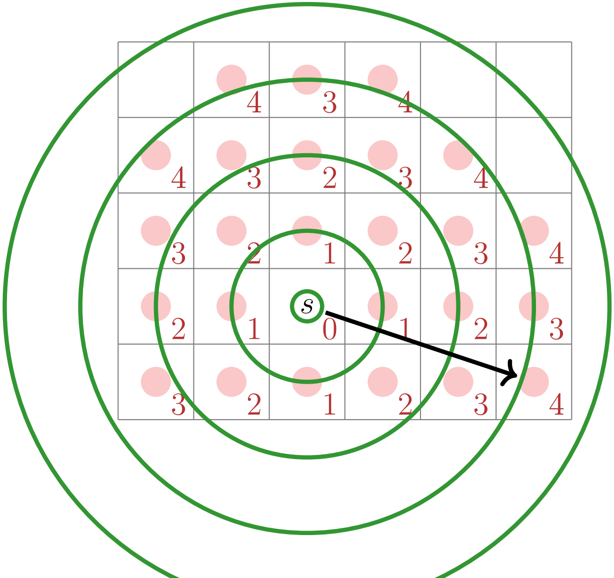
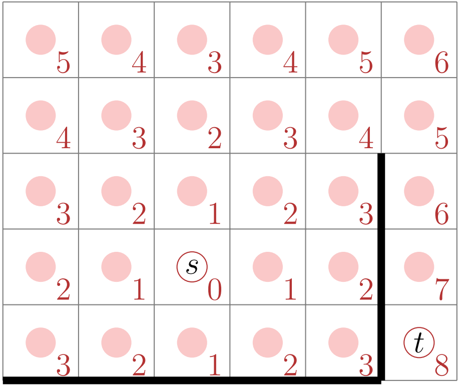
26.5 A*-Algorithm

Motivation A*



- Dijkstra Algorithm searches for all shortest paths, in all directions.

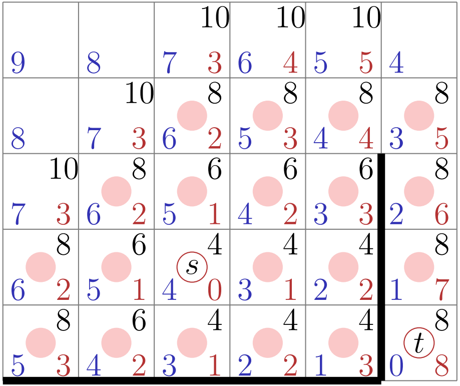
Motivation A*



- Dijkstra Algorithm searches for all shortest paths, in all directions.
- which is correct, because the algorithm does not know about the graph's structure.

A* in Action

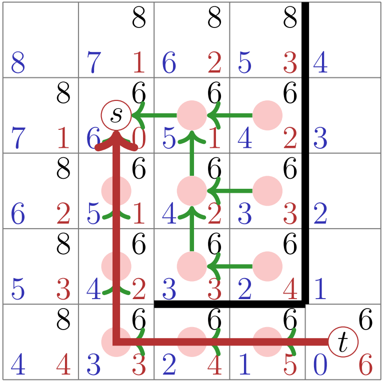
$\hat{f}(u) = d_s[u] + \hat{h}(u)$ $(\hat{h} = \delta_x + \delta_y$ Manhattan-Distance)



- Idea: equip algorithm with a preferred direction by ways of a distance heuristic $\hat{h}$
- The value of this heuristics needs to underestimate the distance to t and is added to the found distance d_s to s

Keep backward path

$\hat{f}(u) = d_s[u] + \hat{h}(u) \quad (\hat{h} = \delta_x + \delta_y)$



- The algorithm works like the Dijkstra-algorithm
- For finding the next candidate of R instead of the value d_s the value of $\hat{f} = \hat{h} + d_s$ is used

A*-Algorithm

Prerequisites

- Positively weighted, finite graph $G = (V, E, c)$
- $s \in V$, $t \in V$
- Distance estimate $\widehat{h}_t(v) \leq h_t(v) := \delta(v, t)\ \forall\ v \in V$.
- Wanted: shortest path $p : s \rightsquigarrow t$

A*-Algorithm($G, s, t, \hat{h}$)

Input: Positively weighted Graph $G = (V, E, c)$, starting point $s \in V$, end point $t \in V$, estimate $\widehat{h}(v) \leq \delta(v, t)$

