

ΕΙΣΑΓΩΓΗ ΣΤΟΝ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟ (2006-07)

Άσκηση 2

Γράψτε ένα πρόγραμμα C (έστω ότι το πηγαίο αρχείο του ονομάζεται “nocomm.c”) το οποίο να διαβάζει από την είσοδο ένα πηγαίο πρόγραμμα C και να εκτυπώνει στην έξοδο το ίδιο πρόγραμμα, έχοντας αφαιρέσει τα σχόλια που περιέχει.

Στην C, μπορούμε να διατυπώσουμε σχόλια, εκτός από τον κλασικό τρόπο που γνωρίζετε, μεταξύ /* και */, και με έναν εναλλακτικό τρόπο. Οτιδήποτε βρίσκεται σε μία γραμμή μετά από το ζευγάρι χαρακτήρων //, και μέχρι το τέλος της γραμμής, θεωρείται σαν σχόλιο. Το πρόγραμμα που θα γράψετε να καλύπτει και τις δύο περιπτώσεις εισαγωγής σχολίων.¹

Προσέξτε ότι τα /* και // όταν βρίσκονται μέσα σε συμβολοσειρές "...” δεν ορίζουν την αρχή σχολίου. Επίσης, αν θέλουμε να βάλουμε μέσα σε μία συμβολοσειρά τον χαρακτήρα "\", τον γράφουμε σαν \".

Θα πρέπει ακόμα να ξέρετε ότι όταν στο τέλος μίας γραμμής, δηλαδή πριν τον χαρακτήρα \n, υπάρχει ο χαρακτήρας \, τότε θεωρείται ότι η επόμενη γραμμή είναι συνέχεια της τρέχουσας. Δηλαδή, είναι σαν να μην έχει γίνει καθόλου αλλαγή γραμμής. Λάβετε υπόψη σας αυτό το δεδομένο, γιατί επηρεάζει το πότε τελειώνει ένα σχόλιο που έχει εισαχθεί με //.

Τέλος, θεωρήστε ότι τα προγράμματα C που θα δίνονται στην είσοδο του προγράμματός σας δεν έχουν συντακτικά λάθη.

Μία ενδεικτική εκτέλεση φαίνεται στη συνέχεια:

```
% ls -l pr.c
-rw-----  1 ip      other      261 Nov 14 12:05 pr.c
% cat pr.c
#include <stdio.h>
/* A "test" program */
main() // The main function
{ /* Just a * in
    a comment */
    printf("%d\" - This /* is not */ a comment\n", 12/3);
    printf("Neither this // is a comment\n");
// And another comment \
    with a /* and a */ in it
}
% ./nocomm < pr.c
#include <stdio.h>

main()
{
    printf("%d\" - This /* is not */ a comment\n", 12/3);
    printf("Neither this // is a comment\n");
}
% ./nocomm < pr.c > prnocomm.c
```

¹Σημειώνεται ότι η δυνατότητα εισαγωγής σχολίων στην C με το ζευγάρι χαρακτήρων // υιοθετήθηκε στο τελευταίο πρότυπο (standard) της γλώσσας, αυτό του 1999. Πιθανώς κάποιοι μεταγλωττιστές που ακολουθούν ακόμα το προηγούμενο πρότυπο, αυτό του 1989, να μην δέχονται τέτοιου είδους σχόλια. Αυτό, βέβαια, το γεγονός δεν αφορά τις απαιτήσεις της άσκησης, που περιγράφονται με πολύ συγκεκριμένο τρόπο.

```
% ls -l prnocomm.c
-rw-----  1 ip      other      135 Nov 14 12:06 prnocomm.c
%
```

Μπορείτε να αντιγράψετε το δοκιμαστικό αρχείο που φαίνεται στην ενδεικτική εκτέλεση από το `~iphw/samples/pr.c`.

Τρέξτε το πρόγραμμά σας και με είσοδο από το αρχείο `/usr/local/doc/gtk/tests/testgtk.c` που μπορείτε να το βρείτε στους σταθμούς εργασίας Suns του Τμήματος. Τι μέγεθος έχει, σε πλήθος bytes, το αποτέλεσμα;

Η παράδοση της άσκησης αυτής συνίσταται στην υποβολή του πηγαίου αρχείου `nocomm.c` με διαδικασία που θα ανακοινωθεί σύντομα.

Υπόδειξη 1: Ένας καλός τρόπος για να αντιμετωπίσετε την άσκηση είναι να την υλοποιήσετε σταδιακά. Σε πρώτη φάση, ασχοληθείτε μόνο με τα σχόλια `/* ... */`, αγνοώντας το δεύτερο είδος σχολίων, την περίπτωση των συμβολοσειρών και το ενδεχόμενο κάποια γραμμή να τελειώνει με `\`. Αφού εξασφαλίσετε ότι το πρόγραμμά σας δουλεύει σωστά για τα σχόλια `/* ... */`, προσθέστε του και τη δυνατότητα να χειρίζεται και σχόλια που αρχίζουν με `//`. Στη συνέχεια, έχοντας ολοκληρώσει και αυτό το στάδιο, ασχοληθείτε με την περίπτωση των συμβολοσειρών. Και τέλος, λάβετε υπόψη σας και την περίπτωση κάποια γραμμή να συνεχίζει στην επόμενη, επειδή τελειώνει με `\`. Αν το σκεφτείτε λίγο προσεκτικά, θα δείτε ότι αυτό το τελευταίο ενδιαφέρει μόνο για γραμμές σχολίων που έχουν εισαχθεί με `//`, όπως έχει ήδη αναφερθεί προηγουμένως.

Υπόδειξη 2: Εκτός από τη γνωστή σας συνάρτηση `getchar`, που διαβάζει ένα χαρακτήρα από την είσοδο, είναι πολύ πιθανό να σας χρησιμεύσει και μία συνάρτηση της βιβλιοθήκης εισόδου/εξόδου, που δεν γνωρίζετε, η `ungetc`, για την αναίρεση της ανάγνωσης ενός χαρακτήρα από την είσοδο. Δηλαδή, αν διαβαστεί ο χαρακτήρας `ch` από την είσοδο και διαπιστωθεί ότι κακώς διαβάστηκε και πρέπει να επιστραφεί, ώστε να ξαναδιαβαστεί στη συνέχεια, αυτό επιτυγχάνεται με την κλήση της `ungetc(ch, stdin)`. Μην σας μπερδεύει εδώ αυτό το `stdin`. Θα εξηγηθεί αναλυτικά σε επόμενα μαθήματα. Προς το παρόν, θεωρήστε ότι είναι ένα όνομα για την είσοδο από την οποία διαβάζονται τα δεδομένα του προγράμματός σας. Πάντως, θα πρέπει να διευκρινισθεί ότι χρήση της συνάρτησης `ungetc` είναι επιτρεπτή, αλλά σε καμία περίπτωση δεν είναι υποχρεωτική. Η άσκηση μπορεί να λυθεί και χωρίς την χρήση της `ungetc`.

Υπόδειξη 3: Αν θυμηθείτε τα περί Turing μηχανών που έχετε ακούσει στην *Εισαγωγή της Επιστήμης της Πληροφορικής και των Τηλεπικοινωνιών*, θα έχετε στη διάθεσή σας ένα πολύ καλό εργαλείο για την αντιμετώπιση της άσκησης.