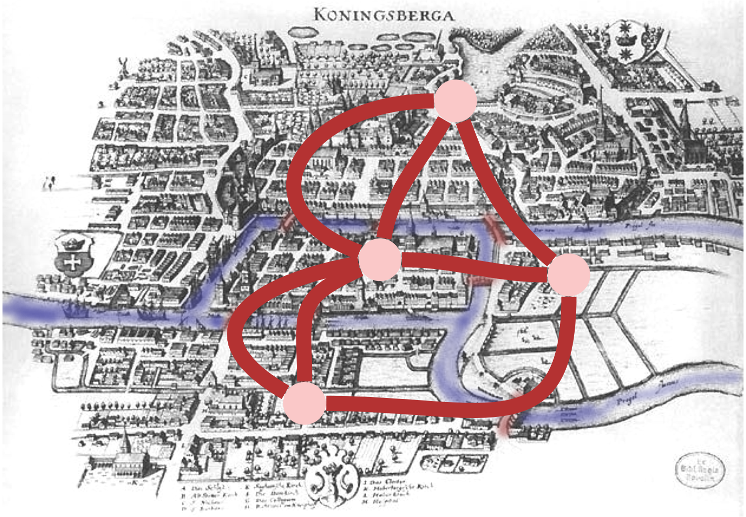
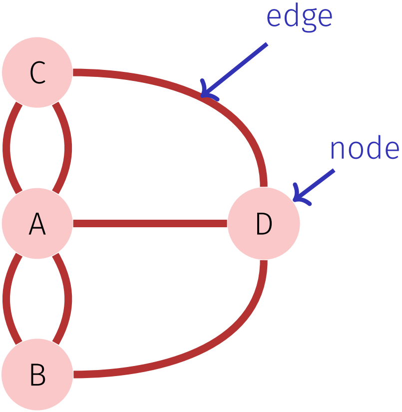
25. Graphs

Notation, Representation, Graph Traversal (DFS, BFS), Topological Sorting , Reflexive transitive closure, Connected components [Ottman/Widmayer, Kap. 9.1 - 9.4,Cormen et al, Kap. 22]

Königsberg 1736



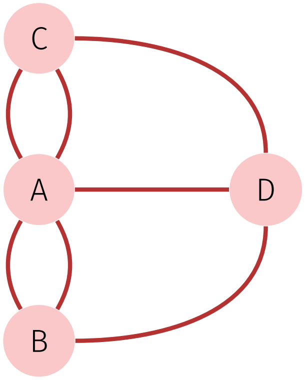
[Multi]Graph



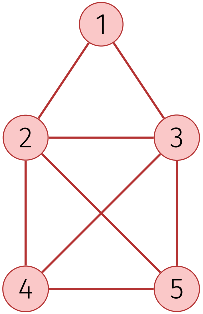
Cycles

- Is there a cycle through the town (the graph) that uses each bridge (each edge) exactly once?
- Euler (1736): no.
- Such a *cycle* is called *Eulerian path*.
- Eulerian path $\Leftrightarrow$ each node provides an even number of edges (each node is of an *even degree*).

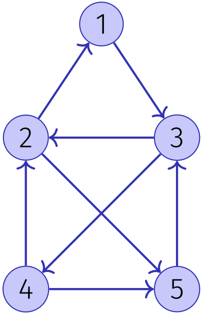
  '$\Rightarrow$" is straightforward, "$\Leftarrow$" ist a bit more difficult but still elementary.



Notation



undirected



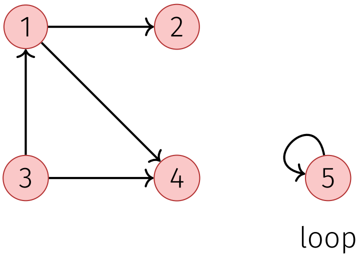
directed

$$V = \{1, 2, 3, 4, 5\}$$
$$E = \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{2, 4\}, \{2, 5\}, \{3, 4\}, \{3, 5\}, \{4, 5\}\}$$

$$V = \{1, 2, 3, 4, 5\}$$
$$E = \{(1, 3), (2, 1), (2, 5), (3, 2), (3, 4), (4, 2), (4, 5), (5, 3)\}$$

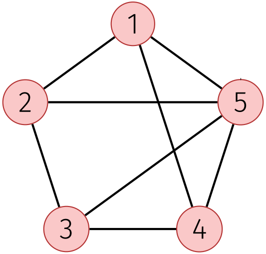
Notation

A **directed graph** consists of a set $V = \{v_1, \dots, v_n\}$ of nodes (*Vertices*) *and a set* $E \subseteq V \times V$ *of* Edges. The same edges may not be contained more than once.



