

Μαρία Μαλακοπούλου, Π2014182

Περιγραφή Κώδικα:

Στην αρχή παίρνουμε την λίστα strings που επιστρέφει η κανονική έκφραση και την κάνουμε κανονικό string με το join και σαν ενδιάμεσο χαρακτήρα το κενό (' '). Μετά διαχωρίζουμε το string με την μέθοδο split, χρησιμοποιούμε ανάλογα τον χαρακτήρα ":" ή "," για τον διαχωρισμό για να εξάγουμε απο το string τις ώρες, λεπτα, κτλ.

Μετα παίρνουμε τις ωρες και τις μετατρέπουμε σε δευτερόλεπτα, κανουμε το ίδιο και για τα λεπτά. Τελικά έχουμε τον χρόνο σαν δευτερόλεπτα (float) και σε αυτά προσθέτουμε το offset που δίνει ο χρήστης. Τελικά ξαναμετατρέπουμε το αποτέλεσμα απο δευτερόλεπτα σε ώρες, λεπτά και δευτερόλεπτα και κάνουμε print στο stdout.

Περιγραφή κανονικής έκφρασης:

xronoseira = re.findall("(\\d\\d:\\d\\d:\\d\\d,\\d\\d\\d --> \\d\\d:\\d\\d:\\d\\d,\\d\\d\\d)", line)

Η κανονική έκφραση βρίσκει εαν υπάρχει το πρότυπο της στο αρχείο (κατά σειρά που δίνεται) και επιστρέφει τι βρήκε στην μεταβλητή "xronoseira".

Για παράδειγμα μια κανονική έκφραση τύπου \\d\\d θα βρίσκει δύο ψηφία σε σειρά και θα τα επιστρέφει σαν στοιχείο λίστας στην μεταβλητή. Η δική μου θα "ταιριάζει" μόνο την χρονική πληροφορία κάθε διαφορετικής σειράς υποτίτλων.

Πηγές:

https://pyformat.info/#number_padding

<https://docs.python.org/3/library/string.html>

Βιβλιοθήκες python για κανονικές εκφράσεις