Output: Existence and value of a shortest path from s to t

foreach $u \in V$ **do**

  $d[u] \leftarrow \infty$; $\widehat{f}[u] \leftarrow \infty$; $\pi[u] \leftarrow$ null

$d[s] \leftarrow 0$; $\widehat{f}[s] \leftarrow \widehat{h}(s)$; $N \leftarrow \{s\}$; $K \leftarrow \{\}$

while $N \neq \emptyset$ **do**

  $u \leftarrow$ ExtractMin$_{\widehat{f}}(N)$; $K \leftarrow K \cup \{u\}$

  if $u = t$ **then return** success

  foreach $v \in N^+(u)$ with $d[v] > d[u] + c(u, v)$ **do**

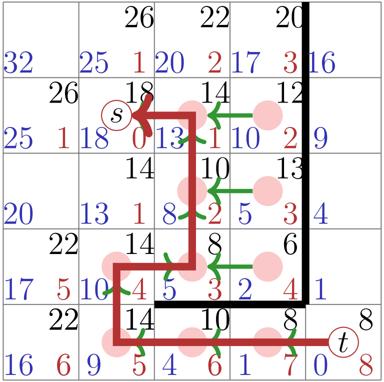
    $d[v] \leftarrow d[u] + c(u, v)$; $\widehat{f}[v] \leftarrow d[v] + \widehat{h}(v)$; $\pi[v] \leftarrow u$

    $N \leftarrow N \cup \{v\}$; $K \leftarrow K - \{v\}$

return failure

What if $\hat{h}$ does not underestimate

$$\hat{f}(u) = d_s[u] + \hat{h}(u) \qquad (\hat{h} = \delta_x^2 + \delta_y^2)$$



- Algorithm can terminate with the wrong result when $\hat{h}$ does not under-estimate the distance to t.
- although the heuristics looks reasonable otherwise (it is monotonic, for instance)

Revisiting nodes

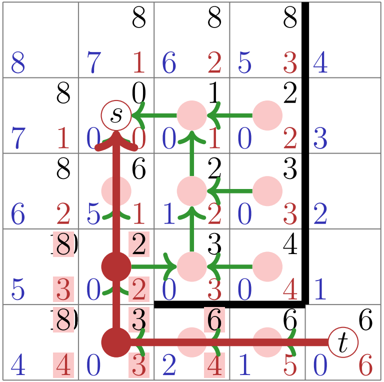
- The A*-algorithm can re-insert nodes that had been extracted from R before.
- This can lead to suboptimal behavior (w.r.t. running time of the algorithm).
- If $\widehat{h}$, in addition to being admissible ($\widehat{h}(v) \leq h(v)$ for all $v \in V$), fulfils monotonicity, i.e. if for all $(u, u') \in E$:

$$\widehat{h}(u') \leq \widehat{h}(u) + c(u', u)$$

  then the A*-Algorithm is equivalent to the Dijsktra-algorithm with edge weights $\tilde{c}(u, v) = c(u, v) + \widehat{h}(u) - \widehat{h}(v)$, and no node is re-inserted into R.
- It is not always possible to find monotone heuristics.

A crazy $\hat{h}$

$\hat{f}(u) = d_s[u] + \hat{h}(u)$



- Algorithm terminates correctly even if the distance heuristic is not monotonic
- It is then possible that nodes are removed and re-inserted into R multiple times.

Conclusion

- The A*-Algorithm is an extension of the Dijkstra algortihm by a distance heuristic $\hat{h}$.
- A* = Dijkstra if $\hat{h} \equiv 0$
- If $\hat{h}$ underestimates the real distance, the algorithm works correctly.
- If $\hat{h}$ is monotone in addition, then the algorithm works efficiently.
- In practical applications (e.g. routing), the choice of $\hat{h}$ is often intuitive and leads to a significant improvement over Dijkstra.

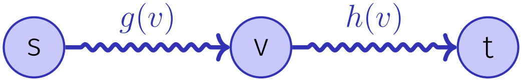
26.6 A*-Algorithm

Proof of correctness Not relevant for the exam

Notation

Let $f(v)$ be the distance of a shortest path from s to t via v, thus

$$f(v) := \underbrace{\delta(s,v)}_{g(v)} + \underbrace{\delta(v,t)}_{h(v)}$$



let p be a shortest path from s to t.
It holds that $f(s) = \delta(s,t)$ and $f(v) = f(s)$ for all $v \in p$.
Let $\widehat{g}(v) := d[v]$ be an estimate of $g(v)$ in the algorithm above. It holds that $\hat{g}(v) \geq g(v)$.
$\widehat{h}(v)$ is an estimate of $h(v)$ with $\widehat{h}(v) \leq h(v)$.

