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



Albert-Ludwigs-Universität Freiburg

Prof. Dr. Rolf Backofen

Bioinformatics Group / Department of Computer Science Algorithms and Datastructures, January 2017

Structure



Graphs

Introduction Implementation Application example



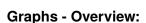
NE NE



Besides arrays, lists and trees the most common datastructure (Trees are a special type of graph)

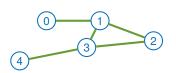
- Besides arrays, lists and trees the most common datastructure (Trees are a special type of graph)
- Representation of graphs in the computer

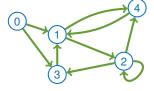
- Besides arrays, lists and trees the most common datastructure (Trees are a special type of graph)
- Representation of graphs in the computer
- Breadth first search (BFS)

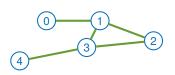


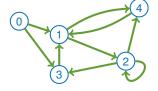
- Besides arrays, lists and trees the most common datastructure (Trees are a special type of graph)
- Representation of graphs in the computer
- Breadth first search (BFS)
- Depth first search (DFS)

- Besides arrays, lists and trees the most common datastructure (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

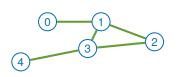


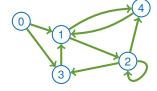




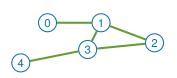


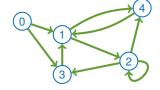
■ Each Graph G = (V, E) consists of:



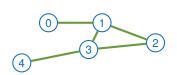


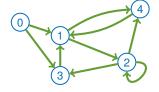
- Each Graph G = (V, E) consists of:
 - A set of vertices (nodes) $V = \{v_1, v_2, \dots\}$



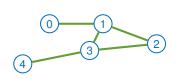


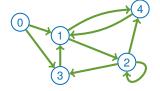
- Each Graph G = (V, E) consists of:
 - A set of vertices (nodes) $V = \{v_1, v_2, \dots\}$
 - A set of edges (arcs) $E = \{e_1, e_2, ...\}$



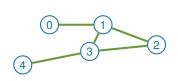


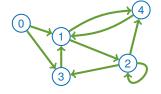
- Each Graph G = (V, E) consists of:
 - A set of vertices (nodes) $V = \{v_1, v_2, ...\}$
 - A set of edges (arcs) $E = \{e_1, e_2, ...\}$
- Each edge connects two vertices $(u, v \in V)$



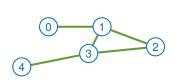


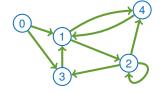
- Each Graph G = (V, E) consists of:
 - A set of vertices (nodes) $V = \{v_1, v_2, \dots\}$
 - A set of edges (arcs) $E = \{e_1, e_2, ...\}$
- Each edge connects two vertices $(u, v \in V)$
 - Undirected edge: $e = \{u, v\}$ (set)



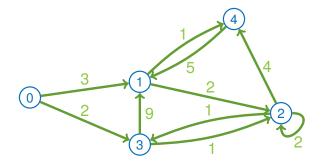


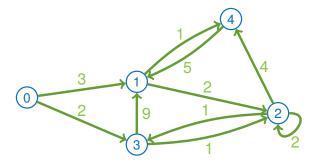
- Each Graph G = (V, E) consists of:
 - A set of vertices (nodes) $V = \{v_1, v_2, ...\}$
 - A set of edges (arcs) $E = \{e_1, e_2, \dots\}$
- Each edge connects two vertices $(u, v \in V)$
 - Undirected edge: $e = \{u, v\}$ (set)
 - Directed edge: e = (u, v) (tuple)



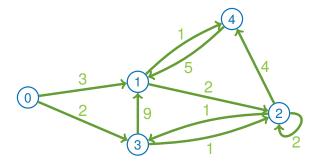


- Each Graph G = (V, E) consists of:
 - A set of vertices (nodes) $V = \{v_1, v_2, ...\}$
 - A set of edges (arcs) $E = \{e_1, e_2, ...\}$
- Each edge connects two vertices $(u, v \in V)$
 - Undirected edge: $e = \{u, v\}$ (set)
 - Directed edge: e = (u, v) (tuple)
- Self-loops are also possible: e =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





Each edge is marked with a real number named weight



- 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

Graphs Introduction



Example: Road network



Intersections:

vertices



Intersections:

vertices

■ Roads: edges

Intersections:

vertices

■ Roads: edges

Travel time:

costs of the edges

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



Figure: Map of Freiburg © OpenStreetMap

Structure



Graphs

Introduction

Implementation

Application example

Graphs Implementation



How to represent this graph computationally?