Notation

An **undirected graph** consists of a set $V = \{v_1, \ldots, v_n\}$ of nodes a and a set $E \subseteq \{\{u, v\} | u, v \in V\}$ of edges. Edges may not be contained more than once.[38]

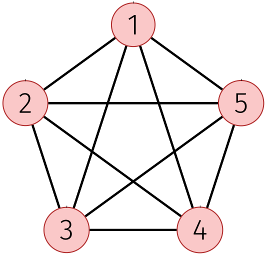


undirected graph

[38] As opposed to the introductory example – it is then called multi-graph.

Notation

An undirected graph $G = (V, E)$ without loops where E comprises all edges between pairwise different nodes is called **complete**.

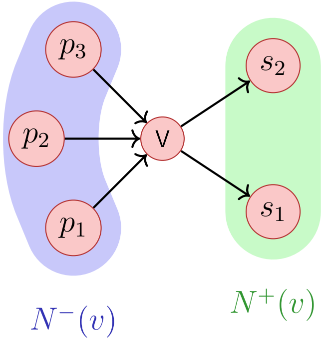


a complete undirected graph

Notation

For directed graphs $G = (V, E)$

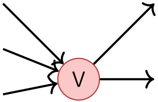
- $w \in V$ is called adjacent to $v \in V$, if $(v, w) \in E$
- **Predecessors** of $v \in V$: $N^-(v) := \{u \in V | (u, v) \in E\}$.
  Successors: $N^+(v) := \{u \in V | (v, u) \in E\}$



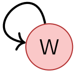
Notation

For directed graphs $G = (V, E)$

- **In-Degree**: $\deg^-(v) = |N^-(v)|$,
  Out-Degree: $\deg^+(v) = |N^+(v)|$



$\deg^-(v) = 3$, $\deg^+(v) = 2$



$\deg^-(w) = 1$, $\deg^+(w) = 1$

Notation

For undirected graphs $G = (V, E)$:

- $w \in V$ is called **adjacent** to $v \in V$, if $\{v, w\} \in E$
- **Neighbourhood** of $v \in V$: $N(v) = \{w \in V | \{v, w\} \in E\}$
- **Degree** of v: $\deg(v) = |N(v)|$ with a special case for the loops: increase the degree by 2.

$\deg(v) = 5$ $\deg(w) = 2$

Node Degrees $\leftrightarrow$ Number of Edges

Handshaking Lemma:

For each graph $G = (V, E)$ it holds

1. $\sum_{v \in V} \deg^-(v) = \sum_{v \in V} \deg^+(v) = |E|$, for G directed
2. $\sum_{v \in V} \deg(v) = 2|E|$, for G undirected.

Paths

- **Path**: a sequence of nodes $\langle v_1, \ldots, v_{k+1} \rangle$ such that for each $i \in \{1 \ldots k\}$ there is an edge from v_i to v_{i+1} .
- **Length** of a path: number of contained edges k.
- **Simple path**: path without repeating vertices

Connectedness

- An undirected graph is called **connected**, if for each pair $v, w \in V$ there is a connecting path.
- A directed graph is called **strongly connected**, if for each pair $v, w \in V$ there is a connecting path.
- A directed graph is called **weakly connected**, if the corresponding undirected graph is connected.

