

Κέρκυρα, 12/5/2017

# ΕΡΓΑΣΙΑ ΣΤΟΥΣ ΜΕΤΑΓΛΩΤΤΙΣΤΕΣ

ΤΟΥ

ΓΙΑΝΝΙΟΥ ΑΝΤΩΝΙΟΥ

A.M. Π2013153

Θέμα: "Προγραμματιστική Εργασία #1"

Η εργασία αφορά στην επεξεργασία αρχείου .srt, στην εύρεση της χρονικής πληροφορίας που υπάρχει στο αρχείο και στην μετατροπή της, προσθέτοντας ένα χρονικό offset πάνω σε αυτήν.

Ο κώδικας του προγράμματος είναι γραμμένος σε γλώσσα Python.

Ο χρήστης δίνει σαν ορίσματα εισόδου ένα αρχείο σε μορφή srt και μια τιμή offset, int ή float. Ο χειρισμός των ορισμάτων γίνεται με την χρήση του module argparse.

Η εύρεση της χρονικής πληροφορίας που υπάρχει στο αρχείο γίνεται με την χρήση regular expressions. Χρησιμοποιώ κατάλληλο pattern "**(([0-9]{2}):([0-9]{2}):([0-9]{2},[0-9]{3})) --> ([0-9]{2}):([0-9]{2}):([0-9]{2},[0-9]{3}))**" για να αναζητήσω με την μέθοδο findeiter την χρονική πληροφορία.

Το pattern αυτό αποτελείται από την πρώτη χρονική πληροφορία και την δεύτερη χρονική πληροφορία και αντιστοιχείται στην παρακάτω μορφή

**hour1:min1:sec1 --> hour2:min2:sec2**

(hour1,min1,sec1 = πρώτη χρονική πληροφορία)

(hour2,min2,sec2 = δεύτερη χρονική πληροφορία)

Με την χρήση των κατάλληλων groups πάνω στο αντικείμενο finditer απομονώνω τις αντίστοιχες πληροφορίες και τις αντιστοιχώ στις μεταβλητές hour1,min1,sec1,hour2,min2,sec2. Στην μεταβλητή off αντιστοιχώ το όρισμα εισόδου offset.

Έπειτα υπολογίζω το συνολικό άθροισμα των hour1,min1,sec1 (αφού τα έχω μετατρέψει σε sec πρώτα) και προσθέτω σε αυτό το offset που έχει δώσει ο χρήστης. Ομοίως κάνω και για τις hour2,min2,sec2.

Όμως για να αποφύγω προβλήματα λόγω ακρίβειας που προκύπτουν από την χρήση float numbers πολλαπλασιάζω όλες τις μεταβλητές επί 1000 (ουσιαστικά τα μετατρέπω σε sec) και έπειτα κάνω τις μετατροπές και τα αθροίσματα που χρειάζονται. Ύστερα, αφού προσθέσω στα συνολικά αθροίσματα (totalsec1,totalsec2) το offset, κάνω την αντίστροφη εργασία υπολογίζοντας πάλι τα νέα hour1,min1,sec1,hour2,min2,sec2 και τέλος, φυσικά τα διαιρώ με το 1000.

Έπειτα χρησιμοποιώντας την συνάρτηση stringReconstruct (η οποία δέχεται σαν όρισμα ένα αριθμό (string) και έναν αριθμό int (length) και επιστρέφει ένα string (string)) μετατρέπω την κάθε χρονική μεταβλητή σε string που έχει μορφή δυο ψηφίων. Η τιμή length χρησιμοποιείται για να ελέγξω αν πρόκειται για μονό ψηφίο όποτε προσθέτω το 0 μπροστά ή όχι (το length είναι 1 για hours και minutes και 5 για seconds).

Τέλος, αντικαθιστώ τις νέες τιμές των hour1,min1,sec1,hour2,min2,sec2 στα αντίστοιχα groups και τα εκτυπώνω στο stdout.

Επειδή η χρονική πληροφορία στο src αρχείο για τα δευτερόλεπτα είναι της μορφής xx,xxx έχει δηλαδή κόμμα, αντί για τελεία, χρησιμοποιώ επίσης regular expressions για να αντικαταστήσω το κόμμα με την τελεία (για να κάνω μετατροπή από string σε float) και το αντίστροφο (για την μετατροπή από float σε string), πριν την τελική αντικατάσταση των νέων τιμών.