

Inhaltsverzeichnis

1	Komplexität und Laufzeiten	1
1.1	Definitionen	1
1.2	Worst-, Best- und Average-Case-Laufzeit	1
1.3	Asymptotische Betrachtung	1
1.4	L'Hospital	1
1.5	Die Klasse Gross- \mathcal{O}	1
1.6	Die Klasse Gross-Omega	1
1.7	Die Klasse Gross-Theta	2
1.8	Die Klasse Klein-O und Klein-Omega	2
1.9	Elementare Eigenschaften	2
1.10	Platzkomplexität	2
2	Suchalgorithmen	2
2.1	Bilineare Suche	2
2.2	Binäre Suche	3
3	Rekursionsgleichungen	3
3.1	Mastertheorem	3
4	Sortieren	3
4.1	Grundlegendes	3
4.2	Insertionsort	3
4.3	Divide-and-Conquer-Algorithmen	3
4.4	Mergesort	4
4.5	Heaps	4
4.6	Heapsort	4
4.7	Prioritätswarteschlange	5
4.8	Quicksort	5
4.9	Bubblesort	5
4.10	Countingsort	5
5	Elementare Datenstrukturen	6
5.1	Abstrakter Datentyp (ADT)	6
5.2	Binärbaum	6
5.3	Binäre Suchbäume	6
5.4	Rotationen in Bäumen	7
5.5	AVL-Bäume	7
5.6	Rot-Schwarz-Bäume (RBT)	7
6	Hashing	8
6.1	Dictionary	8
6.2	Direkte Adressierung	9
6.3	Grundlagen des Hashings	9

6.4	Hashfunktionen	9
6.5	Universelles Hashing	9
6.6	MD5	9
6.7	Verkettung	10
6.8	Offene Adressierung	10
6.9	Sondierung	10
7	Graphalgorithmen	11
7.1	Graphen	11
7.2	Pfade und Kreise	12
7.3	Repräsentation von Graphen	12
7.4	Graphendurchlauf	13
7.5	Zusammenhangskomponenten in ungerichteten Graphen	14
7.6	Zusammenhangskomponenten in gerichteten Graphen	14
7.7	Erreichbarkeitsanalysen	15
7.8	Kosaraju-Sharir-Algorithmus	15
7.9	Gerichtete azyklische Graphen (DAGs)	15
7.10	Topologisches Sortieren	15
7.11	Gewichtete Graphen	16
7.12	Kritische-Pfad-Problem	16
7.13	Eulersche Kreise	16
8	Minimale Spannbäume	16
8.1	Spannbäume	16
8.2	Minimale Spannbäume	17
8.3	Prüfer-Code	17
8.4	Prim und Kruskal	17
9	Kürzeste Pfade	17
9.1	Notation	18
9.2	Problemvarianten	18
9.3	Bellman-Ford-Algorithmus	18
9.4	Dijkstra-Algorithmus	18
9.5	All-Pairs Shortest Paths	19
9.6	Binäre Relationen	19
9.7	Warshall-Algorithmus	19
9.8	Floyd-Algorithmus	19
10	Matching	20
10.1	Definition	20
10.2	Algorithmus für bipartites Matching	20
10.3	Satz von Philip Hall	20
10.4	Satz von Dénes König	20
10.5	Algorithmus für gewichtetes Matching	21

11 Maximaler Fluss	21
11.1 Flussnetzwerke	21
11.2 Restnetzwerke	21
11.3 Ford-Fulkerson-Theorem	22
11.4 Ford-Fulkerson-Methode	22
11.5 Schnitte in Flussnetzwerken	22
11.6 Max-flow Min-cut Theorem	22
11.7 Edmonds-Karp-Algorithmus	23
12 Dynamische Programmierung	23
12.1 Memoization	23
12.2 The principle of optimality	23
12.3 Ketten von Matrixmultiplikation	23
12.4 Das Rucksackproblem	24
13 Greedy-Algorithmen	24
13.1 Grundlegendes	24
13.2 Sequentielle Speicherung	25
13.3 Stundenplanung	25
13.4 Korrektheitsbeweis	25
13.5 Huffman-Codes	25
14 Algorithmische Geometrie	26
14.1 Geometrische Interpretation	26
14.2 Links oder rechts?	26
14.3 Punkt, Strecke, Polygon	26
15 Abkürzungsverzeichnis	27

1 Komplexität und Laufzeiten

1.1 Definitionen

D_n : Menge aller Eingaben der Länge n

$t(I)$: Benötigte Anzahl elementarer Operationen für Eingabe I

$Pr(I)$: Wahrscheinlichkeit, dass Eingabe I auftritt

1.2 Worst-, Best- und Average-Case-Laufzeit

$$W(n) = \max \{t(I) | I \in D_n\}$$

$$B(n) = \min \{t(I) | I \in D_n\}$$

$$A(n) = \sum_{I \in D_n} Pr(I) \cdot t(I)$$

1.3 Asymptotische Betrachtung

Koeffizienten und Terme niedriger Ordnung werden ignoriert

1.4 L'Hospital

Seien $f, g: \mathbb{N} \rightarrow \mathbb{R}$ differenzierbare Funktionen mit $\lim_{n \rightarrow \infty} f(n) = \lim_{n \rightarrow \infty} g(n) = \infty$

Dann gilt: $\lim_{n \rightarrow \infty} \frac{g(n)}{f(n)} = \lim_{n \rightarrow \infty} \frac{g'(n)}{f'(n)}$

1.5 Die Klasse Gross- \mathcal{O}

$\mathcal{O}(f)$ ist die Menge von Funktionen, die **nicht schneller** als die Funktion f wachsen.

Diese Eigenschaft gilt ab einer Konstanten n_0 , Werte unter n_0 werden vernachlässigt.

$g \in \mathcal{O}(n^2)$ sagt mehr aus als $g \in \mathcal{O}(n^3)$.

Beispiel: $g \in \mathcal{O}(f)$ heißt, dass $c \cdot f(n)$ eine obere Schranke für $g(n)$ ist.

$$\mathcal{O}(cf(n)) = \mathcal{O}(f(n))$$

$$\mathcal{O}(f(n)) + \mathcal{O}(g(n)) = \mathcal{O}(f(n) + g(n))$$

$$\mathcal{O}(f(n)) + \mathcal{O}(g(n)) = \mathcal{O}(\max \{f(n), g(n)\})$$

$$\mathcal{O}(f(n))\mathcal{O}(g(n)) = \mathcal{O}(f(n)g(n))$$

$f(\mathcal{O}(1)) = \mathcal{O}(1)$ (das ist jetzt weder formal korrekt noch irgendwie sinnvoll, aber es stand da so)

$$\sum_{k=1}^n \mathcal{O}(f(k)) \subseteq \mathcal{O}(nf(n)), \text{ wenn } f(n) \text{ monoton steigt}$$

1.6 Die Klasse Gross-Omega

$\Omega(f)$ ist die Menge von Funktionen, die **nicht langsamer** als die Funktion f wachsen.

