

Dialogue search Database Project

CS315A

Dhawal Upadhyay (14218), Kaushal Agrawal (14313), Vibhor Goyal (14793)

Problem Description

We often come across a dialogue of a tv show/movie but are unable to recall the episode/movie. Googling the query is seldom helpful.

The project aims to create a webpage where one can search for a query dialogue/a list of random words (s)he can remember occurring in the show, and will be answered which episode it occurs in. If it does not match any episode, the closest match shall be reported.

Solution Outline

We will create tables where each row contains episode name, string of words, and its count.

We will have separate tables for strings of length 1,2,...upto 10.

Sample dialogue and corresponding tables –

Charlie: How lucky I am! Lucky, lucky, lucky...

Table 1			Table 2			Table 3		
Episode	String	Count	Episode	String	Count	Episode	String	Count
s06e09	how	1	s06e09	how lucky	1	s06e09	how lucky I	1
s06e09	lucky	4	s06e09	lucky I	1	s06e09	lucky I am	1
s06e09	I	1	s06e09	I am	1	s06e09	lucky lucky lucky	1
s06e09	am	1	s06e09	lucky lucky	2			

User inputs query dialogues as a set of strings. We query the strings in the corresponding database according to the length of strings. We output the best match (i.e. the episode which matches maximum number of strings). The best match algorithm might get optimized according to weightage.

Final Project

We will build a website where there will be fields to input strings, and the tv series name (if known). We will use SQL to answer the queries.

Additionally – We can add a file upload system where users can upload a file for a missing episode to be entered in the database.