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

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

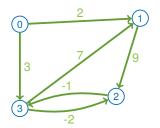


Figure: Weighted graph with

$$|V| = 4$$
, $|E| = 6$

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

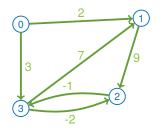


Figure: Weighted graph with |V| = 4, |E| = 6

	end-vertice			
	0	1	2	3
<u>0</u>		2		3
ert (1)			9	
start-vertice				-1
sta ③		7	-2	

Figure: Adjacency matrix

Graphs

Implementation



How to represent this graph computationally?

Graphs

Implementation



How to represent this graph computationally?

2 Adjacency list / fields with space consumption $\Theta(|V| + |E|)$

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

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

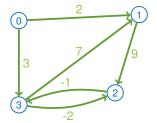


Figure: Weighted graph with |V| = 4, |E| = 6

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

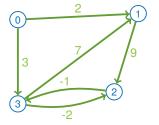


Figure: Weighted graph with |V| = 4, |E| = 6

<u>0</u> <u>8</u>	1, 2	3, 3
start-verti	2, 9	
± 2	3, -1	
sta ③	1, 7	2, -2

Figure: Adjacency list

Graphs Implementation

IBURG

NE NE

Graphs

Implementation



Graph: Arrangement

■ Graph is fully defined through the adjacency matrix / list

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

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

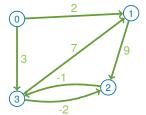


Figure: Weighted graph with

$$|V| = 4$$
, $|E| = 6$

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

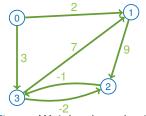


Figure: Weighted graph with |V| = 4, |E| = 6

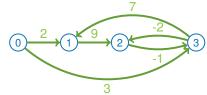


Figure: Same graph ordered by number - outer planar graph



```
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))
```

Graphs Degrees (Valency)

BURG

FREIB

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

Graphs Degrees (Valency)

NI EIBURG

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



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

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

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

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

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

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

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

Graphs Degrees (Valency)



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

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



Figure: Vertex with degree of 4

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

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

Graphs Paths





Paths

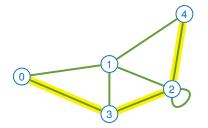


Figure: Undirected path of length 3 P = (0,3,2,4)

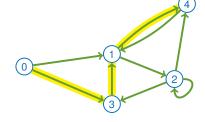


Figure: Directed path of length 3 P = (0,3,1,4)

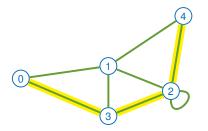


Figure: Undirected path of length 3 P = (0,3,2,4)

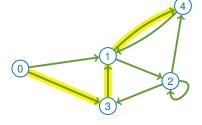


Figure: Directed path of length 3 P = (0,3,1,4)

■ A path of G is a sequence of edges $u_1, u_2, ..., u_i \in V$ with

Paths

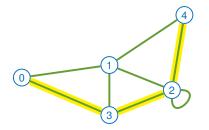


Figure: Undirected path of length 3 P = (0, 3, 2, 4)

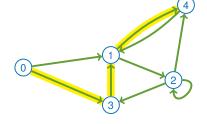


Figure: Directed path of length 3 P = (0, 3, 1, 4)

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



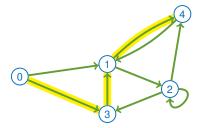


Figure: Directed path of length 3 P = (0,3,1,4)

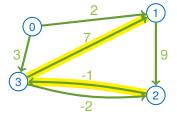


Figure: Weighted path with cost 6 P = (2,3,1)



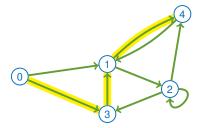


Figure: Directed path of length 3 P = (0,3,1,4)

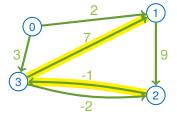


Figure: Weighted path with cost 6 P = (2,3,1)



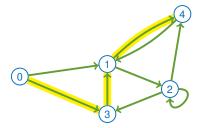


Figure: Directed path of length 3 P = (0,3,1,4)

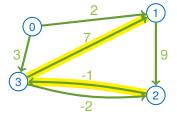


Figure: Weighted path with cost 6 P = (2,3,1)