Diese Eigenschaft gilt ab einer Konstanten n_0 , Werte unter n_0 werden vernachlässigt.

$g \in \Omega(n^2)$ sagt mehr aus als $g \in \Omega(n)$.

Beispiel: $g \in \Omega(f)$ heißt, dass $c \cdot f(n)$ eine untere Schranke für $g(n)$ ist.

1.7 Die Klasse Gross-Theta

$\Theta(f)$ ist die Menge von Funktionen, die **genauso schnell** wie die Funktion f wachsen.
Diese Eigenschaft gilt ab einer Konstanten n_0 , Werte unter n_0 werden vernachlässigt.

Beispiel: $g \in \Theta(f)$ heißt, dass $c_1 \cdot f(n)$ eine untere Schranke und $c_2 \cdot f(n)$ eine obere Schranke für $g(n)$ ist.

$g \in \Theta(f)$ wenn $\lim_{n \rightarrow \infty} \frac{g(n)}{f(n)} = c$ für ein $0 < c < \infty$
 $g \in \Theta(f)$ gdw. $g \in \mathcal{O}(f)$ und $g \in \Omega(f)$

1.8 Die Klasse Klein-O und Klein-Omega

$o(f)$ ist die Menge der Funktionen, die **echt langsamer** als f wachsen.
 $\omega(f)$ ist die Menge der Funktionen, die **echt schneller** als f wachsen.

$f \in o(g)$ gdw. $g \in \omega(f)$

1.9 Elementare Eigenschaften

Reflexivität: $f \in \mathcal{O}(f)$, $f \in \Omega(f)$, $f \in \Theta(f)$

Transitivität: Aus $f \in \mathcal{O}/\Omega/\Theta(g)$ und $g \in \mathcal{O}/\Omega/\Theta(h)$ folgt $f \in \mathcal{O}/\Omega/\Theta(h)$.

Symmetrie von Θ : $f \in \Theta(g)$ gdw. $g \in \Theta(f)$

Beziehung zwischen \mathcal{O} und Ω : $f \in \mathcal{O}(g)$ gdw. $g \in \Omega(f)$

1.10 Platzkomplexität

Wir definieren $S(n)$ als Platzkomplexität.

Bei der Speicherung eines Liedes ist $S(n) \in \mathcal{O}(n)$ und $S(n) \in \Omega(1)$.

2 Suchalgorithmen

2.1 Bilineare Suche

Array wird aus beiden Richtungen „gleichzeitig“ durchsucht.
Zeitkomplexität wie bei linearer Suche $\mathcal{O}(n)$.

2.2 Binäre Suche

Sortiertes Array E mit n Elementen, gesucht ist das Element K .

Wir beginnen in der Mitte und **halbieren** den Suchraum in jedem Durchlauf.

Zeitkomplexität ist $\mathcal{O}(\log n)$, allerdings ist die Initialisierung mit $\mathcal{O}(n \log n)$ aufwändig.

3 Rekursionsgleichungen

3.1 Mastertheorem

$T(n) = b \cdot T(\frac{n}{c}) + f(n)$ mit $b \geq 1$ und $c > 1$

Anzahl der Blätter im Rekursionsbaum: n^E mit $E = \frac{\log(b)}{\log(c)} = \log_c(b)$

Fall 1

Bedingung: $f(n) \in \mathcal{O}(n^{E-\epsilon})$ für ein $\epsilon > 0$

Es folgt: $T(n) \in \Theta(n^E)$

Fall 2

Bedingung: $f(n) \in \Theta(n^E)$

Es folgt: $T(n) \in \Theta(n^E \cdot \log n)$

Fall 3

1. Bedingung: $f(n) \in \Omega(n^{E+\epsilon})$ für ein $\epsilon > 0$

2. Bedingung: $b \cdot f(\frac{n}{c}) \leq d \cdot f(n)$ für ein $d < 1$ und ein hinreichend großes n

Es folgt: $T(n) \in \Theta(f(n))$

4 Sortieren

4.1 Grundlegendes

Ein Sortieralgorithmus ist **stabil**, wenn er die Reihenfolge der Elemente, deren Sortierschlüssel gleich sind, aufrecht erhält. Wir bezeichnen einen Sortieralgorithmus als **in-place**, wenn er außer dem für die Speicherung der zu bearbeitenden Daten benötigten Speicher nur eine konstante Menge von Speicher benötigt (die also von der zu bearbeitenden Datenmenge unabhängig ist).

4.2 Insertionsort

Unsortierte Elemente werden in sortiertes Array vor dem nächstgrößeren Element eingefügt.

Insertionsort ist im Worst- und Average-Case quadratisch, im Best-Case linear.

4.3 Divide-and-Conquer-Algorithmen

Teile-und-Beherrsche Algorithmen (divide-and-conquer) teilen das Problem in mehrere Teilprobleme auf, die dem Ausgangsproblem ähneln, jedoch von kleinerer Größe sind.

Sie lösen die Teilprobleme rekursiv und kombinieren diese Lösungen dann, um die Lösung des eigentlichen Problems zu erstellen.

4.4 Mergesort

Array wird in zwei möglichst gleich große Hälften geteilt, die durch rekursive Mergesort-Aufrufe sortiert werden.

Es ist $W(n) = B(n) = A(n) \in \Theta(n \cdot \log(n))$.

Speicherbedarf $\Theta(n)$ für die Kopie des Arrays und $\Theta(\log(n))$ für die Verwaltung der Rekursion.

4.5 Heaps

Ein Heap ist ein Binärbaum, der Elemente mit Schlüsseln enthält. Für **Max-Heaps** gelten folgende Bedingungen:

- Der Schlüssel eines Knotens ist stets mindestens so groß wie die Schlüssel seiner Kinder.
- Alle Ebenen (mit Ausnahme der untersten) sind komplett gefüllt.
- Die Blätter befinden sich damit auf einer (oder zwei) Ebene(n).
- Die Blätter der untersten Ebene sind linksbündig angeordnet.

Heaps lassen sich in Arrays einbetten, dabei liegt die Wurzel in $a[0]$, das linke Kind von $a[i]$ in $a[2 \cdot i + 1]$ und das rechte Kind in $a[2 \cdot i + 2]$.

Die Höhe eines Heaps in Abhängigkeit von seinen Elementen ist gegeben durch $H = \lfloor \log_2(n) \rfloor$. Es ist $H \in \mathcal{O}(\log n)$.

Für die Anzahl der Knoten im linken Teilbaum l und die Knotenzahl im rechten Teilbaum r gilt stets: $r \leq l \leq 2r + 1$.

Ein Heap wird Reihe für Reihe von oben nach unten aufgebaut, indem neue Elemente möglichst weit links angefügt werden. Ist ein neues Element größer als sein Elternknoten, wird rekursiv nach oben getauscht.

4.6 Heapsort

Im Max-Heap enthält die Wurzel immer die größte Ziffer, deshalb tauschen wir die Wurzel mit dem letzten Element des Heaps, das in der untersten Ebene ganz rechts steht. Nun ist die letzte Position im Heap sortiert, wir können den Heap nun als Array aufschreiben und haben die letzte Position sortiert.

