

## Μεταγλωττιστές 2020

### Προγραμματιστική Εργασία #2

**Ονοματεπώνυμο:** Κουμπουλίδης Παντελής

**AM:** Π2016176

Αρχικά ξεκίνησα κάνοντας όλες τις απαλοιφές, διότι θεώρησα ότι ήταν το πιο δύσκολο κομμάτι, έπειτα συνέχισα με τις εξαγωγές των απαραίτητων στοιχείων και τέλος αφαίρεσα όλα τα κενά. Η έξοδος μου, όσο δοκίμαζα το πρόγραμμα ήταν σε ένα αρχείο txt, για να υπάρχει άμεση σύγκριση με το πρωτότυπο. Δυσκολεύτηκα αρκετά στην δημιουργία της callback συνάρτησης, αλλά συμβουλευόντας τις σημειώσεις του μαθήματος έγινε κατανοητή.

**Re.compile:** μηχανή ταιριάσματος

**Finditer:** επιστρέφει πολλαπλά ταιριάσματα σε αντίθεση με την search που επιστρέφει το πρώτο ταίριασμα.

**Sub(a,b):** Η εντολή sub, βρίσκει τις εκφράσεις της μηχανής ταιριάσματος και τις αντικαθιστά με a στο κείμενο b.

**Callback Συνάρτηση:** Για την πιο ευέλικτη λειτουργία του προγράμματος χρησιμοποιείται αυτή η συνάρτηση. Βρίσκει τους χαρακτήρες που έχουν δοθεί από τη μηχανή ταιριάσματος και ανάλογα τον χαρακτήρα κάνει match στο λεξικό. Έπειτα η συνάρτηση επιστρέφει τον ανάλογο χαρακτήρα ο οποίος αντικατασταθείτε με την εντολή sub.

**<title>(.+?)</title>** : Επιστρέφει στο group(1) το περιεχόμενο του title

**'<a(.+?)</a>'** : Επιστρέφει στο group(1) το περιεχόμενο των Υπερσυνδέσμων

**(<script.+?</script>)|(<style.+?</style>)|(<!--.+?-->), re.DOTALL** : Αρχικά δεν είχα χρησιμοποιήσει το DOTALL, αλλά παρατήρησα πως στην έξοδο του προγράμματος υπήρχε ένα κομμάτι από το περιεχόμενο ενός script, το οποίο ήταν multiline.

**'<[>]+>'** : Οτιδήποτε βρίσκεται μεταξύ < >

Οι μόνες πηγές που χρησιμοποίησα ήταν οι σημειώσεις του μαθήματος, που ήταν στις περισσότερες περιπτώσεις αρκετές. Για τη δημιουργία της callback βοήθησαν και μερικά παραδείγματα που βρήκα στο stackoverflow