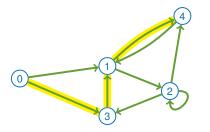
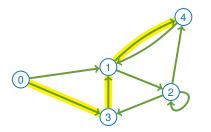


Figure: Directed path of length 3 P = (0, 3, 1, 4)

Figure: Weighted path with cost 6 P = (2, 3, 1)

The length of a path is: (also costs of a path)

Paths



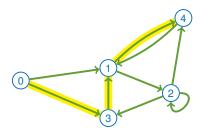
3 -1 -2 2

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

Paths



3 -1 2

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

Graphs Paths



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



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

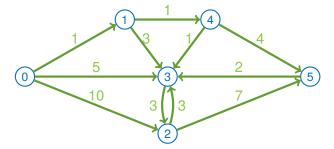


Figure: Shortest path from 0 to 2 with cost / distance d(0,2) = ?

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

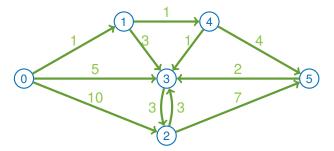


Figure: Shortest path from 0 to 2 with cost / distance d(0,2) = ?

 \blacksquare 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)

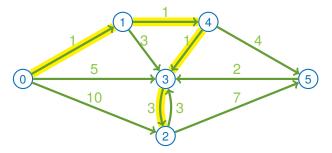


Figure: Shortest path from 0 to 2 with cost / distance d(0,2) = 6P = (0,1,4,3,2)

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

Graphs Paths



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



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

$$d = \max_{u,v \in V} d(u,v)$$

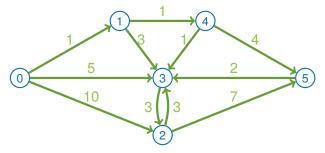


Figure: Diameter of graph is d = ?

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

$$d = \max_{u,v \in V} d(u,v)$$

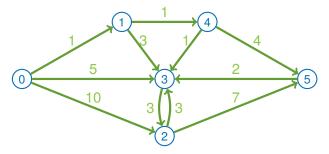


Figure: Diameter of graph is d = ?

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



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

$$d = \max_{u,v \in V} d(u,v)$$

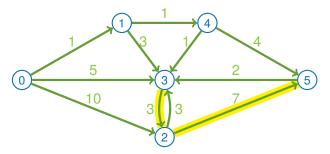


Figure: Diameter of graph is d = 10, P = (3, 2, 5)

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

Graphs Connected Components

EIBURG

Connected components: G = (V, E)

Connected components: G = (V, E)

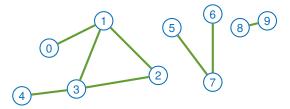


Figure: Three connected components

Undirected graph:

Connected components: G = (V, E)

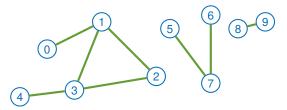


Figure: Three connected components

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

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

Connected components: G = (V, E)

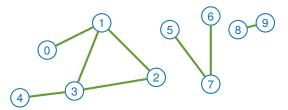


Figure: Three connected components

- 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

Graphs Connected Components



Connected components: G = (V, E)

Graphs Connected Components



N N N

Connected components: G = (V, E)

■ Directed graph:

Connected components: G = (V, E)

- Directed graph:
 - Named strongly connected components

Connected components: G = (V, E)

- Directed graph:
 - Named strongly connected components
 - Direction of edge has to be regarded



Connected components: G = (V, E)

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

Graphs

Connected Components - Graph Exploration



Graph Exploration: (Informal definition)

Graphs

Connected Components - Graph Exploration



Graph Exploration: (Informal definition)

■ Let G = (V, E) be a graph and $s \in V$ a start vertex

- Let G = (V, E) be a graph and $s \in V$ a start vertex
- We visit each reachable vertex connected to s

- 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

- 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

- 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

- 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

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

We start with all vertices unmarked and mark visited vertices

- We start with all vertices unmarked and mark visited vertices
- 2 Mark the start vertex s (level 0)

- We start with all vertices unmarked and mark visited vertices
- Mark the start vertex s (level 0)
- Mark all unmarked connected vertices (level 1)

- We start with all vertices unmarked and mark visited vertices
- Mark the start vertex s (level 0)
- Mark all unmarked connected vertices (level 1)
- Mark all unmarked vertices connected to a level 1-vertex (level 2)