Dieses Verfahren wird rekursiv angewendet, indem das letzte, sortierte Element gestrichen wird. Anschließend muss wieder ein Max-Heap aufgebaut werden und wir beginnen von vorne, bis im Heap keine Elemente mehr enthalten sind, sodass das Array sortiert ist.

Auch Heapsort hat eine Laufzeitkomplexität von $\mathcal{O}(n \cdot \log(n))$, der Speicherplatzbedarf ist konstant. Heapsort ist ein Inplace-Sortieralgorithmus und nicht stabil.

4.7 Prioritätswarteschlange

Elemente werden mit Schlüsseln versehen, wobei jeder Schlüssel höchstens an ein Element vergeben wird. Die Schlüssel betrachten wir nun als Priorität. Prioritätswarteschlangen können mit einem Array, einem sortierten Array oder einem Heap implementiert werden.

4.8 Quicksort

Wir wählen ein beliebiges Pivotelement und teilen das Array danach in zwei Teile - der linke Teil enthält Elemente, die kleiner als das Pivotelement sind, der rechte Teil enthält Elemente, die mindestens so groß sind wie das Pivotelement, allerdings nicht das Pivotelement selbst. Das wiederholen wir rekursiv. Die Pivotelemente sind am Ende in korrekter Reihenfolge.

Die Platzkomplexität von Quicksort ist $\Theta(\log(n))$, es ist $W(n) \in \Theta(n^2)$, $B(n) \in \mathcal{O}(n \cdot \log(n))$ und $A(n) \in \mathcal{O}(n \cdot \log(n))$.

4.9 Bubblesort

Es werden alle nebeneinanderstehenden Paare in der Liste vertauscht, die momentan in der falschen Ordnung stehen. Die Liste rechts von der letzten Tauschposition des letzten Durchgangs ist dabei stets bereits sortiert. Sobald keine Änderungen mehr auftreten, ist der Algorithmus fertig - ansonsten muss der Algorithmus erneut durchlaufen.

Damit ist $W(n) \in \mathcal{O}(n^2)$.

4.10 Countingsort

Die Elemente des Array sind aus dem Bereich $\{0, \dots, B\}$ für einen bekannten Wert B . Wir erstellen ein Counting-Array, das so viele Positionen wie annehmbare Werte hat - jedem annehmbaren Wert wird (fortlaufend) eine Position im Counting-Array zugeordnet. Dieses Array wird mit Nullen initialisiert.

Nun gehen wir das zu sortierende Array durch. Für jedes Element im Array wird der zum Element gehörende Wert im Counting-Wert um 1 inkrementiert. Nun ersetzen wir im Counting-Array die Elemente durch die jeweiligen Teilsummen ($c[0] = c[0]$, $c[1] = c[0] + c[1]$, ...).

Abschließend gehen wir das unsortierte Array von hinten nach vorne durch. Bei jedem Element des unsortierten Arrays steht nun die neue, sortierte Position des Elements im Counting-Array. Bei jedem Einfügen eines Elements muss der zu diesem Element gehörige Wert im Counting-Array um 1 dekrementiert werden.

Die Laufzeit von Countingsort ist $\in \mathcal{O}(n + B)$, der Speicherbedarf $\in \mathcal{O}(n + B)$.

5 Elementare Datenstrukturen

5.1 Abstrakter Datentyp (ADT)

Ein abstrakter Datentyp besteht aus einer Datenstruktur (Menge von Werten) und einer Menge von Operationen darauf, zum Beispiel aus Konstruktor, Zugriffsfunktionen und Bearbeitungsfunktionen. Damit sind die Daten außerhalb des ADT nur über wohldefinierte Operationen zugänglich (Datenkapselung), während die Repräsentation der Daten nur für die Implementierung des ADT relevant ist.

Es gibt zum Beispiel *Stack*, *Queue*, *Priority-Queue* und einfach bzw. doppelt verkettete Listen.

5.2 Binärbaum

Binärbäume enthalten Knoten mit einem linken und einem rechten Nachfolger (die jeweils leer sein können). Es gibt eine Wurzel (ganz oben), innere Knoten mit mindestens einem Nachfolger und Blätter ohne Nachfolger.

Die Tiefe eines Knotens ist der Abstand des Knoten zur Wurzel, die Höhe eines Baumes ist die (maximale) Tiefe des untersten Knotens. Eine Ebene d enthält maximal 2^d Knoten, ein Binärbaum der Höhe h enthält höchstens $2^{h+1} - 1$ Knoten. Ein Binärbaum ist vollständig, wenn die möglichen Knotenzahl für die Höhe des Baumes voll ausgefüllt ist.

Bei einer Traversierung werden alle Elemente des Baumes einmal besucht, jede Kante wird dabei zweimal durchlaufen. Die Traversierung beginnt und endet an der Wurzel des Baumes. Die Teilbäume der Knoten werden in einer festen Reihenfolge (zuerst linker, dann rechter Teilbaum) besucht.

Es gibt verschiedene Traversierungen, die sich alleine dadurch unterscheiden, wann sie die Elemente des Knotens „besuchen“. So gibt es Inorder-, Preorder- und Postorder-Traversierung.

5.3 Binäre Suchbäume

Ein binärer Suchbaum (BST) ist ein Binärbaum, der Elemente mit Schlüsseln enthält, wobei der Schlüssel jedes Knotens mindestens so groß ist wie jeder Schlüssel im linken Teilbaum und höchstens so groß ist wie jeder Schlüssel im rechten Teilbaum.

Im binären Suchbaum gibt es neben dem Element des Knotens und einem Verweis auf die Nachfolger auch noch einen Verweis auf den Vaterknoten - dieser ist bei der Wurzel leer.

Die Inorder-Traversierung gibt aufgrund des Aufbaus des Suchbaums alle Schlüssel des Baumes in sortierter Reihenfolge aus.

Kompliziert ist im BST vor allem das Löschen, wenn das zu löschende Element in einem Knoten mit zwei Kindern steht (bei einem Kind wird der Knoten entfernt und der Zeiger vom Vaterknoten auf das Kind gelegt, ein Blatt wird einfach entfernt). Bei zwei Kindern finden wir den Nachfolger des zu löschenden Knotens, entfernen ihn aus seiner ursprünglichen Position und ersetzen den zu löschenden Knoten mit ihm.

Alle Operationen auf einem BST sind linear und $\in \Theta(h)$.

5.4 Rotationen in Bäumen

Bei einer Linksrotation rückt der rechte Nachfolger (N) des Knotens (K) an die Position von K. K wird linker Nachfolger von N, der ehemalige linke Nachfolger von N wird rechter Nachfolger von K. Die Rechtsrotation verläuft analog.

Durch die Rotation bleibt ein BST weiterhin ein BST, auch die Inorder-Traversierung bleibt unverändert. Die Zeitkomplexität der Rotation ist $\in \Theta(1)$.

