Breath First search

Start at the tree root (or some arbitrary node of a graph, sometimes referred to as a 'search key'[1]) and explores the neighbor nodes first, before moving to the next level neighbors.

Logic behind my program:
create 5 different depth lists. while the length of the main_link list is less than 1000.
extract a link which has not been visited yet crawl to find the links.

step 1:

when a page is visited, extract all the links in the page that match our criteria and add it at its list
list L=[l1,l2,l3].

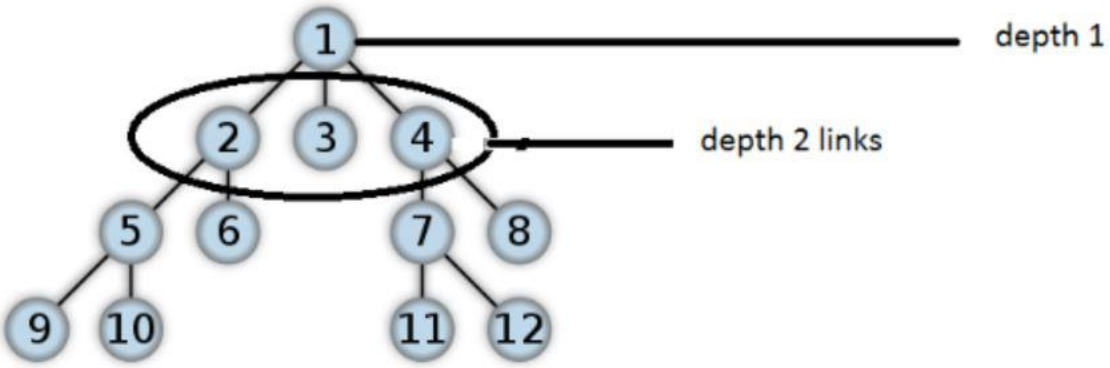
step 2:

check for the length of the main_link list . crawl the sequential links in the list. 1. L1
2. L2

step 3:

repeat the steps till we extract 1000 unique links.

Links obtained: 1000



[google bfs page]

Links visited: 1->2->3->4…

https://en.wikipedia.org/wiki/Sustainable_energy
https://en.wikipedia.org/wiki/Greenpower
https://en.wikipedia.org/wiki/Energy_conservation
https://en.wikipedia.org/wiki/Cogeneration
https://en.wikipedia.org/wiki/Efficient_energy_use
https://en.wikipedia.org/wiki/Green_building

Depth First Crawling:

One starts at the root (selecting some arbitrary node as the root in the case of a graph) and explores as far as possible along each branch before backtracking.

In Depth First Crawling, we do the following:

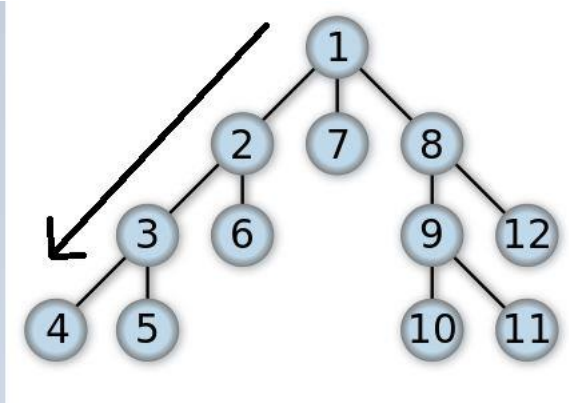
Step 1: When a page is visited, get all the links in page that match our criteria and add it to list L = [L1, L2, L3] and increment Depth (D =1) by 1.

Step 2: Go to L1 and crawl all the links and add to another list S = [S1, S2, S3] and increment Depth by 1 (now D = 2).

Step 3: Go to Step 4 if depth reaches 6 else Go to Step 2 and do the same with new links from S1.

Step 4: If Depth has reached 6 for all the links or if 1000 unique links are crawled then stop Crawling.

Total number of links obtained: 679



Links visited: 1->2->3->4->5….
https://en.wikipedia.org/wiki/Sustainable_energy
https://en.wikipedia.org/wiki/Passive_solar_building_design
https://en.wikipedia.org/wiki/Solar_energy
https://en.wikipedia.org/wiki/Solar_Energy_(journal)
https://en.wikipedia.org/wiki/Solar_heating