- We start with all vertices unmarked and mark visited vertices
- Mark the start vertex s (level 0)
- Mark all unmarked connected vertices (level 1)
- Mark all unmarked vertices connected to a level 1-vertex (level 2)
- 5 Iteratively mark reachable vertices for all levels

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

Graphs

Connected Components - Breadth-First Search



■ The marked vertices create a "spanning tree" containing all reachable nodes

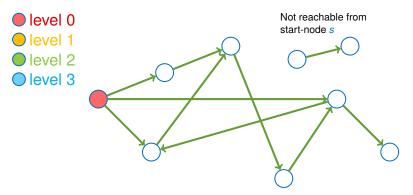


Figure: spanning tree of a breadth-first search

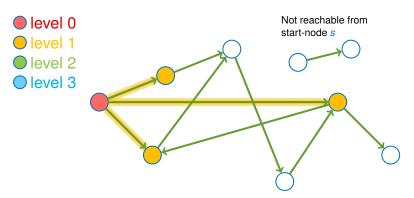


Figure: spanning tree of a breadth-first search

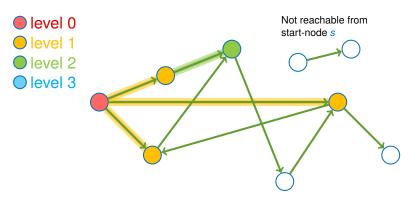


Figure: spanning tree of a breadth-first search

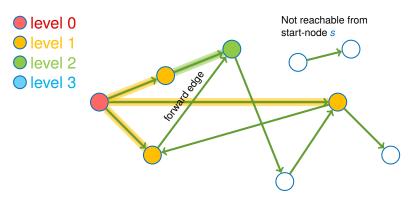


Figure: spanning tree of a breadth-first search

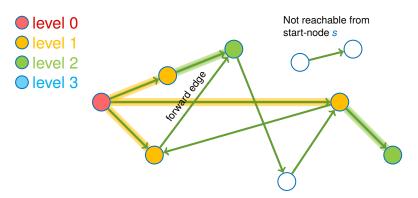


Figure: spanning tree of a breadth-first search

■ The marked vertices create a "spanning tree" containing all reachable nodes

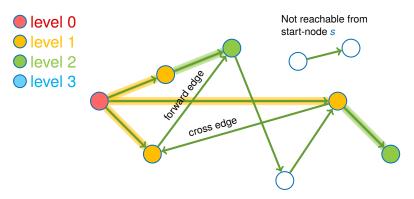


Figure: spanning tree of a breadth-first search

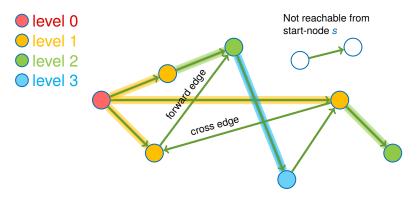


Figure: spanning tree of a breadth-first search

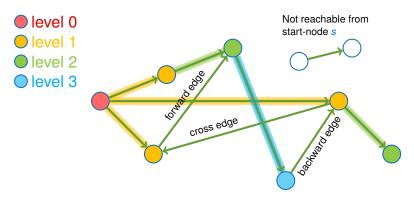


Figure: spanning tree of a breadth-first search

We start with all vertices unmarked and mark visited vertices

- We start with all vertices unmarked and mark visited vertices
- Mark the start vertex s

- 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)

- We start with all vertices unmarked and mark visited vertices
- 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)

Search starts with long paths (searching with depth)

- Search starts with long paths (searching with depth)
- Marks like breadth-first search all connected vertices

- Search starts with long paths (searching with depth)
- Marks like breadth-first search all connected vertices
- If the graph is acyclic we get a topological sorting

Depth-first search:

- Search starts with long paths (searching with depth)
- Marks like breadth-first search all connected vertices
- If the graph is acyclic we get a topological sorting
 - Each newly visited vertex gets marked by an increasing number

Depth-first search:

- Search starts with long paths (searching with depth)
- Marks like breadth-first search 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

Graphs

Connected Components - Depth-First Search

JNI REIBURG

- The marked vertices create a different spanning tree containing all reachable nodes
- start-node
- opath 1
- path 2
- opath 3

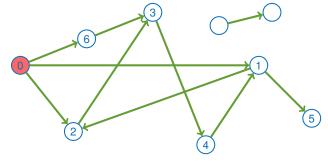


Figure: spanning tree of a depth-first search

- The marked vertices create a different spanning tree containing all reachable nodes
- start-node
- path 1
- path 2
- opath 3