Simple Observations

- generally: $0 \leq |E| \in \mathcal{O}(|V|^2)$
- connected graph: $|E| \in \Omega(|V|)$
- complete graph: $|E| = \frac{|V|\cdot(|V|-1)}{2}$ (undirected)
- Maximally $|E| = |V|^2$ (directed),$|E| = \frac{|V|\cdot(|V|+1)}{2}$ (undirected)

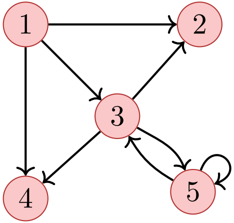
Cycles

- **Cycle**: path $\langle v_1, \ldots, v_{k+1} \rangle$ with $v_1 = v_{k+1}$
- **Simple cycle**: Cycle with pairwise different $v_1, \ldots, v_k$, that does not use an edge more than once.
- **Acyclic**: graph without any cycles.

Conclusion: undirected graphs cannot contain cycles with length 2 (loops have length 1)

Representation using a Matrix

Graph $G = (V, E)$ with nodes $v_1 \ldots, v_n$ stored as **adjacency matrix** $A_G = (a_{ij})_{1 \le i,j \le n}$ with entries from $\{0, 1\}$. $a_{ij} = 1$ if and only if edge from v_i to v_j.

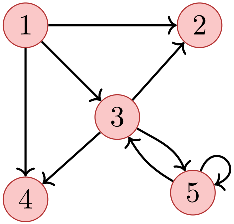


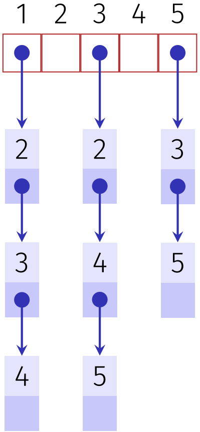
$$\begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Memory consumption $\Theta(|V|^2)$. A_G is symmetric, if G undirected.

Representation with a List

Many graphs $G = (V, E)$ with nodes $v_1, \dots, v_n$ provide much less than n^2 edges. Representation with **adjacency list**: Array $A[1], \dots, A[n]$, A_i comprises a linked list of nodes in $N^+(v_i)$.





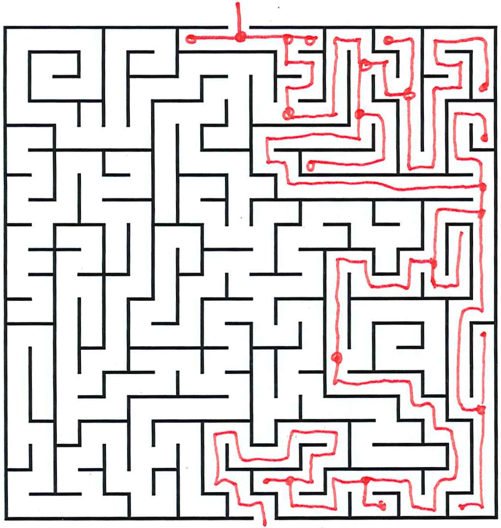
Memory Consumption $\Theta(|V| + |E|)$.

Runtimes of simple Operations

Operation	Matrix	List
Find neighbours/successors of $v \in V$		
find $v \in V$ without neighbour/successor		
$(v, u) \in E$?		
Insert edge		
Delete edge (v, u)		

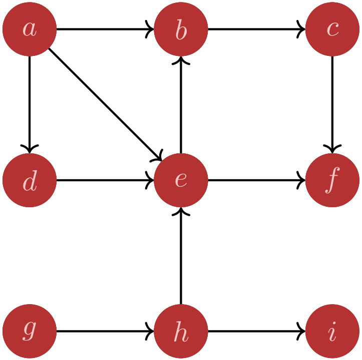
Exercise Class

Depth First Search



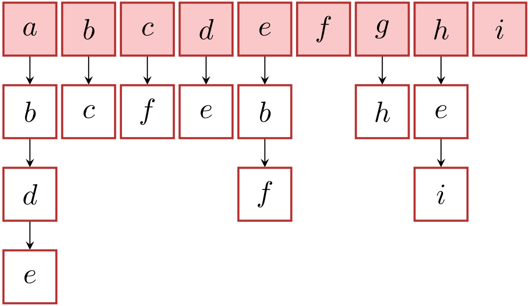
Graph Traversal: Depth First Search

Follow the path into its depth until nothing is left to visit.