5.5 AVL-Bäume

Ein AVL-Baum ist ein balanzierter BST, bei dem sich für jeden Knoten die Höhen der beiden Teilbäume um höchstens 1 unterscheiden - die Höhe der Teilbäume wird balanziert. Es gibt ein zusätzliches Datenfeld für jeden Knoten, in dem über die Höhe des Unterbaums Buch geführt wird. Nach jeder Operation muss die Bilanz des AVL-Baums wiederhergestellt werden, dies ist in $\Theta(h)$ möglich. Dadurch ist stets $h \in \Theta(\log(n))$ gewährleistet, $\Theta(\log(n))$ kann für alle Operationen auf dem Baum garantiert werden.

Es kann - wie bereits angerissen - passieren, dass der AVL-Baum nach dem Einfügen eines Knotens unbalanziert ist und die Bedingungen damit verletzt sind. Dann muss der Baum neu balanziert werden. Dafür gibt es folgende Regeln:

- Wenn das neue Blatt außen am Baum hängt, reicht eine einfache Rotation nach innen um den Vaterknoten.
- Wenn das neue Blatt innen am Baum hängt, muss eine Doppelrotation durchgeführt werden. Dazu rotieren wir erst um den Vaterknoten nach außen, sodass sich dann eine Situation wie im ersten Fall ergibt. Es wird nun um den neuen Vaterknoten nach innen rotiert.

Auch durch Löschen (von Blättern) kann der AVL-Baum unbalanziert werden, dann muss ebenfalls durch Rotieren Abhilfe geschaffen werden.

5.6 Rot-Schwarz-Bäume (RBT)

In Rot-Schwarz-Bäumen gibt es zwei Arten von Knoten: Schwarze Knoten werden strikt balanziert, während rote Knoten nur als „Schlupf“ dienen. Dabei muss die Anzahl der Schlupfknoten beschränkt bleiben. Die Balanzierungsanforderungen an die AVL-Bäume werden damit aufgeweicht, außerdem reduziert sich der benötigte Speicherplatz für zusätzliche Informationen auf 1 Bit (die Farbe), da der Balanzierungsfaktor nicht mehr gespeichert wird.

Die Wurzel eines RBT ist immer schwarz - jeder Knoten muss entweder rot oder schwarz sein. Außerdem darf jeder rote Knoten nur schwarze Kinder haben. Externe Knoten (Null-Zeiger, also quasi nicht existente Knoten) sind immer schwarz. Für jeden Knoten müssen alle Pfade, die an diesem Knoten starten und in einem externen Knoten enden, die gleiche Anzahl schwarzer Knoten enthalten. In Zeichnungen werden die externen Knoten häufig weggelassen.

Die Schwarzhöhe $bh(x)$ eines Knotens x ist die Anzahl schwarzer Knoten bis zu einem (externen) Blatt (x zählt nicht mit, das Blatt/der externe Knoten schon), die Schwarzhöhe $bh(t)$ eines RBT t ist die Schwarzhöhe seiner Wurzel.

Ein Rot-Schwarz-Baum t mit der Schwarzhöhe $h = bh(t)$ mindestens $2^h - 1$ und höchstens $4^h - 1$ innere Knoten - das sind alle „internen“ Knoten.

Neue Knoten werden immer als rote Knoten eingefügt, weil das Einfügen von schwarzen Knoten im Regelfall zur Verletzung der Schwarz-Höhen-Bedingung führt. Die Position für das Einfügen suchen wir uns wie in einem BST. Allerdings kann durch das Einfügen eines roten Knotens eine Rot-Rot-Verletzung auftreten, wenn bereits der Vater des neu einzufügenden Knotens rot ist:

1. Wenn der Onkel des Vaters ebenfalls rot ist, können wir den Vater und den Onkel schwarz sowie den Großvater rot färben, um die Rot-Rot-Verletzung zu beheben. Eventuell entstehen dadurch überhalb des Vaters erneut Rot-Rot-Verletzungen die iterativ genauso behoben werden müssen. Sofern die Wurzel durch der Vertauschung rot gefärbt wurde, können wir diese einfach wieder schwarz färben - die Schwarz-Höhe des Baumes erhöht sich dadurch um eins, eine Schwarz-Höhen-Verletzung tritt dadurch aber nicht auf.
2. Wenn der Onkel des Vaters nicht rot ist und das neue Blatt im Baum innen hängt, müssen wir über den Vater nach außen rotieren. Dann haben wir Fall 3.
3. Wenn der Onkel des Vaters nicht rot ist und das neue Blatt im Baum außen hängt, muss zunächst nach rechts um den Großvater rotiert werden. Sodann wird der ehemalige Großvater rot gefärbt, damit die Schwarz-Höhe von $(x + 1)$ wieder auf x korrigiert wird. Der neue Vater wird dann schwarz gefärbt.

Auch das Löschen von schwarzen Knoten in einem RBT ist problematisch, weil dadurch in der Regel eine Schwarz-Höhen-Verletzung entsteht - hingegen ist das Löschen von roten Knoten unproblematisch. Auch hier gibt es mehrere Fälle:

1. Hat der zu löschende Knoten zwei Kinder, suchen wir seinen Nachfolger und entfernen ihn aus dem Rot-Schwarz-Baum. Eine eventuell neu auftretende Farbverletzung ist dabei zu beheben. Nun ersetzen wir den zu löschenden Knoten durch seinen Nachfolger und übernehmen die Farbe des Nachfolgers.
2. Hat der zu löschende Knoten nur ein Kind, kann der Knoten gelöscht werden. Anschließend kann die Farbverletzung durch einfaches Vertauschen der Farben oder die „Weitergabe“ des Schwarzwertes Richtung Wurzel behoben werden.
3. Hat der zu löschende Knoten keine Kinder, muss der Onkel und der Vater unter Beachtung der Schwarzhöhe umgefärbt werden. Die eventuell dadurch entstehenden Farbverletzungen müssen iterativ behoben werden.

Die Laufzeit der Algorithmen für Einfügen und Löschen liegt in $\Theta(\log(n))$, alle anderen Operationen sind mit denen in einem BST algorithmisch und bezüglich der Laufzeit identisch.

6 Hashing

6.1 Dictionary

Ein *Dictionary* (auch *assoziatives Array*) speichert Daten, die jederzeit anhand ihres Schlüssels abgerufen werden können. Die Daten sind dabei dynamisch gespeichert.

6.2 Direkte Adressierung

Bei der direkten Adressierung gibt es ein Array (die Direkte-Adressierungs-Tabelle) so, dass es für jeden möglichen Schlüssel eine entsprechende Position gibt. Jedes Array-Element enthält dabei einen Pointer auf die gespeicherte Information.

6.3 Grundlagen des Hashings

Ein Hash ist ein Algorithmus, der aus einer großen Datenmenge eine sehr kleine Zusammenfassung (Identifikation; Fingerabdruck) berechnet. Dabei ist er klar definiert und einfach sowie ressourcenschonend zu berechnen.

