Algorithms and Data Structures Graphs, Depth-/Breadth-first Search, Graph-Connectivity

Albert-Ludwigs-Universität Freiburg

Prof. Dr. Rolf Backofen

Bioinformatics Group / Department of Computer Science Algorithms and Data Structures, January 2019

UNI FREIBURG





Graphs

Introduction Implementation Application example





Graphs - Overview:

 Besides arrays, lists and trees the most common data structure

(Trees are a special type of graph)

- Representation of graphs in the computer
- Breadth-first search (BFS)
- Depth-first search (DFS)
- Connected components of a graph



Terminology:







- Each graph G = (V, E) consists of:
 - A set of vertices (nodes) $V = \{v_1, v_2, ...\}$
 - A set of edges (arcs) $E = \{e_1, e_2, ...\}$
- Each edge connects two vertices $(u, v \in V)$
 - Undirected edge: $e = \{u, v\}$ (set)
 - Directed edge: e = (u, v) (tuple)
- Self-loops are also possible: e = (u, u) or $e = \{u, u\}$



Weighted graph:





Each edge is marked with a real number named weight
 The weight is also named length or cost of the edge depending on the application





Example: Road network

- Intersections: vertices
- Roads: edges
- Travel time: costs of the edges



Figure: Map of Freiburg © OpenStreetMap





How to represent this graph computationally?

1 Adjacency matrix with space consumption $\Theta(|V|^2)$





Figure: Adjacency matrix





How to represent this graph computationally?

Adjacency list / fields with space consumption $\Theta(|V| + |E|)$ Each list item stores the target vertex and the cost of the edge



Figure: Weighted graph with |V| = 4, |E| = 6



Figure: Adjacency list



Graph: Arrangement

- Graph is fully defined through the adjacency matrix / list
- The arrangement is not relevant for visualisation of the graph



a

Graphs Implementation - Python

. . .



```
class Graph:
    def __init__(self):
        self.vertices = []
        self.edges = []
    def addVertice(self, vert):
        self.vertices.append(vert)
    def addEdge(self, fromVert, toVert, cost):
        self.edges.append( \
            (fromVert, toVert, cost))
```



Degree of a vertex: Directed graph: G = (V, E)



Figure: Vertex with in- / outdegree of 3 / 2

Indegree of a vertex u is the number of edge head ends adjacent to the vertex

$$\deg^+(\boldsymbol{u}) = |\{(\boldsymbol{v},\boldsymbol{u}): (\boldsymbol{v},\boldsymbol{u}) \in \boldsymbol{E}\}|$$

Outdegree of a vertex u is the number of edge tail ends adjacent to the vertex

$$\deg^{-}(\boldsymbol{u}) = |\{(\boldsymbol{u},\boldsymbol{v}): (\boldsymbol{u},\boldsymbol{v}) \in \boldsymbol{E}\}|$$

Graphs Degrees (Valency)



Degree of a vertex: Undirected graph: G = (V, E)



Figure: Vertex with degree of 4

Degree of a vertex u is the number of vertices adjacent to the vertex

$$\deg(\boldsymbol{u}) = |\{\{\boldsymbol{v}, \boldsymbol{u}\} : \{\boldsymbol{v}, \boldsymbol{u}\} \in \boldsymbol{E}\}|$$



Paths in a graph: G = (V, E)





Figure: Undirected path of length 3 P = (0,3,2,4) Figure: Directed path of length 3 P = (0,3,1,4)

A path of *G* is a sequence of edges $u_1, u_2, \ldots, u_i \in V$ with Undirected graph: $\{u_1, u_2\}, \{u_2, u_3\}, \ldots, \{u_{i-1}, u_i\} \in E$

Directed graph: $(u_1, u_2), (u_2, u_3), \dots, (u_{i-1}, u_i) \in E$







Paths in a graph: G = (V, E)





Figure: Directed path of length 3 P = (0,3,1,4) Figure: Weighted path with cost 6 P = (2,3,1)

- The length of a path is: (also costs of a path)
 - Without weights: number of edges taken
 - With weights: sum of weigths of edges taken





Shortest path in a graph: G = (V, E)



Figure: Shortest path from 0 to 2 with cost / distance d(0,2) = ?

The shortest path between two vertices u, v is the path P = (u, ..., v) with the shortest length d(u, v) or lowest costs





Shortest path in a graph: G = (V, E)



Figure: Shortest path from 0 to 2 with cost / distance d(0,2) = 6P = (0,1,4,3,2)

The shortest path between two vertices u, v is the path P = (u, ..., v) with the shortest length d(u, v) or lowest costs



Diameter of a graph: G = (V, E)





Figure: Diameter of graph is *d* =?

