Matthew Ulmer

**Degrees of Wikipedia**

**Algorithm**

For this assignment we are given two text documents. The first contains Wikipedia link titles on separate rows and the second contains links to other Wikipedia articles, represented by the row the link title appears on. The goal is to find the shortest possible path between two articles and to do so I used Dijkstra’s algorithm. But before using Dijkstra’s algorithm, it is necessary to populate the graph. In order to do this I iterate through the links and set directed edges from each article to the article they point to. It is important that they are directed edges as it is not necessarily true that there will be a link to go back. Afterwards, Dijkstra’s algorithm is performed and the path array is returned. In this path array, the value at a specific index represents the next index to jump to in the path. If a value at a particular index is null, that means that the link represented by that index value cannot reach the target link as they are in disjoint sets. Because of this, it is only necessary for me to check if the value in path at index source is equal to null once. If it is equal to null, that means that there is no path from the source to target. Otherwise, I just keep updating the value of source with the value at position path[source] until the target is reached.

**Format**

* Start of first line is the source
* End of last line is the target
* In the middle of two link titles there is: (cur\_source\_index -> path[cur\_source\_index])

**1.**

Invincible\_(Michael\_Jackson\_album) (651->4328) Venezuela

Venezuela (4328->4575) World\_War\_I

World\_War\_I (4575->2571) Peter\_O'Toole

Peter\_O'Toole (2571->1260) Laurence\_Olivier\_Awards

**2.**

Halle\_Berry (320->501) Hugh\_Jackman

Hugh\_Jackman (501->935) John\_Travolta

John\_Travolta (935->1958) Moment\_by\_Moment

**3.**

Jennifer\_Dunn (836->1019) July\_29

July\_29 (1019->799) January\_7

January\_7 (799->2183) Nicolas\_Cage

Nicolas\_Cage (2183->7) Forrest\_Gump\_(film)

**4.**

President\_of\_Argentina (2705->1782) May\_25

May\_25 (1782->2342) October\_28

October\_28 (2342->980) Julia\_Roberts

Julia\_Roberts (980->3928) The\_Player

**5.**

Peter\_Twinn (2578->1486) London

London (1486->1500) Los\_Angeles,\_California

Los\_Angeles,\_California (1500->1502) Los\_Angeles\_Daily\_News

**6.**

The\_Peninsula\_Manila (3923->4262) United\_States

United\_States (4262->2320) October\_17

October\_17 (2320->2524) Paxton\_Whitehead

**7.**

Laura\_Branigan (1255->3203) Saturday\_Night\_Live

Saturday\_Night\_Live (3203->1887) Mike\_Myers\_(actor)

Mike\_Myers\_(actor) (1887->3116) Rupert\_Everett

**8.**

Guy\_Bolton (286->3310) September\_6

September\_6 (3310->2620) Pink\_Floyd

Pink\_Floyd (2620->4217) UK\_Singles\_Chart

**9.**

Thriller\_(genre) (4037->3557) Stephen\_King

Stephen\_King (3557->995) July\_16

July\_16 (995->1032) June\_10

June\_10 (1032->2760) Prince\_Hashim\_bin\_Al\_Hussein

**10.**

Varsity\_Blues\_(film) (4322->2481) Paramount\_Pictures

Paramount\_Pictures (2481->4342) Viacom