

Josh Sapers and Subhash Gubba
3/3/14
Software Design
Paul Ruvolo

Project Overview:

Create a program that plays the “Wikiracing game” which navigates through Wikipedia given a starting article with the goal to find a given ending article. It is to only use hyperlinks on the page in order to get to the next article.

Implementation:

Using BeautifulSoup4 we parsed all the hyperlinks into a dictionary, and then we checked the amount of words in common that each had with the goal article, and we also checked to see if the link page contained the name of the goal. If it contained the name, we had the program immediately recursively repeat the process for the new link. If it did not appear we used the link that had the most words in common. In order to circumvent infinite loops, we made sure to cross check with a list of previously visited articles.

Test Cases / Examples:

Below are the outputs of some tests we used to determine the functionality of our code. Included are the paths that it took from the first article to the end and the time it took to get there.

George_W._Bush,Case_sensitivity,Poland,The_Holocaust,List_of_conspiracy_theories#Antisemitic_conspiracy_theories,Nuclear_industry,Mineral,Dietary_mineral,Food_chemistry,Sugar,Carbohydrates 2:52

Satellite,Portugal,Globalization,United_Kingdom,Mary_Shelley,Europe,Wolf,Carpet,France,China,Australia,Guernsey,Donkey 3:09

Zelda, Manhattan,Al_Capone,Mae_Capone,The_St._Valentine
%27s_Day_Massacre_(film),Joe_Aiello,John_Scalise,List_of_Mafia_crime_families,Canada,History_of_Canada_(1982%E2%80%93),1964_in_Canada,1989_in_Canada,1771_in_Canada,Great_U
pheaval,Fort_Rouill
%C3%A9,List_of_towns_in_New_France,Code_Noir,French_colonial_empire,Intendant_of_New_
France,Beaver_Wars,Franco-Turkish_War,1980_Turkish_coup_d
%27%C3%A9tat,1958_Lebanon_crisis,Sinai_insurgency,List_of_modern_conflicts_in_the_Middle_
East,2007_Lebanon_conflict,Iranian_Revolution,Iranian_Cultural_Continent,1987_Mecca_incide
nt,Shi%27a
%E2%80%93Sunni_relations,Basra,Kanpur,Kanpur#Development_of_industries,Meerut,Indo-
Pakistani_wars_and_conflicts,Instrument_of_Surrender_(1971),Bengali_nationalism,English_natio
nalism,Traditional_Unionist_Voice,Northern_Ireland,Ulster_loyalism,Glenanne_gang,Deal_barrack
s_bombing,Birmingham_pub_bombings,Ulster_Defence_Association,John_White_(loyalist),Davy_
Fogel,Robert_Ford_(British_Army_officer),World_War_II,Italian_resistance_movement,Bologna,B
ologna_(disambiguation),Bologna_massacre,Massacre 2:51

**The code takes really long when given common words which aren't hyperlinked often in articles.

Reflection:

We could have split writing of code better, as well as given ourselves more time to write it. That being said, we did scope our project well for the time given, as we have completed our goal. We did lots of unit testing to iterate upon our code more, which proved to be effective. Overall, it was fun.