Das Ziel von Hashing ist es, einen extrem großen Schlüsselraum auf einen vernünftigen, kleinen Bereich von ganzen Zahlen abzubilden. Dabei soll es möglichst unwahrscheinlich sein, dass zwei Schlüssel auf die gleiche Zahl abgebildet werden. In der Praxis wird hierbei oft nur ein kleiner Teil der Schlüssel verwendet, damit ist bei der direkten Adressierung ein Großteil des Arrays verschwendet. $h(k)$ sei der Hashwert von k . Das Auftreten von $h(k) = h(k')$ für $k \neq k'$ nennt man Kollision.

6.4 Hashfunktionen

Eine gute Hash-Funktion ist einfach zu berechnen, surjektiv auf der Zielmenge und verwendet alle Indizes mit möglichst gleicher Häufigkeit. Ferner sollen ähnliche Schlüssel möglichst breit auf die Hashtabelle verteilt werden. Es gibt einige Basistechniken, um eine gute Hashfunktion zu erhalten:

Bei der Divisionsmethode wird die Hashfunktion mit $h(k) = k \bmod m$ gebildet. Dabei muss man m sorgfältig wählen, da zum Beispiel bei $m = 2^p$ einfach die letzten p Bits genommen werden. Eine gute Wahl ist m prim und nicht zu nah an einer Zweierpotenz.

Die Multiplikationsmethode stellt $h(k) = \lfloor m \cdot (k \cdot c \bmod 1) \rfloor$ mit $0 < c < 1$. Der Parameter m ist dabei nicht kritisch, Knuth empfiehlt $c = \frac{(\sqrt{5}-1)}{2} \approx 0,618$.

6.5 Universelles Hashing

Das größte Problem beim Hashing ist, dass es immer eine ungünstige Sequenz von Schlüsseln gibt, die auf den gleichen Hashwert abgebildet werden. Die Idee zur Lösung dieses Problem ist, zufällig eine Hashfunktionen aus einer gegebenen, kleinen Menge H auszuwählen - unabhängig von den verwendeten Schlüsseln.

6.6 MD5

MD5 übersetzt Nachrichten beliebiger Länge in einen 128-Bit-Hashwert. Inzwischen gilt MD5 als sehr unsicher, Kollisionen sind verhältnismäßig einfach zu berechnen.

6.7 Verkettung

Bei der Verkettung wird in der Hash-Tabelle an der Stelle des gehashten Schlüssels nicht nur der Wert gespeichert, sondern auch der Original-Schlüssel (vor dem Hashen). Die Speicherung erfolgt in der Hash-Tabelle als verkettete Liste. Dadurch erhöht sich natürlich den Speicheraufwand, außerdem muss beim Speichern und Abrufen eine Suche durchgeführt werden.

6.8 Offene Adressierung

Bei der offenen Adressierung werden alle Schlüssel direkt an einer Position in der Hashtabelle gespeichert (im Gegensatz zur Verkettung). Wenn ein Schlüssel eingefügt werden soll, wird geprüft, ob die entsprechende Position in der Hashtabelle frei ist. Ist dies nicht der Fall, kann der Schlüssel nicht (wie bei der Verkettung) einfach angehängt werden. Es muss dann - in Abhängigkeit des einzufügenden Schlüssels - nach einer freien Position gesucht werden.

Damit das funktioniert, bekommt die Hashfunktion $h(k)$ nun einen zweiten Parameter, der in $[0, m-1]$ (m ist die Größe der Hashtabelle) liegt. Dieser Parameter gibt die Nummer der Sondierung (das ist die Nummer des Versuchs, eine freie Position zu finden) an. Die Hashfunktion liefert nun in der Regel eine andere Position, die wir in der Hashtabelle prüfen. Sobald wir eine leere Position finden, wird der Schlüssel dort in der Hashtabelle eingefügt und wir sind fertig. Finden wir keine leere Position, führen wir das Verfahren iterativ durch, erhöhen also den zweiten Parameter der Hashfunktion in jedem Durchlauf.

Über den eindeutigen Index des Schlüssels in der Hashtabelle kann man dann mithilfe einer Direkten-Adressierungs-Tabelle auf das eigentlich assoziativ zu speichernde Element verweisen.

Die Suche in der Hashtabelle nach einem Schlüssel verläuft ähnlich - wir gehen wieder nach dem Schema durch die Tabelle und sind fertig, sobald wir den gesuchten Schlüssel gefunden haben. Es wird also nicht geprüft, ob die jeweilige Position leer ist, sondern ob der Inhalt der Hashtabelle an dieser Position dem gesuchten Schlüssel entspricht.

Aufpassen müssen wir beim Löschen eines Elements. Man könnte einfach wieder die Position des Elements suchen und dann die Hashtabelle an dieser Stelle leeren. Das führt aber zu dem Problem, dass alle eventuell nach diesem Element eingefügte Elemente nicht mehr gefunden werden, da die Sondierung früher abgebrochen wird. Daher ersetzen wir das Element beim Löschen durch z.B. "DELETED". Dieser Wert wird bei der Sondierung als leer interpretiert - beim Einfügen wird der Slot wiederverwendet, bei der Suche übersprungen.

Dadurch sind die Suchzeiten nicht mehr allein vom Füllgrad der Hashtabelle abhängig, deswegen verwendet man zur Kollisionsauflösung meist Verkettung, wenn Schlüssel auch gelöscht werden sollen.

6.9 Sondierung

Wir gehen nun näher auf das Verfahren der Sondierung ein. Für jeden Schlüssel k benötigen wir eine Sondierungssequenz $\langle h(k, 0), h(k, 1), \dots, h(k, m-1) \rangle$, um die offene Adressierung nutzen zu können. Dabei sollte die Sondierungssequenz eine Permutation von $\langle 0, \dots, m-1 \rangle$ sein, damit jeder Slot auch wirklich verwendet wird. Ideal wäre gleichverteiltes Hashing, sodass jede der $m!$ Permutationen als Sondierungssequenz gleich wahrscheinlich ist. In der Praxis ist das aber zu aufwändig, daher wird

approximiert.

Lineares Sondieren

$h(k, i) = (h'(k) + i) \bmod m$ ist eine Hashfunktion für lineares Sondieren. Dabei ist h' eine normale Hashfunktion (siehe 6.4). Die Verschiebung der nachfolgenden Sondierungen hängt damit linear von i ab. Damit wird die gesamte Sequenz bereits durch die erste Sondierung bestimmt und es gibt nur m verschiedene Sequenzen.

Das führt zum Problem des Clusterings: Lange Folgen von belegten Slots tendieren dazu, noch länger zu werden, weil $h'(k)$ konstant bleibt und nur der Offset jedes Mal um eins größer wird.

Quadratisches Sondieren

$h(k, i) = (h'(k) + c_1 \cdot i + c_2 \cdot i^2) \bmod m$ ist eine Hashfunktion für quadratisches Sondieren. Dabei ist h' eine normale Hashfunktion (siehe 6.4) und $c_1, c_2 \in \mathbb{N}$ sind geeignete Konstanten. Die Verschiebung der nachfolgenden Sondierungen hängt damit quadratisch von i ab. Damit wird die gesamte Sequenz auch bei der quadratischen Sondierung bereits durch die erste Sondierung bestimmt und es gibt nur m verschiedene Sequenzen.