The diameter of a graph is the length / the costs of the longest shortest path



Diameter of a graph: G = (V, E)





Figure: Diameter of graph is d = 10, P = (3, 2, 5)

The diameter of a graph is the length / the costs of the longest shortest path

Graphs Connected Components





BURG

20/65

Figure: Three connected components

- Undirected graph:
 - All connected components are a partition of V

$V = V_1 \cup \cdots \cup V_k$

Two vertices u, v are in the same connected component if a path between u and v exists



Connected components: G = (V, E)

- Directed graph:
 - Named strongly connected components
 - Direction of edge has to be regarded
 - Not part of this lecture

UNI FREIBURG

Graph Exploration: (Informal definition)

- Let G = (V, E) be a graph and $s \in V$ a start vertex
- We visit each reachable vertex connected to s
- Breadth-first search: in order of the smallest distance to s
- Depth-first search: in order of the largest distance to s
- Not a problem on its own but is often used as subroutine of other algorithms
 - Searching of connected components
 - Flood fill in drawing programms

Breadth-First Search:

- 1 We start with all vertices unmarked and mark visited vertices
- 2 Mark the start vertex *s* (level 0)
- 3 Mark all unmarked connected vertices (level 1)
- 4 Mark all unmarked vertices connected to a level 1-vertex (level 2)
- 5 Iteratively mark reachable vertices for all levels
- 6 All connected nodes are now marked and in the same connected component as the start vertex *s*



Figure: spanning tree of a breadth-first search



Figure: spanning tree of a breadth-first search

IBURG



Figure: spanning tree of a breadth-first search

IBURG



Figure: spanning tree of a breadth-first search



Figure: spanning tree of a breadth-first search



Figure: spanning tree of a breadth-first search



Figure: spanning tree of a breadth-first search

BURG



Figure: spanning tree of a breadth-first search

Depth-First Search:

- We start with all vertices unmarked and mark visited vertices
- 2 Mark the start vertex s
- Pick an unmarked connected vertex and start a recursive depth-first search with the vertex as start vertex (continue on step 2)
- If no unmarked connected vertex exists go one vertex back and continue recursive search (reduce the recursion level by one)



33/65

Depth-first search:

- Search starts with long paths (searching with depth)
- Like breadth-first search marks all connected vertices
- If the graph is acyclic we get a topological sorting
 - Each newly visited vertex gets marked by an increasing number
 - The numbers increase with path length from the start vertex



Figure: spanning tree of a depth-first search



Figure: spanning tree of a depth-first search



Figure: spanning tree of a depth-first search



Figure: spanning tree of a depth-first search



Figure: spanning tree of a depth-first search



Figure: spanning tree of a depth-first search



Figure: spanning tree of a depth-first search

BURG



Figure: spanning tree of a depth-first search



Figure: spanning tree of a depth-first search

IBURG



Figure: spanning tree of a depth-first search

BURG

Graphs Why is this called Breadth- and Depth-First Search?



UNI FREIBURG

Runtime complexity:

- Constant costs for each visited vertex and edge
- We get a runtime complexity of $\Theta(|V'| + |E'|)$
- Let V' and E' be the reachable vertices and edges
- All vertices of V' are in the same connected component as our start vertex s
- This can only be improved by a constant factor

- Connected component labeling
- Counting of objects in an image







What is object, what is background?



Convert to black and white using threshold:

value = 255 if value > 100 else 0





Interpret image as graph:

- Each white pixel is a node
- Edges between adjacent pixels (normally 4 or 8 neighbors)
- Edges are not saved externally, algorithm works directly on array
- Breadth- / depth-first search find all connected components (particles)



- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1

UNI



- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels

UNI FREIBURG



- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

Find connected components:



- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

JRG

8

UNI FREIBURG



- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

Find connected components:



- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

JRG

8



- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

Find connected components:



- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

JRG

8

UNI FREIBURG



- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

Find connected components:



- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

JRG

8

UNI FREIBURG



- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

Find connected components:



- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 2

. . .