Order $a, b, c, f, d, e, g, h, i$

adjacency list



Colors

Conceptual coloring of nodes

- **white:** node has not been discovered yet.
- **grey:** node has been discovered and is marked for traversal / being processed.
- **black:** node was discovered and entirely processed.

Algorithm Depth First visit DFS-Visit(G, v)

Input: graph $G = (V, E)$, Knoten v.

$v.color \leftarrow$ grey
// visit v
foreach $w \in N^+(v)$ **do**
  if $w.color =$ white **then**
    DFS-Visit(G, w)

$v.color \leftarrow$ black

Depth First Search starting from node v. Running time (without recursion): $\Theta(\deg^+ v)$

Algorithm Depth First visit DFS-Visit(G)

Input: graph $G = (V, E)$

foreach $v \in V$ **do**
    $v.color \leftarrow$ white

foreach $v \in V$ **do**
    if $v.color =$ white **then**
        DFS-Visit(G,v)

Depth First Search for all nodes of a graph. Running time:
$\Theta(|V| + \sum_{v \in V}(\deg^+(v) + 1)) = \Theta(|V| + |E|)$.

Interpretation of the Colors

When traversing the graph, a tree (or Forest) is built. When nodes are discovered there are three cases

- White node: new tree edge
- Grey node: cycle (“back-edge”)
- Black node: forward- / cross edge

[Iterative DFS-Visit(G, v)]

Input: graph $G = (V, E)$, $v \in V$ with $v.color =$ white

Stack $S \leftarrow \emptyset$
$v.color \leftarrow$ grey; S.push(v) // invariant: grey nodes always on stack
while $S \neq \emptyset$ **do**
  $w \leftarrow$ nextWhiteSuccessor(v) // code: next slide
  if $w \neq$ null **then**
    $w.color \leftarrow$ grey; S.push(w)
    $v \leftarrow w$ // work on w. parent remains on the stack
  else
    $v.color \leftarrow$ black // no grey successors, v becomes black
    if $S \neq \emptyset$ **then**
      $v \leftarrow S$.pop() // visit/revisit next node
      if $v.color =$ grey **then** S.push(v)

Memory Consumption Stack $\Theta(|V|)$

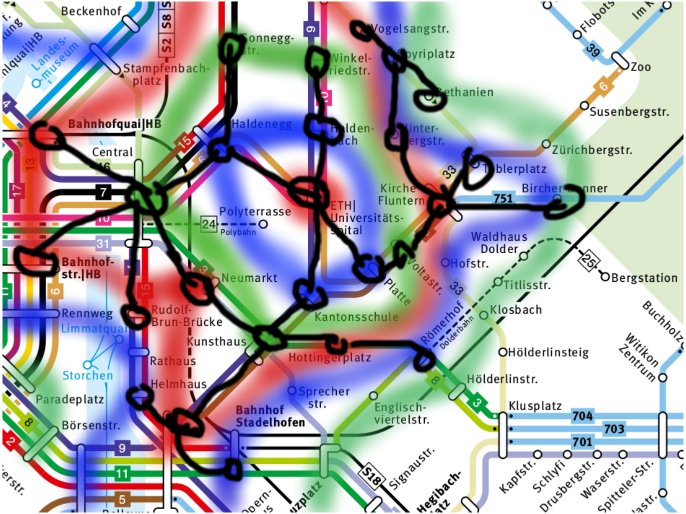
[nextWhiteSuccessor(v)]

Input: node $v \in V$
Output: Successor node u of v with $u.color =$ white, null otherwise

```
foreach u ∈ N⁺(v) do
    if u.color = white then
        return u
return null
```

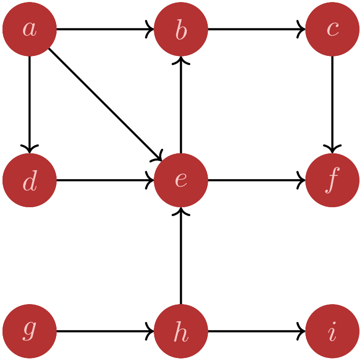
There are simpler variants of iterative depth first search. Howeber, they do not admit the same kind of interpretation of the edges between colored nodes. Moreover, they usually have a worst-case memory consumption of $\Theta(|E|)$