Das Problem des Clusterings wird vermieden oder zumindest abgeschwächt. Allerdings tritt immer noch sekundäres Clustering auf, denn $h(k, 0) = h(k', 0)$ verursacht $h(k, i) = h(k', i)$ für alle i .

Doppeltes Hashing

$h(k, i) = (h_1(k) + i \cdot h_2(k)) \bmod m$ ist eine Hashfunktion für doppeltes Hashing. Dabei sind h_1 und h_2 normale Hashfunktionen (siehe 6.4). Die Verschiebung der nachfolgenden Sondierungen hängt damit nur von h_2 ab. Damit wird die gesamte Sequenz nicht bereits durch die erste Sondierung bestimmt. Deswegen werden die Schlüssel besser in der Hashtabelle verteilt und es wird eine Approximierung von gleichverteiltem Hashing erzielt.

Sind h_2 und m relativ prim, wird die gesamte Hashtabelle abgesucht. Man kann zum Beispiel $m = 2^k$ wählen und h_2 nur ungerade Zahlen erzeugen lassen, um ein gutes Ergebnis zu erzielen. Dann können m^2 verschiedene Permutationen erzeugt werden, weil jedes mögliche Paar von $h_1(k)$ und $h_2(k)$ eine andere Sequenz erzeugt.

7 Graphalgorithmen

7.1 Graphen

Graphen haben wir in der Vorlesung *Diskrete Strukturen* definiert, diese bestehen aus dem Paar (V, E) (Ecken/Knoten und Kanten). Die Kanten können gerichtet sein (eine Richtung haben) oder nicht - wir haben dann einen gerichteten bzw. ungerichteten Graphen. Einen gerichteten Graphen nennt man auch Digraph.

Der Knoten u ist adjazent zu Knoten v gdw. $(u, v) \in E$ und inzident zu der Kante $e = (u, v)$. Der Grad $\deg(u)$ ist die Anzahl der zu u inzidenten Kanten. Wir definieren - insbesondere für gerichtete Graphen - ferner den In-Grad $\deg^+(u)$ und den Aus-Grad $\deg^-(u)$. Ein Graph ist k -regulär, wenn alle Knoten den Grad k haben.

In ungerichteten Graphen $G = (V, E)$ gilt das Handshake-Lemma $\sum_{v \in V} \deg(v) = 2 \cdot |E|$, bei gerichteten Graphen ist $\sum_{v \in V} \deg^+(v) = \sum_{v \in V} \deg^-(v)$.

Ein Teilgraph eines Graphen $G = (V, E)$ ist ein Graph $G' = (V', E')$ mit $V' \subseteq V$ und $E' \subseteq E \cap (V' \times V')$. Ist $V' \subset V$ und/oder $E' \subset E$, so nennen wir G' einen echten Teilgraphen von G . Wenn $E' = E \cap (V' \times V')$ ist, so ist G' der durch V' induzierte Teilgraph.

Ein gerichteter Graph G heißt symmetrisch, wenn aus $(v, w) \in E$ immer auch $(w, v) \in E$ folgt. Zu jedem ungerichteten Graphen gibt es einen korrespondierenden symmetrischen, gerichteten Graphen. Der Graph G heißt vollständig, wenn jedes Paar von Knoten mit einer Kante verbunden ist. Es ist K_n der ungerichtete, vollständige Graph mit n Knoten.

Ferner können wir einen Graphen **transponieren**, der transponierte Graph von $G = (V, E)$ ist gegeben durch $G^T = (V, E')$ mit $(v, w) \in E'$ gdw. $(w, v) \in E$.

7.2 Pfade und Kreise

Ein Spaziergang (walk) von einem Knoten v zu einem Knoten w ist eine Folge v_0, v_1, \dots, v_k mit dem Startknoten $v_0 = v$ und dem Endknoten $v_k = w$. Ferner muss (v_i, v_{i+1}) für $0 \leq i \leq k-1$ eine Kante des Graphen sein.

Hat der Spaziergang nur paarweise verschiedene Knoten (d.h. $v_i \neq v_j$ für $i \neq j$), nennt man ihn simpel/einfach oder Pfad. Die Länge eines Spaziergangs ist die Anzahl der durchlaufenen Kanten.

Ein Kreis ist eine Spezialisierung des Pfades, dieser darf nicht-leer sein und es muss $v_0 = v_k$ sein - der Spaziergang muss also an der Stelle wieder enden, an der er begonnen hat, wir machen also eine Rundwanderung. Ein Graph ist kreisfrei oder azyklisch, wenn er keine Kreise als Teilgraphen hat.

Ein Knoten w ist vom Knoten v aus erreichbar, wenn es einen Pfad von v nach w gibt. Ein ungerichteter Graph G ist zusammenhängend, wenn jeder Knoten von jedem anderen Knoten aus erreichbar ist. Eine Zusammenhangskomponente (connected component) von G ist ein maximaler zusammenhängender Teilgraph von G .

In einem gerichteten Graphen G' differenzieren wir genauer. G' ist stark zusammenhängend (strongly connected), wenn jeder Knoten von jedem anderen aus erreichbar ist. Hingegen ist G' schwach zusammenhängend, wenn der zugrunde liegende ungerichtete Graph (in dem alle Kanten ungerichtet gemacht worden sind) zusammenhängend ist. Eine starke Zusammenhangskomponente von G' ist ein maximaler stark zusammenhängender Teilgraph von G' .

Jeder ungerichtete Graph kann eindeutig in Zusammenhangskomponenten, jeder gerichtete Graph eindeutig in starke Zusammenhangskomponenten aufgeteilt werden.

Ein Baum ist ein kreisfreier und zusammenhängender Graph. Hat ein Baum n Knoten, so hat er $n-1$ Kanten. Ein ungerichteter Graph mit n Knoten und $n-2$ oder weniger Kanten kann nicht zusammenhängend sein, während ein ungerichteter Graph mit n Knoten und n oder mehr Kanten einen Zyklus enthalten muss - in beiden Fällen kann dieser Graph also kein Baum sein.

7.3 Repräsentation von Graphen

Es stellt sich die Frage, wie man einen Graphen als Datenstruktur am sinnvollsten repräsentiert. Hierzu gibt es verschiedene Möglichkeiten, wir machen uns die Adjazenz zu Nutzen.

Adjazenzmatrix

Wir haben einen Graphen $G = (V, E)$ mit $|V| = n$, $|E| = m$ und $V = v_1, \dots, v_n$. Dann ist der Graph in der Adjazenzmatrix-Darstellung gegeben durch eine $n \times n$ -Matrix A mit $A(i, j) = 1$ wenn

$(v_i, v_j) \in E$ (und 0 wenn nicht). Wenn G ungerichtet ist, ist A symmetrisch ($A^{tr} = A$), dann muss nur die Hälfte der Matrix gespeichert werden. Der Platzbedarf ist $\Theta(n^2)$.