JRG

2



Result of connected component labeling:

			c	P			6 1			к		M	N	0	P	0		6	T	u i	v	w	3						c	D			GH		1.1	к	L.	M	N	0	P 1	0	8	A T	U	V	W
35	255	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0	0	0 1	á r	0	0	0	0	0			35	10	0	0	0	0	0 1	0 0	0	0	0	0	0	0	0	0 0	0 1	0 0	3 0	0	0	0
36	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0	0	0 1	5 1	0	0	0	0	0			36	0	0	0	0	0	0 1	0 0	0	0	0	0	0	0	0	0 0	0 1	0 0	3 0	0	0	0
37	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0	0	0 1	5 1	0	0	0	0	0			37	0	0	0	0	0	0 1	0 0	0	0	0	0	0	0	0	0 0	0 1	0 0	3 0	0	0	0
30	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0	0	0 1	5 1	0	0	0	0	0			38	0	0	0	0	0	0 0	0 0	0	0	0	0	0	0	0	0 0	5 1	0 0	3 0	0	0	0
39	0	0	0	0	0	0	0 0	0 25	5.25	5 255	5.0	0	0	0	0	0	0 1	5 1	0	0	0	0	0			39	0	0	0	0	0	0 1	0 0	13	13	13	0	0	0	0	0 0	5 1	0 0	0 6	0	0	0
40	0	0	0	0	0	0 2	55 25	55 25	525	5 255	5 255	0	0	0	0	0	0 1	3 1	0	0	0	0	0			40	0	0	0	0	0	0 1	3 12	3 13	13	13	13	0	0	0	0 0	0 0	0 0	3 0	0	0	0
41	0	0	0	0	0	255 2	55 21	55 25	5 25	5 255	\$ 255	255	255	0	0	0	0	3 1	0	0	0	0	0			41	0	0	0	0	0 1	13 1	3 12	1 13	13	13	13	13	13	0	0 0	0 1	0 0	0 6	0	0	0
42	0	0	0	0	255	255.2	55 21	55 25	525	5 255	\$ 255	255	255	0	0	0	0 1	5 6	0	0	0	0	0			42	0	0	0	0	13 1	13 1	3 12	1 11	13	13	13	13	13	0	0 0	5 1	0 0	0 6	0	0	0
43	0	0	0	0	255	255 2	55 25	55 25	525	5 255	5 2 5 5	255	255	0	0	0	0 1	3 1	0	0	0	0	0			43	0	0	0	0	13 1	13 1	3 12	1 13	13	13	13	13	13	0	0 0	0 1	0 0	3 0	0	0	0
44	0	0	0	255	255	255.2	55 21	55 25	5 25	5 255	5 2 5 5	255	255	0	0	0	0 1	3 1	0	0	0	0	0			44	0	0	0	13	13 1	13 1	3 12	3 13	13	13	13	13	13	0	0 0	0 1	0 0	3 0	0	0	0
45	0	0	0	255	255	255.2	55 25	55 25	5 25	5 255	5 2 5 5	255	255	0	0	0	0 1	5 1	0	0	0	0	0			45	0	0	0	13	13 1	3 1	3 12	13	13	13	13	13	13	0	0 0		0 0	0 6	0	0	0
46	0	0	0	255	255	255 2	55 25	55 25	525	5 255	5255	255	255	0	0	0	0 1	3 1	0	0	0	0	25			46	0	0	0	13	13 1	13 1	3 13	3 13	13	13	13	13	13	0	0 0	0 1	0 0	3 0	0	0	0 1
47	0	0	0	255	255	255.2	55 25	55 25	5 25	5 255	5 2 5 5	255	255	0	0	0	0 1	5 1	0	0	0 3	255	25			47	0	0	0	13	13 1	13 1	3 12	3 13	13	13	13	13	13	0	0 0	0 1	0 0	0 6	0	0	17 1
48	0	0	0	255	255:	255 2	55 25	55 25	5 25	5 255	5 2 5 5	255	0	0	0	0	0 1	3 1	0	0 2	55 2	255	25	_	~	48	0	0	0	13	13 1	13 1	3 13	3 13	13	13	13	13	0	0	0 0	0 1	0 0	3 0	0	17	17 1
49	0	0	0	255:	255:	255 2	55 25	55 25	525	5 255	\$ 255	255	0	0	0	0	0 1	3 1	0	0 2	55:	255	25		-	49	0	0	0	13	13 1	13 1	3 12	3 13	13	13	13	13	0	0	0 0	0 1	0 0	3 0	0	17	17 1
50	0	0	0	255	255	255 2	55 25	55 25	525	5 255	\$255	255	0	0	0	0	0 1	3 1	0 2	552	55 2	255	25			50	0	0	0	13	13 1	13 1	3 12	1 13	13	13	13	13	0	0	0 0	0 1	0 0	0 6	17	17	17 1
51	0	0	0	0	255:	255 2	55 25	55 25	525	5 255	5 2 5 5	0	0	0	0	0	0 1	3 21	552	552	55 2	255	25			51	0	0	0	0	13 1	13 1	3 12	1 13	13	13	13	0	0	0	0 0	0 1	0 0	3 17	17	17	17 1
52	0	0	0	0	0 3	255 2	55 21	55 25	525	5255	5 0	0	0	0	0	0	0 1	3 2'	552	552	55:	255	25			52	0	0	0	0	0 1	13 1	3 12	1 13	13	13	0	0	0	0	0 0	0 1	0 0	3 17	17	17	17 1