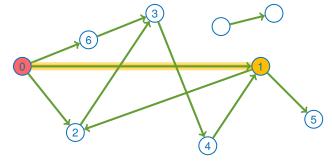


Figure: spanning tree of a depth-first search

- The marked vertices create a different spanning tree containing all reachable nodes
- start-node
- path 1
- path 2
- opath 3

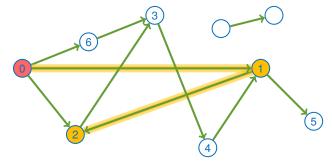


Figure: spanning tree of a depth-first search

- The marked vertices create a different spanning tree containing all reachable nodes
- start-node
- path 1
- path 2
- opath 3

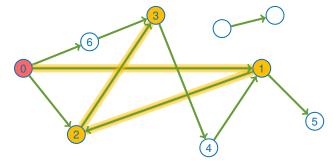


Figure: spanning tree of a depth-first search

- start-node
- path 1
- path 2
- opath 3

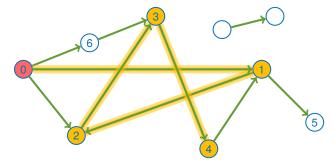


Figure: spanning tree of a depth-first search

- start-node
- path 1
- path 2
- opath 3

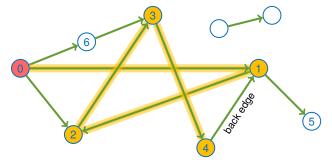


Figure: spanning tree of a depth-first search

- start-node
- path 1
- path 2
- opath 3

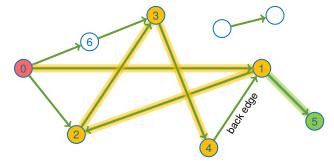


Figure: spanning tree of a depth-first search

- The marked vertices create a different spanning tree containing all reachable nodes
- start-node
- path 1
- path 2
- opath 3

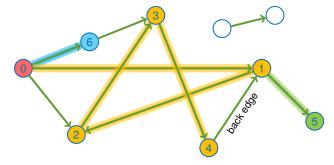


Figure: spanning tree of a depth-first search

- The marked vertices create a different spanning tree containing all reachable nodes
- start-node
- path 1
- path 2
- opath 3

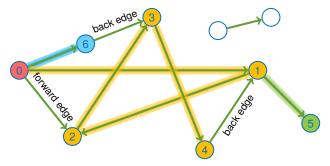


Figure: spanning tree of a depth-first search

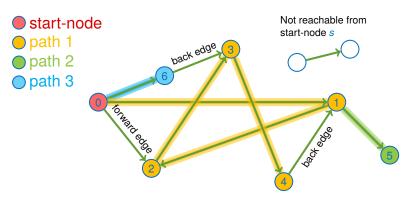


Figure: spanning tree of a depth-first search

Graphs

Why is this called Breadth - and Depth First Search?



Constant costs for each visited vertex and edge

- Constant costs for each visited vertex and edge
- We get a runtime complexity of $\Theta(|V'| + |E'|)$

- 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

- 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

- 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

Structure



Graphs

Introduction Implementation

Application example

Image processing



Image processing



■ Connected component labeling

Image processing

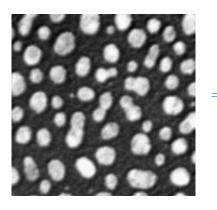


- Connected component labeling
- Counting of objects in an image

Image processing



- Connected component labeling
- Counting of objects in an image



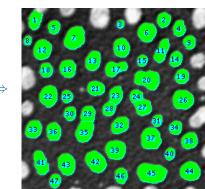


Image processing



What's object, what's background?

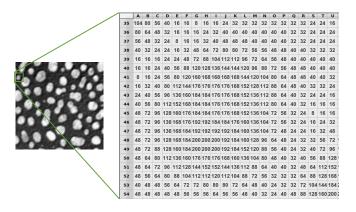
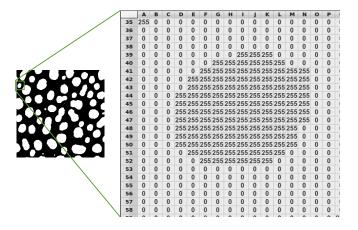


Image processing

Convert to black white using threshold:

value = 255 if value > 100 else 0



Application example Image processing



Application example Image processing

FEE

Interpret image as graph:

Each white pixel is a node

Application example Image processing



- Each white pixel is a node
- Edges between adjacent pixels (normally 4 or 8 neighbors)

- 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

- 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)

Image processing



Find connected components:

Image processing



Find connected components:

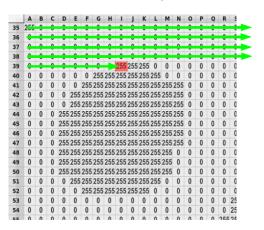
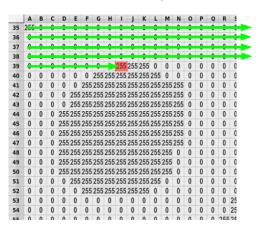


Image processing



Find connected components:

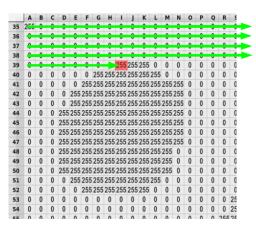


Search pixel-by-pixel for non-zero intensity

Image processing



Find connected components:

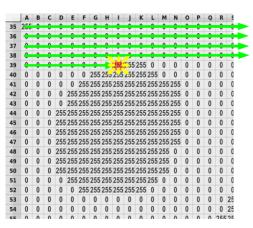


- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1

Image processing



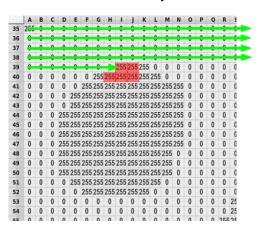
Find connected components:



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

Image processing

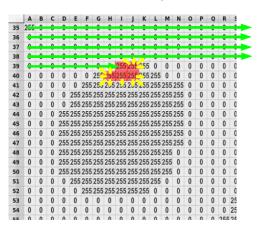




- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

Image processing

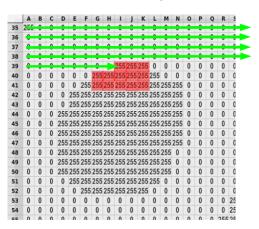




- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

Image processing

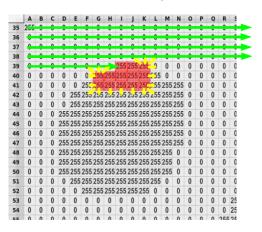




- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

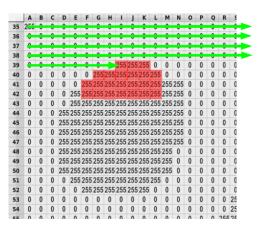
Image processing





- 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

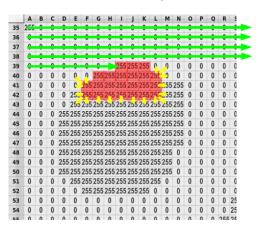




- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

Image processing

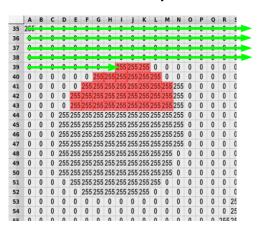




- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

Image processing

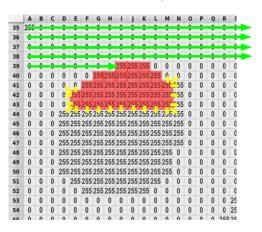




- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

Image processing

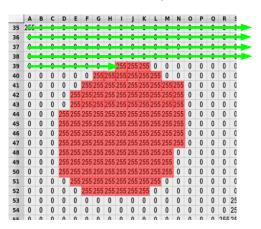




- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

Image processing

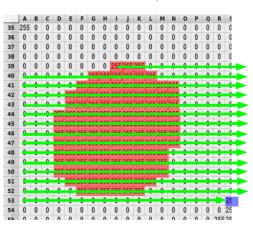




- Search pixel-by-pixel for non-zero intensity
- Label found pixel as component 1
- Check neighbors of all new labeled pixels
- Label non-zero pixels as component 1

Image processing





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

Result of connected component labeling:

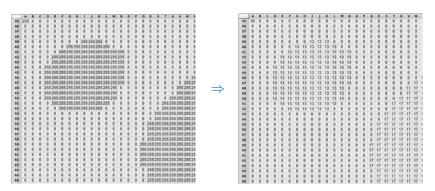


Figure: Result: particle indices instead of intensities

■ 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.

■ Graph-Search

■ Graph-Connectivity

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