Adjazenzliste

Wir können einen Graphen auch als Array von Adjazenzlisten darstellen. Dabei ist die Arrayposition gegeben durch die Nummer des Knotens i , an dieser Stelle enthält das Array eine verkettete Liste mit allen Kanten mit Startknoten v_i . Wenn G ungerichtet ist, wird jede Kante zweimal gespeichert. Der Platzbedarf ist $\Theta(n + m)$.

7.4 Graphendurchlauf

Viele Algorithmen untersuchen jeden einzelnen Knoten und jede einzelne Kante eines Graphen. Dazu gibt es verschiedene Graphendurchlaufstrategien, die jeden Knoten (oder jede Kante) genau einmal besuchen: Die Breitensuche (BFS) und die Tiefensuche (DFS). Dabei handelt es sich im Wesentlichen um Verallgemeinerungen der Strategien zur Baumtraversierung.

Allerdings müssen wir uns alle bereits besuchten Knoten explizit merken. Wir arbeiten mit der Adjazenzlisten-Darstellung, Algorithmen auf dieser Basis kosten $\mathcal{O}(|V| + |E|)$ Zeit.

Breitensuche

Bei der Breitensuche (Breadth-First Search, BFS) sind anfangs alle Knoten als unbesucht (WHITE) markiert. Es wird in einem beliebigen Knoten v begonnen. Der aktuelle Knoten wird immer als aktiv (GRAY) markiert. Sodann wird für jede Kante (v, w) mit unbesuchtem Nachfolger w von allen Kanten dieses Nachfolgers aus weitergesucht. Insbesondere gibt es keinerlei Backtracking. Der Knoten v wird dann als besucht (BLACK) markiert. Damit sind die schwarzen Knoten dann genau jene Knoten, die vom Startknoten aus erreichbar sind.

Für einen Knoten $v \in V$ bezeichnet $d(v)$ den Abstand vom Startknoten zum Knoten v , also die Anzahl der Kanten auf dem kürzesten Weg vom Startknoten nach v . Wenn ein Knoten w in die Queue gegeben wird, so färben wir die dafür verantwortliche Kante (v, w) gelb. Dann ist v der Vater vom Knoten w .

BFS besucht die Knoten also in einer Reihenfolge mit ansteigendem Abstand vom Startknoten. Erst wenn alle Knoten mit Abstand d verarbeitet sind, werden die Knoten mit Abstand $d+1$ angegangen. Die Suche terminiert, wenn für einen Abstand d keine Knoten auftreten. Die zu verarbeitenden Knoten werden als FIFO-Queue organisiert, es gibt nur eine einzige Verarbeitungsmöglichkeit für v (wenn es aus der Queue entnommen wird). Die gelben Kanten induzieren den Breitensuchbaum, der den Startknoten als Wurzel hat.

Die Zeitkomplexität der BFS ist $\mathcal{O}(|V| + |E|)$, der Platzbedarf ist $\Theta(|V|)$.

Tiefensuche

Auch bei der Tiefensuche (Depth-First Search, DFS) markieren wir zunächst alle Knoten als unbesucht (WHITE). Der aktuelle Knoten v wird immer als aktiv (GRAY) markiert, es ist beliebig, mit welchem Knoten wir beginnen. Für jede Kante (v, w) mit unbesuchtem Nachfolger w suchen wir nun rekursiv von w aus. Das bedeutet also, dass wir den neu entdeckten Knoten w besuchen und von dort aus forschen, bis es nicht mehr weiter geht. Sodann gehen wir von w nach v zurück. Für jede Kante (v, w) , für die der Nachfolger w schon besucht wurde, überprüfen wir die Kante, besuchen

w aber nicht. Abschließend markieren wir den Knoten v als besucht (BLACK). Damit sind auch bei der DFS die schwarzen Knoten diejenigen, die vom Startknoten aus erreichbar sind.

Die Tiefensuche erforscht also einen Pfad so weit wie möglich und setzt dann zurück. Die zu verarbeitenden Knoten werden also in LIFO-Reihenfolge verarbeitet. Es gibt zwei mögliche Verarbeitungszeitpunkte für jeden Knoten - wenn dieser entdeckt oder wenn er abgeschlossen wird.

Bei der Tiefensuche gibt es verschiedene Klassen von Kanten. Baum-Kanten treten im DFS-Baum auf. Rückwärts-Kanten gehen von Knoten u zum Vorfahren v , Vorwärts-Kanten gehen von einem Vorfahren v zu einem Knoten u . Alle restlichen Kanten sind Quer-Kanten. In einem ungerichteten Graphen ist jede Kante entweder Baum-Kante oder Rückwärts-Kante.

Die Zeitkomplexität der DFS ist $\mathcal{O}(|V| + |E|)$, der Platzbedarf ist $\Theta(|V|)$.

7.5 Zusammenhangskomponenten in ungerichteten Graphen

Was tun wir, wenn wir alle Zusammenhangskomponenten (SCCs) eines ungerichteten Graphen bestimmen möchten?

- Konstruktion des dazugehörigen symmetrischen, gerichteten Graphen
- Start der Tiefensuche bei einem beliebigen Knoten
- Suche aller anderen Knoten (und aller Kanten) in der selben Komponente mit DFS
- Wenn es weitere, unbesuchte Knoten gibt, Wiederholung des Verfahrens mit diesen

7.6 Zusammenhangskomponenten in gerichteten Graphen

Wiederholung (7.2): Es sei G ein gerichteter Graph. G heißt stark zusammenhängend, wenn jeder Knoten von jedem anderen Knoten aus erreichbar ist. Eine starke Zusammenhangskomponente von G ist ein maximaler stark zusammenhängender Teilgraph von G . Der transponierte Graph von G ist G^T , hier ist die Richtung aller Kanten von G umgedreht.

Die starken Zusammenhangskomponenten eines gerichteten Graphen induzieren den dazugehörigen **Kondensationsgraphen**. Diesen bilden wir, indem wir jede Zusammenhangskomponente quasi zu einem Knoten zusammenfassen und die Kanten zwischen den Zusammenhangskomponenten übernehmen. Der Kondensationsgraph $G \downarrow$ ist azyklisch.

Aus der Vorlesung stammt die folgende Definition für den Kondensationsgraphen, die zunächst wegen der Definition von V' verwirrt: Es sei $G = (V, E)$ ein gerichteter Graph mit k starken Zusammenhangskomponenten $S_i = (V_i, E_i)$ für $1 \leq i \leq k$. Der Kondensationsgraph $G \downarrow = (V', E')$ ist dann definiert durch $V' = \{V_1, \dots, V_k\}$ (es gibt für jede Zusammenhangskomponente einen Zustand, der quasi nach einer Menge benannt ist). Es ist $(V_i, V_j) \in E'$ gdw. $i \neq j$ und $\exists (v, w) \in E : v \in V_i \wedge w \in V_j$.

Die starken Komponenten von G und G^T sind die gleichen. Die Kondensation und die Transposition kommutieren: $(G \downarrow)^T = (G^T) \downarrow$.