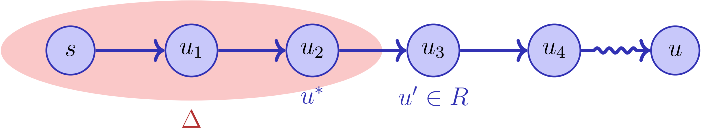
Why the Algorithm Works

Lemma 24

Let $u \in V$ and, at a time during the execution of the algorithm, $u \notin M$. Let p be a shortest path from s to u. Then there is a $u' \in p$ with $\widehat{g}(u') = g(u')$ and $u' \in R$.

The lemma states that there is always a node in the open set R with the minimal distance from s already computed and that belongs to a shortest path (if existing).

Illustration and Proof



Proof: If $s \in R$, then $\widehat{g}(s) = g(s) = 0$. Therefore, let $s \notin R$.

Let $p = \langle s = u_0, u_1, \ldots, u_k = u \rangle$ and $\Delta = \{u_i \in p, u_i \in M, \widehat{g}(u_i) = g(u_i)\}$.
$\Delta \neq \emptyset$, because $s \in \Delta$.

Let $m = \max\{i : u_i \in \Delta\}$, $u^* = u_m$. Then $u^* \neq u$, since $u \notin M$. Let $u' = u_{m+1}$.

1. $\widehat{g}(u') \leq \widehat{g}(u^*) + c(u^*, u')$ because u' has already been relaxed
2. $\widehat{g}(u^*) = g(u^*)$ (because $u^* \in \Delta$)
3. $\hat{g}(u') \geq g(u')$ (construction of $\hat{g}$)
4. $g(u') = g(u^*) + c(u^*, u')$ (because p optimal)

Therefore: $\widehat{g}(u') = g(u')$ and thus also $u' \in R$ because $u' \notin \Delta$. ■

Corollary

Corollary 25

If $\widehat{h}(u) \leq h(u)$ for all $u \in V$ and A- Algorithmus has not yet terminated. The for each shortest path p from s t t there is some node $u' \in p$ with $\hat{f}(u') \leq \delta(s,t) = f(t)$.*

If there is a shortest path p from s to t, then there is always a node in the open set R that underestimates the overal distance and that is on the shortest path.

Proof of the Corollary

Proof:

From the lemma: $\exists u' \in p$ with $\widehat{g}(u') = g(u')$.

Therefore:

$$\begin{aligned}\widehat{f}(u') &= \widehat{g}(u') + \widehat{h}(u') \\ &= g(u') + \widehat{h}(u') \\ &\leq g(u') + h(u') = f(u')\end{aligned}$$

Because p is shortest path: $f(u') = \delta(s, t)$. ■

Admissibility

Theorem 26

If there is a shortest path from s to t and $\hat{h}(u) \leq h(u) \; \forall \; u \in V$ then A terminates with $\hat{g}(t) = \delta(s,t)$*

Proof: If the algorithm terminates, then it termines with t with $f(t) = \widehat{g}(t) + 0 = g(t)$. That is because $\widehat{g}$ overestimates g at most and by the corollary above that algorithm always finds an element $v \in R$ with $f(v) \leq \delta(s,t)$.

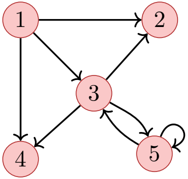
The algorithm terminates in finitely many steps. For finite graphs the maximal number of relaxing steps is bounded. [41] ■

[41] For a δ-graph the maximum number of relaxing steps before R contains only nodes with $\hat{f}(s) > \delta(s,t)$ is limited as well. The exact argument can be found in the seminal article Hart, P. E.; Nilsson, N. J.; Raphael, B. (1968). "A Formal Basis for the Heuristic Determination of Minimum Cost Paths".

