Assignment No: 07

Name: Omkar Kulkarni
PRN: B25CE2012
Class: SY-I
Batch: C

Problem Statement:

Web Crawling:
A. **Breadth First Search (BFS):**
A web crawler uses BFS to visit web pages systematically, starting from a seed URL and exploring links level by level.
Write a program to simulate the indexing of web pages for a search engine using BFS.

Program:

```
#include <stdio.h>
#define MAX 10
int queue[MAX], front = -1, rear = -1, visited[MAX] = {0},
adj[MAX][MAX];

void enqueue(int v){ if(rear==MAX-1) printf("Queue Overflow\n");
else { if(front==-1) front=0; queue[++rear]=v; } }

int dequeue(){ if(front==-1||front>rear) return -1; return
queue[front++]; }

void bfs(int start, int n){
    enqueue(start); visited[start]=1;
    printf("\nIndexing pages starting from: %d\n",start);
    while(front<=rear){
        int v=dequeue(); printf("Indexed page: %d\n",v);
        for(int i=0;i<n;i++)
            if(adj[v][i]==1 && !visited[i]){ enqueue(i);
visited[i]=1; }
    }
}

int main(){
```

```
    int n,e,from,to,start;
    printf("Enter number of web pages (max 10): "); scanf("%d",&n);
    printf("Enter number of hyperlinks: "); scanf("%d",&e);
    for(int i=0;i<n;i++) for(int j=0;j<n;j++) adj[i][j]=0;
    printf("Enter links (from to):\n");
    for(int i=0;i<e;i++){ scanf("%d %d",&from,&to); adj[from][to]=1;
}
    printf("Enter seed/start page number: "); scanf("%d",&start);
    bfs(start,n); return 0;
}
```

Output:

```
Enter number of web pages (max 10): 6
Enter number of hyperlinks: 5
Enter links (from to):
1 2
1 4
2 3
4 6
5 6
Enter seed/start page number: 1
Indexing pages starting from: 1
Indexed page: 1
Indexed page: 2
Indexed page: 4
Indexed page: 3
Indexed page: 6
```

B. Depth First Search (DFS):

Web crawlers use DFS to explore pages deeply before backtracking.
Write a program to simulate DFS web crawling.

Program:

```
#include <stdio.h>
#define MAX 10
int adj[MAX][MAX], visited[MAX];
```

```
void dfs(int v,int n){
    visited[v]=1; printf("\nIndexed page: %d",v);
    for(int i=0;i<n;i++) if(adj[v][i]==1 && !visited[i]) dfs(i,n);
}

int main(){
    int n,e,from,to,start;
    printf("Enter number of web pages (max 10): "); scanf("%d",&n);
    printf("Enter number of hyperlinks: "); scanf("%d",&e);
    for(int i=0;i<n;i++) for(int j=0;j<n;j++) adj[i][j]=0;
    printf("Enter links (from to):\n");
    for(int i=0;i<e;i++){ scanf("%d %d",&from,&to); adj[from][to]=1;
}
    printf("Enter seed(start) page number: "); scanf("%d",&start);
    printf("\nStarting DFS web crawling from page: %d\n",start);
    dfs(start,n); return 0;
}
```

Output:

```
Enter number of web pages (max 10): 6
Enter number of hyperlinks: 5
Enter links (from to):
1 2
1 4
2 3
4 6
5 6
Enter seed(start) page number: 1
Starting DFS web crawling from page: 1
Indexed page: 1
Indexed page: 2
Indexed page: 3
Indexed page: 4
Indexed page: 6
```