Breadth First Search



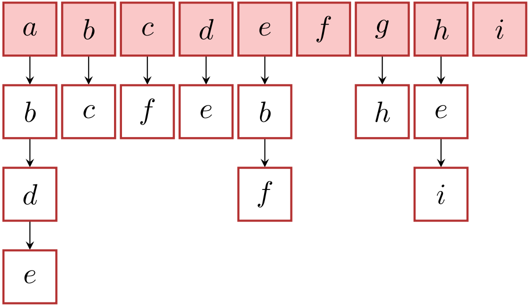
Graph Traversal: Breadth First Search

Follow the path in breadth and only then descend into depth.



Order $a, b, d, e, c, f, g, h, i$

Adjacency List



(Iterative) BFS-Visit(G, v)

Input: graph $G = (V, E)$

Queue $Q \leftarrow \emptyset$
enqueue(Q, v)
v.visited $\leftarrow$ true
while $Q \neq \emptyset$ **do**
    $w \leftarrow$ dequeue(Q)
    // visit w
    foreach $c \in N^+(w)$ **do**
        if c.visited $=$ false **then**
            c.visited $\leftarrow$ true
            enqueue(Q, c)

Algorithm requires extra space of $\mathcal{O}(|V|)$.

Main program BFS-Visit(G)

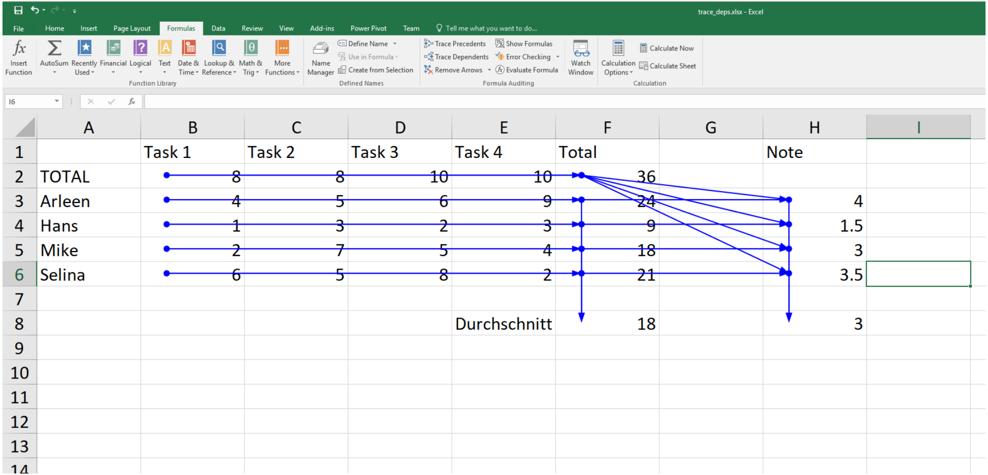
Input: graph $G = (V, E)$

foreach $v \in V$ **do**
    v.visited $\leftarrow$ false

foreach $v \in V$ **do**
    if v.visited $=$ false **then**
        BFS-Visit(G,v)

Breadth First Search for all nodes of a graph. Running time: $\Theta(|V| + |E|)$.

Topological Sorting



Evaluation Order?

Topological Sorting

Topological Sorting of an acyclic directed graph $G = (V, E)$:
Bijective mapping

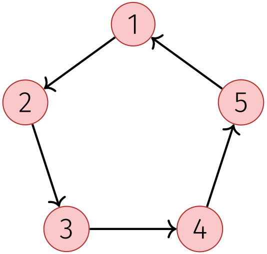
$$\mathrm{ord} : V \to \{1, \ldots, |V|\}$$

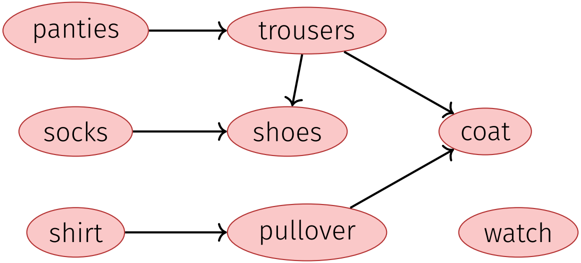
such that

$$\mathrm{ord}(v) < \mathrm{ord}(w) \ \forall \ (v, w) \in E.$$

Identify i with Element $v_i := \mathrm{ord}^1(i)$. Topological sorting $\widehat{=} \langle v_1, \ldots, v_{|V|} \rangle$.