53	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0	0	0 2	55.2	552	552	55:	255	25			53	0	0	0	0	0	0 1	0 0	0	0	0	0	0	0	0	0 0	0 1	0 1	7 17	17	17	17 1
54	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0	0	0 2	55 24	552	552	55 2	255	25			54	0	0	0	0	0	0 1	0 0	0	0	0	0	0	0	0	0 0	0 1	0 1	7 17	17	17	17 1
55	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0	0 2	552	55.24	552	552	55 2	255	25			55	0	0	0	0	0	0 1	0 0	0	0	0	0	0	0	0	0 0	0 1	17 1	7 17	17	17	17 1
56	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0	0 2	552	55 2	552	552	55:	255	25			56	0	0	0	0	0	0 1	0 0	0	0	0	0	0	0	0	0 0	0 1	17 1	7 17	17	17	17 1
57	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0	0 2	552	55 24	552	552	55 2	255	25			57	0	0	0	0	0	0 1	0 0	0	0	0	0	0	0	0	0 0	0 1	17 1	7 17	17	17	17 1
58	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0	0 2	552	55.2	552	552	55 2	255	25			58	0	0	0	0	0	0 1	0 0	0	0	0	0	0	0	0	0 0	11	17 1	7 17	17	17	17 1
59	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0 ;	2552	552	55 21	552	552	55 2	255	25			59	0	0	0	0	0	0 1	0 0	0	0	0	0	0	0	0	0 1	7 1	17 1	7 17	17	17	17 1
60	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0 ;	2552	552	\$5.2	552	552	55 2	255	25			60	0	0	0	0	0	0 0	0 0	0	0	0	0	0	0	0	0 1	7 1	17 1	7 17	17	17	17 1
61	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0 ;	2552	552	55.21	552	552	55 2	255	25			61	0	0	0	0	0	0 0	0 0	0	0	0	0	0	0	0	0 1	7 1	17 1	7 17	17	17	17 1
62	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0 ;	2552	552	55 21	552	552	55 2	255	25			62	0	0	0	0	0	0 1	0 0	0	0	0	0	0	0	0	0 1	7 1	17 1	7 17	17	17	17 1
63	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0 ;	2552	552	\$5.2	552	552	55 2	255	25			63	0	0	0	0	0	0 1	0 0	0	0	0	0	0	0	0	0 1	7 1	17 1	7 17	17	17	17 1
64	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0	0 2	552	\$5.2	552	552	55 2	255	25			64	0	0	0	0	0	0 1	0 0	0	0	0	0	0	0	0	0 0	0 1	17 1	7 17	17	17	17 1
65	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0	0 2	552	55 25	552	552	55 2	255	25			65	0	0	0	0	0	0 1	0 0	0	0	0	0	0	0	0	0 0	0 1	17 1	7 17	17	17	17 1
66	0	0	0	0	0	0	0 0	0 0	0 0	0	0	0	0	0	0	0 2	552	\$5.25	552	552	55 2	255	25			66	0	0	0	0	0	0 1	0 0	0	0	0	0	0	0	0	0 0	0 1	17 1	7 17	17	17	17 1

Figure: Result: particle indices instead of intensities

Further Literature



General

 [CRL01] Thomas H. Cormen, Ronald L. Rivest, and Charles E. Leiserson.
 Introduction to Algorithms. MIT Press, Cambridge, Mass, 2001.
 [MS08] Kurt Mehlhorn and Peter Sanders. Algorithms and data structures, 2008.

https://people.mpi-inf.mpg.de/~mehlhorn/
ftp/Mehlhorn-Sanders-Toolbox.pdf.

Further Literature

Graph Search

[Wika] Breadth-first search

https://en.wikipedia.org/wiki/ Breadth-first_search

[Wikb] Depth-first search

https: //en.wikipedia.org/wiki/Depth-first_search

Graph Connectivity

[Wik] Connectivity (graph theory) https://en.wikipedia.org/wiki/Connectivity_ (graph_theory)