Tutorial 06 - CS384 - Web Series Renamer & Regular Expression

Dr. Mayank Agarwal

Assignment Given: 17th Oct 2021,
Deadline 23rd Oct 2021, 23:59
Submission: GitHub

Things to be kept in mind

1. Program will be checked for plagiarism.

2. Using Regex is mandatory

A sample set of **srt+mp4** is given to you in separate folders. You need to rename **mp4 + srt files**. You know that in VLC/any media play will auto include the srt if the filename and srt name is same. You need to ask 3 inputs from the users

1. Initially make a menu based program that will ask which web series to rename. The sample python code I have already given along with the variables.

2. Season Number Padding (take int input): The padding given to the season number. A padding of three digits means if the series have 20 seasons then it will be renamed as season 001, 002, 003..., 019, 020.

3. Episode Number Padding (take int input): For example, a padding of two means if the series have 30 episodes then it will be renamed as episode 001, 002, ..., 030

There are 3 folders that you need to rename:

1. Breaking Bad

2. Game of Thrones

3. Lucifer

Lets say the user provides these inputs:

1. Main Title of the **Web Series**: 2

2. Season Number Padding (take int input): 1

3. Episode Number Padding (take int input): 2

Original:: 'Game of Thrones - 8x05 - The Bells.WEB.REPACK.MEMENTO.en.srt"
New: "Game of Thrones - Season 8 Episode 05 - The Bells.mp4"

   Also look, there is an interesting case: Look at one of the names:
   "Game of Thrones - 8x05 - The Bells.WEB.REPACK.MEMENTO.en.srt"
Original:: 'Game of Thrones - 8x05 - The Bells.WEB.REPACK.MEMENTO.en.srt"
New: "Game of Thrones - Season 8 Episode 05 - The Bells.mp4"

   Here you can see that ".WEB.REPACK.MEMENTO.en" is the string that is not required. Usually every srts have common strings towards the end that needs to be identified manually and stripped. So in case of Game of Thrones, you need to find out manually such non-useful string and strip them in the final code. A **recommended** approach will be **regex**. Like **splitting** on the **WEB** and other such string (e.g., **720p**).
   Task: Based on the received input, the filename should be *seriesname season number (depending on the padding) episode number (depening on the padding) + episode name ()if any)*.
   Following series needs to be renamed (srt+mp4) via the os module.

1. Breaking Bad

2. Game of Thrones

3. Lucifer

   Spaces should be strictly one between characters and numbers.