(Counter-)Examples





Cyclic graph: cannot be sorted topologically.

A possible toplogical sorting of the graph:
shirt, pullover, panties, watch, trousers, coat, socks, shoes

Observation

Theorem 21

A directed graph $G = (V, E)$ permits a topological sorting if and only if it is acyclic.

Proof "$\Rightarrow$"

If G contains a cycle it cannot permit a topological sorting, because in a cycle $\langle v_{i_1}, \dots, v_{i_m} \rangle$ it would hold that $v_{i_1} < \cdots < v_{i_m} < v_{i_1}$.

Proof “$\Leftarrow$”

- Base case ($n = 1$): Graph with a single node without loop can be sorted topologically, set $\text{ord}(v_1) = 1$.
- Hypothesis: Graph with n nodes can be sorted topologically
- Step ($n \rightarrow n + 1$):

    1. G contains a node v_q with in-degree $\deg^-(v_q) = 0$. Otherwise iteratively follow edges backwards – after at most $n + 1$ iterations a node would be revisited. Contradiction to the cycle-freeness.
    2. Graph without node v_q and without its edges can be topologically sorted by the hypothesis. Now use this sorting and set $\text{ord}(v_i) \leftarrow \text{ord}(v_i) + 1$ for all $i \neq q$ and set $\text{ord}(v_q) \leftarrow 1$.

Algorithm Topological-Sort(G)

Input: graph $G = (V, E)$.
Output: Topological sorting ord

Stack $S \leftarrow \emptyset$
foreach $v \in V$ **do** $A[v] \leftarrow 0$
foreach $(v, w) \in E$ **do** $A[w] \leftarrow A[w] + 1$ // Compute in-degrees
foreach $v \in V$ with $A[v] = 0$ **do** push(S, v) // Memorize nodes with in-degree 0
$i \leftarrow 1$
while $S \neq \emptyset$ **do**
    $v \leftarrow$ pop(S); $\mathrm{ord}[v] \leftarrow i$; $i \leftarrow i + 1$ // Choose node with in-degree 0
    foreach $(v, w) \in E$ **do** // Decrease in-degree of successors
        $A[w] \leftarrow A[w] - 1$
        if $A[w] = 0$ **then** push(S, w)

if $i = |V| + 1$ **then return** ord **else return** “Cycle Detected”

Algorithm Correctness

Theorem 22

Let $G = (V, E)$ be a directed acyclic graph. Algorithm **`TopologicalSort`**(G) *computes a topological sorting* ord *for* G *with runtime* $\Theta(|V| + |E|)$.

Proof: follows from previous theorem:

1. Decreasing the in-degree corresponds with node removal.
2. In the algorithm it holds for each node v with $A[v] = 0$ that either the node has in-degree 0 or that previously all predecessors have been assigned a value $\mathrm{ord}[u] \leftarrow i$ and thus $\mathrm{ord}[v] > \mathrm{ord}[u]$ for all predecessors u of v. Nodes are put to the stack only once.
3. Runtime: inspection of the algorithm (with some arguments like with graph traversal)

Algorithm Correctness

Theorem 23

Let $G = (V, E)$ be a directed graph containing a cycle. Algorithm TopologicalSort terminates within $\Theta(|V| + |E|)$ steps and detects a cycle.

Proof: let $\langle v_{i_1}, \ldots, v_{i_k} \rangle$ be a cycle in G. In each step of the algorithm remains $A[v_{i_j}] \geq 1$ for all $j = 1, \ldots, k$. Thus k nodes are never pushed on the stack und therefore at the end it holds that $i \leq V + 1 - k$.

The runtime of the second part of the algorithm can become shorter. But the computation of the in-degree costs already $\Theta(|V| + |E|)$.