27. Transitive Closure, All Pairs Shortest Paths

Reflexive transitive closure [Ottman/Widmayer, Kap. 9.2 Cormen et al, Kap. 25.2] Floyd-Warshall Algorithm [Ottman/Widmayer, Kap. 9.5.3 Cormen et al, Kap. 25.2]

Adjacency Matrix Product



$$B := A_G^2 = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}^2 = \begin{pmatrix} 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 2 \end{pmatrix}$$

Interpretation

Theorem 27

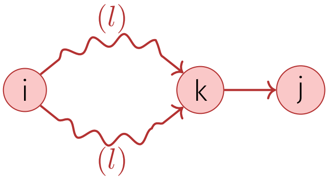
Let $G = (V, E)$ be a graph and $k \in \mathbb{N}$. Then the element $a_{i,j}^{(k)}$ of the matrix $(a_{i,j}^{(k)})_{1 \leq i,j \leq n} = (A_G)^k$ provides the number of paths with length k from v_i to v_j .

[Proof]

By Induction.

Base case: straightforward for $k = 1$. $a_{i,j} = a_{i,j}^{(1)}$.
Hypothesis: claim is true for all $k \leq l$
Step ($l \to l + 1$):

$$a_{i,j}^{(l+1)} = \sum_{k=1}^{n} a_{i,k}^{(l)} \cdot a_{k,j}$$



$a_{k,j} = 1$ iff egde k to j, 0 otherwise. Sum counts the number paths of length l from node v_i to all nodes v_k that provide a direct direction to node v_j, i.e. all paths with length $l + 1$.

Relation

Given a finite set V

(Binary) **Relation** R on V: Subset of the cartesian product
$V \times V = \{(a, b) | a \in V, b \in V\}$

Relation $R \subseteq V \times V$ is called

- **reflexive**, if $(v, v) \in R$ for all $v \in V$
- **symmetric**, if $(v, w) \in R \Rightarrow (w, v) \in R$
- **transitive**, if $(v, x) \in R$, $(x, w) \in R \Rightarrow (v, w) \in R$

The (Reflexive) Transitive Closure R^* of R is the smallest extension
$R \subseteq R^* \subseteq V \times V$ such that R^* is reflexive and transitive.

Graphs and Relations

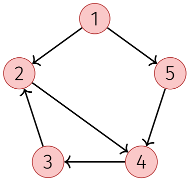
Graph $G = (V, E)$
adjacencies $A_G \hat{=}$ Relation $E \subseteq V \times V$ over V

- **reflexive** $\Leftrightarrow a_{i,i} = 1$ for all $i = 1, \ldots, n.$ (loops)
- **symmetric** $\Leftrightarrow a_{i,j} = a_{j,i}$ for all $i, j = 1, \ldots, n$ (undirected)
- **transitive** $\Leftrightarrow (u, v) \in E$, $(v, w) \in E \Rightarrow (u, w) \in E$. (reachability)

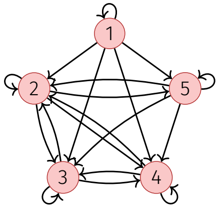
Reflexive Transitive Closure

Reflexive transitive closure of G $\Leftrightarrow$ **Reachability relation** E^*: $(v, w) \in E^*$ iff $\exists$ path from node v to w.

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$



$G = (V, E)$



$G^* = (V, E^*)$

Algorithm $A \cdot A$

Input: (Adjacency-)Matrix $A = (a_{ij})_{i,j=1\ldots n}$
Output: Matrix Product $B = (b_{ij})_{i,j=1\ldots n} = A \cdot A$

```
B ← 0
for r ← 1 to n do
    for c ← 1 to n do
        for k ← 1 to n do
            b_rc ← b_rc + a_rk · a_kc            // Number of Paths
return B
```

Counts number of paths of length 2

Algorithm $A \otimes A$

Input: Adjacency-Matrix $A = (a_{ij})_{i,j=1\dots n}$
Output: Modified Matrix Product $B = (b_{ij})_{i,j=1\dots n} = A \otimes A$

```
B ← A                                          // Keep paths
for r ← 1 to n do
    for c ← 1 to n do
        for k ← 1 to n do
            b_rc ← max{b_rc, a_rk · a_kc}      // Path: yes/no
return B
```

Computes which paths of length 1 and 2 exist

Computation of the Reflexive Transitive Closure

Goal: computation of $B = (b_{ij})_{1 \leq i,j \leq n}$ with $b_{ij} = 1 \Leftrightarrow (v_i, v_j) \in E^*$ First idea:

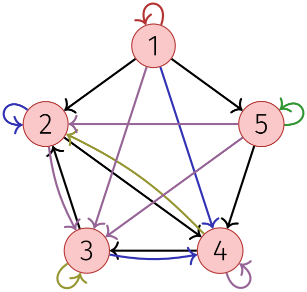
- Start with $B \leftarrow A$ and set $b_{ii} = 1$ for each i (Reflexivity.).
- Compute

$$B_n = \bigotimes_{i=1}^{n} B$$

with powers of 2 $B_2 := B \otimes B$, $B_4 := B_2 \otimes B_2$, $B_8 = B_4 \otimes B_4$ ...
$\Rightarrow$ running time $n^3 \lceil \log_2 n \rceil$

Improvement: Algorithm of Warshall (1962)

Inductive procedure: all paths known over nodes from $\{v_i : i < k\}$. Add node v_k.



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

Algorithm TransitiveClosure(A_G)

Input: Adjacency matrix $A_G = (a_{ij})_{i,j=1...n}$
Output: Reflexive transitive closure $B = (b_{ij})_{i,j=1...n}$ of G

$B \leftarrow A_G$
for $k \leftarrow 1$ **to** n **do**
    $b_{kk} \leftarrow 1$ // Reflexivity
    for $r \leftarrow 1$ **to** n **do**
        for $c \leftarrow 1$ **to** n **do**
            $b_{rc} \leftarrow \max\{b_{rc}, b_{rk} \cdot b_{kc}\}$ // All paths via v_k

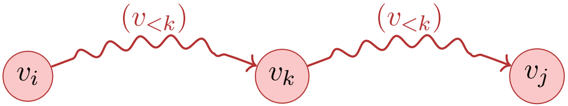
return B

Runtime $\Theta(n^3)$.

Correctness of the Algorithm (Induction)

Invariant (k): all paths via nodes with maximal index $< k$ considered.

- **Base case ($k = 1$)**: All directed paths (all edges) in A_G considered.
- **Hypothesis**: invariant (k) fulfilled.
- **Step** ($k \to k + 1$): For each path from v_i to v_j via nodes with maximal index k: by the hypothesis $b_{ik} = 1$ and $b_{kj} = 1$. Therefore in the k-th iteration: $b_{ij} \leftarrow 1$.



All shortest Paths

Compute the weight of a shortest path for each pair of nodes.

- $|V|\times$ Application of Dijkstra's Shortest Path algorithm $\mathcal{O}(|V|\cdot(|E|+|V|)\cdot\log|V|)$ (with Fibonacci Heap: $\mathcal{O}(|V|^2\log|V|+|V|\cdot|E|)$)
- $|V|\times$ Application of Bellman-Ford: $\mathcal{O}(|E|\cdot|V|^2)$
- There are better ways!

Induction via node number

Consider weights of all shortest paths S^k with intermediate nodes in[42] $V^k := \{v_1, \dots, v_k\}$, provided that weights for all shortest paths S^{k-1} with intermediate nodes in V^{k-1} are given.

- v_k no intermediate node of a shortest path of $v_i \rightsquigarrow v_j$ in V^k: Weight of a shortest path $v_i \rightsquigarrow v_j$ in S^{k-1} is then also weight of shortest path in S^k.
- v_k intermediate node of a shortest path $v_i \rightsquigarrow v_j$ in V^k: Sub-paths $v_i \rightsquigarrow v_k$ and $v_k \rightsquigarrow v_j$ contain intermediate nodes only from S^{k-1}.

[42] like for the algorithm of the reflexive transitive closure of Warshall

Induction via node number

$d^k(u,v)$ = Minimal weight of a path $u \rightsquigarrow v$ with intermediate nodes in V^k

Induktion

$$
\begin{aligned}
&d^k(u,v) = \min\{d^{k-1}(u,v), d^{k-1}(u,k) + d^{k-1}(k,v)\}(k \geq 1)\\
&d^0(u,v) = c(u,v)
\end{aligned}
$$

Algorithm Floyd-Warshall(G)

Input: Graph $G = (V, E, c)$ without negative weight cycles.
Output: Minimal weights of all paths d

