η κλήση του προγράμματος είναι η ακόλουθη:

```
./parseFormula periodicTable.txt -pn <input.txt> <output.txt>
```

VI. Παραδείγματα εκτέλεσης

Έστω το αρχείου εισόδου `testFile.txt` με τα παρακάτω περιεχόμενα:

```
H
KBr
Ca(OH)2
H2O
C6H12O6
Ni(NO3)2
Co3(Fe(CN)6)2
```

```
./parseFormula periodicTable.txt -v testFile.txt
```

```
Verify balanced parentheses in testFile.txt
Parentheses are balanced for all chemical formulas
```

```
./parseFormula periodicTable.txt -ext testFile.txt extFile.txt
```

```
Compute extended version of formulas in chemFormulas.txt
Writing formulas to extFile.txt
```

```
./parseFormula periodicTable.txt -pn testFile.txt pnFile.txt
```

```
Compute total proton number of formulas in testFile.txt
Writing formulas to pnFile.txt
```

Παράδειγμα αρχείου εξόδου (`pnFile.txt`):

```
1
54
38
10
96
90
289
```

Παράδειγμα αρχείου εξόδου (`extFile.txt`):

```
H
K Br
Ca O H O H
H H O
C C C C C H H H H H H H H H H H O O O O O
Ni N O O O N O O O
Co Co Co Fe C N C N C N C N C N C N Fe C N C N C N C N C N
```

Έστω το αρχείου εισόδου `testFile1.txt` με τα παρακάτω περιεχόμενα:

```
H
KBr
```

Ca (OH) 2
 H2O
 C6H12O6
 Ni (NO3) 2
 Co3 (Fe (CN) 6)) 2

```
./parseFormula periodicTable.txt -v testFile1.txt
```

```
Verify balanced parentheses in testFile.txt
Parentheses NOT balanced in line: 3
Parentheses NOT balanced in line: 7
```

VII. Γενικές Οδηγίες

Το πρόγραμμα σας θα πρέπει να συμβαδίζει με το πρότυπο ISO C, να περιλαμβάνει εύστοχα και περιεκτικά σχόλια, να έχει καλή στοίχιση και το όνομα κάθε μεταβλητής, σταθεράς, ή συνάρτησης να είναι ενδεικτικό του ρόλου της. **Να χρησιμοποιήσετε το λογισμικό τεκμηρίωσης doxygen** έτσι ώστε να μπορούμε να μετατρέψουμε τα σχόλια του προγράμματός σας σε HTML αρχεία και να τα δούμε με ένα browser.

Επίσης, σας θυμίζουμε ότι κατά την διάρκεια της εκτέλεσης του προγράμματος σας αυτό θα πρέπει να δίνει τα κατάλληλα μηνύματα σε περίπτωση λάθους.

Το πρόγραμμα σας θα πρέπει να μεταγλωττίζεται στις μηχανές του εργαστηρίου.

Παραδώστε στο Moodle όλα τα αρχεία που έχουν σχέση με την εργασία (κώδικας *.h, *.c με doxygen σχόλια, doxygen configuration file, README.md).

VIII. Κριτήρια αξιολόγησης

Έλεγχος ορθότητας χημικών τύπων	20
Παρουσίαση σε εκτεταμένη μορφή των χημικών τύπων	40
Υπολογισμός αριθμού πρωτονίων	15
Σωστή χρησιμοποίηση απαιτούμενων τύπων δεδομένων, I/O σε αρχεία	10
Γενική εικόνα (ευανάγνωστος κώδικας, σχολιασμός, κλπ.)	15
ΣΥΝΟΛΟ	100

Καλή Επιτυχία!