$d^0 \leftarrow c$
for $k \leftarrow 1$ **to** $|V|$ **do**
    for $i \leftarrow 1$ **to** $|V|$ **do**
        for $j \leftarrow 1$ **to** $|V|$ **do**
            $d^k(v_i, v_j) = \min\{d^{k-1}(v_i, v_j), d^{k-1}(v_i, v_k) + d^{k-1}(v_k, v_j)\}$

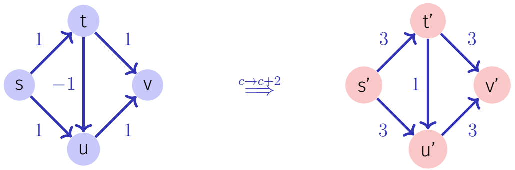
Runtime: $\Theta(|V|^3)$

Remark: Algorithm can be executed with a single matrix d (in place).

Reweighting

Idea: Reweighting the graph in order to apply Dijkstra's algorithm.

The following does **not** work. The graphs are not equivalent in terms of shortest paths.



Reweighting

Other Idea: “Potential” (Height) on the nodes

- $G = (V, E, c)$ a weighted graph.
- Mapping $h : V \to \mathbb{R}$
- New weights

$$\tilde{c}(u, v) = c(u, v) + h(u) - h(v), \ (u, v \in V)$$

Reweighting

Observation: A path p is shortest path in in $G = (V, E, c)$ iff it is shortest path in in $\tilde{G} = (V, E, \tilde{c})$

$$\begin{aligned}\tilde{c}(p) = \sum_{i=1}^{k} \tilde{c}(v_{i-1}, v_i) &= \sum_{i=1}^{k} c(v_{i-1}, v_i) + h(v_{i-1}) - h(v_i) \\ = h(v_0) - h(v_k) + &\sum_{i=1}^{k} c(v_{i-1}, v_i) = c(p) + h(v_0) - h(v_k)\end{aligned}$$

Thus $\tilde{c}(p)$ minimal in all $v_0 \rightsquigarrow v_k \iff c(p)$ minimal in all $v_0 \rightsquigarrow v_k$.

Weights of cycles are invariant: $\tilde{c}(v_0, \dots, v_k = v_0) = c(v_0, \dots, v_k = v_0)$

Johnson's Algorithm

Add a new node $s \notin V$:

$$
\begin{aligned}
G' &= (V', E', c') \\
V' &= V \cup \{s\} \\
E' &= E \cup \{(s, v) : v \in V\} \\
c'(u, v) &= c(u, v),\ u \neq s \\
c'(s, v) &= 0 (v \in V)
\end{aligned}
$$

Johnson's Algorithm

If no negative cycles, choose as height function the weight of the shortest paths from s,

$$h(v) = d(s, v).$$

For a minimal weight d of a path the following triangular inequality holds:

$$d(s, v) \leq d(s, u) + c(u, v).$$

Substitution yields $h(v) \leq h(u) + c(u, v)$. Therefore

$$\tilde{c}(u, v) = c(u, v) + h(u) - h(v) \geq 0.$$

Algorithm Johnson(G)

Input: Weighted Graph $G = (V, E, c)$
Output: Minimal weights of all paths D.

New node s. Compute $G' = (V', E', c')$
if BellmanFord(G', s) $=$ false **then** return "graph has negative cycles"
foreach $v \in V'$ **do**
    $h(v) \leftarrow d(s, v)$ // d aus BellmanFord Algorithmus
foreach $(u, v) \in E'$ **do**
    $\tilde{c}(u, v) \leftarrow c(u, v) + h(u) - h(v)$
foreach $u \in V$ **do**
    $\tilde{d}(u, \cdot) \leftarrow$ Dijkstra($\tilde{G}', u$)
    foreach $v \in V$ **do**
        $D(u, v) \leftarrow \tilde{d}(u, v) + h(v) - h(u)$

Analysis

Runtimes

- Computation of G': $\mathcal{O}(|V|)$
- Bellman Ford G': $\mathcal{O}(|V| \cdot |E|)$
- $|V| \times$ Dijkstra $\mathcal{O}(|V| \cdot |E| \cdot \log |V|)$
  (with Fibonacci Heap: $\mathcal{O}(|V|^2 \log |V| + |V| \cdot |E|)$)

Overal $\mathcal{O}(|V| \cdot |E| \cdot \log |V|)$
($\mathcal{O}(|V|^2 \log |V| + |V| \cdot |E|)$)