7.7 Erreichbarkeitsanalysen

Die Tiefensuche wurde bereits vor Jahrhunderten formal beschrieben als ein Verfahren zum Durchqueren von Labyrinthen. Ein Labyrinth lässt sich dabei als Graph darstellen: Die Stellen, an denen mehr als ein Weg gewählt werden kann, werden als Knoten dargestellt. Die Kanten sind die Wege zwischen den Knoten. Der Eingang ist ein spezieller Knoten *start*, das Ziel ein spezieller Knoten *target*.

Das Problem (die Suche nach einem Weg von einem beliebigen Knoten zum Ausgang) lässt sich mit einer Tiefensuche lösen.

7.8 Kosaraju-Sharir-Algorithmus

Der Kosaraju-Sharir-Algorithmus findet starke Komponenten in einem gerichteten Graphen in zwei Phasen:

1. Führe DFS auf G durch. Speichere jeden Knoten beim Abschließen (d. h. wenn der Knoten BLACK gefärbt wird) auf einem Stack S .
2. Färbe alle Knoten wieder auf WHITE zurück.
3. Wiederhole, solange S noch weiße Knoten enthält. Wähle den obersten noch weißen Knoten v vom Stack S . Führe DFS mit Startknoten v auf transponiertem Graphen G^T aus. Speichere für jeden besuchten Knoten diesen *Leiterknoten* v als Repräsentanten seiner SCC.

Ein Knoten v heißt „Leiter“ (lt. Vorlesung, besser „Anführer“, da engl. leader lt. VL), wenn er als letzter Knoten in seiner SCC bei der DFS BLACK gefärbt wird.

Die Worst-Zeit-Komplexität des Kosaraju-Sharir-Algorithmus ist $\Theta(|V| + |E|)$, die Speicherkomplexität ist $\Theta(|V|)$.

7.9 Gerichtete azyklische Graphen (DAGs)

Gerichtete kreisfreie Graphen (directed acyclic graph, DAG) sind eine wichtige Klasse von Graphen. Viele Probleme lassen sich mit Hilfe von DAGs formulieren und haben auf DAGs eine niedrigere Komplexität als auf Digraphen. Ein DAG hat eine **partielle Ordnung** $<$ auf den Knoten.

Eine Kante (v, w) besagt $v < w$. Da eine partielle Ordnung nach der Vorlesung *Diskrete Strukturen* transitiv und anti-symmetrisch ist, kann sie keine Kreise enthalten.

7.10 Topologisches Sortieren

Es sei $G = (V, E)$ ein gerichteter Graph mit n Knoten. Eine topologische Ordnung von G ist eine Nummerierung der Knoten. $\text{topo}: V \rightarrow \{1, \dots, n\}$, sodass $\text{topo}(v) > \text{topo}(w)$ für jede Kante $(v, w) \in E$ gilt. Eine topologische Ordnung ist die Einbettung einer partiellen Ordnung in eine totale (lineare) Ordnung.

Für einen Digraphen G mit einem Kreis existiert keine topologische Ordnung. Jeder DAG hat mindestens eine topologische Ordnung.

Eine topologische Ordnung kann in $\Theta(|V| + |E|)$ bestimmt werden.

7.11 Gewichtete Graphen

Ein knotengewichteter Graph G ist ein Tripel (V, E, w) , wobei (V, E) ein Graph ist. Zusätzlich ist $w : V \rightarrow \mathbb{R}$ die Gewichtsfunktion und $w(v)$ mit $v \in V$ das Gewicht des Knotens v .

Ein kantengewichteter Graph G ist ein Tripel (V, E, w) , wobei (V, E) ein Graph ist. Zusätzlich ist $w : E \rightarrow \mathbb{R}$ die Gewichtsfunktion und $w(e)$ mit $e \in E$ das Gewicht der Kante e .

Wir können einen knotengewichteten Graphen (V, E, w) einfach in einen kantengewichteten Graphen (V, E, w') überführen, indem alle von einem Knoten v ausgehenden Kanten $e = (v, -) \in E$ das Gewicht $w'(e) = w(v)$ erhalten.

Auch gewichtete Graphen können wir durch eine Adjazenzliste oder -matrix darstellen. Bei knotengewichteten Graphen speichern wir die Zusatzinformation zu den Knoten üblicherweise in einem weiteren Array. Kantengewichte können bei der Matrixdarstellung direkt in der Matrix gespeichert werden (ein besonderer Wert, etwa ∞ , besagt, dass keine Kante existiert).

7.12 Kritische-Pfad-Problem

Als Gewicht eines Pfades bezeichnen wir die Summe der Kantengewichte der besuchten Kanten bei kantengewichteten Graphen und die Summe der Knotengewichte der besuchten Knoten bei knotengewichteten Graphen. Wir sollen nun den längsten Pfad in einem gewichteten DAG finden - das ist das Kritische-Pfad-Problem.

7.13 Eulersche Kreise

Ein eulerscher Kreis in einem Graphen $G = (V, E)$ ist ein Kreis, der jede Kante genau einmal durchläuft.

Ein ungerichteter Graph G besitzt einen eulerschen Kreis gdw. jeder Knoten einen geraden Grad hat und wenn G zusammenhängend ist.

Ein gerichteter Graph G besitzt einen eulerschen Kreis gdw. der Eingangsgrad jedes Knotens dem Ausgangsgrad entspricht und G (stark oder schwach) zusammenhängend ist.

Wir können einen Eulerkreis in einem Graphen $G = (V, E)$ in $\mathcal{O}(|V| + |E|)$ Zeit berechnen.

8 Minimale Spannbäume

8.1 Spannbäume

Ein Spannbaum eines ungerichteten, zusammenhängenden Graphen $G = (V, E)$ ist ein Teilgraph von G , der ein ungerichteter Baum ist und alle Knoten von G enthält. Übrigens ist ein Baum maximal kreisfrei und minimal zusammenhängend. Fügt man eine Kante zu einem Baum hinzu, so entsteht ein Kreis. Löscht man eine Kante eines Baumes, so erhält man zwei Zusammenhangskomponenten. Der vollständige Graph mit n Knoten hat n^{n-2} Spannbäume.

8.2 Minimale Spannbäume

Zuerst definieren wir das Gewicht $w(G')$ eines Teilgraphen $G' = (V', E')$ des gewichteten Graphen G . Es ist $w(G') = \sum_{e \in E'} w(e)$. Ein Spannbaum des Graphen G mit minimalem Gewicht heißt minimaler Spannbaum (minimum spanning tree, MST) von G .

Ein Spannbaum von G kann, muss aber nicht eindeutig sein.

8.3 Prüfer-Code

Für Bäume gibt es einen sogenannten *Prüfer-Code*, in diesen können wir einen Baum übersetzen. Um einen Baum in einen Prüfer-Code zu übersetzen, müssen wir iterativ das **Blatt** x mit kleinster Nummer bestimmen, x zusammen mit der inzidenten Kante (x, y) löschen und y im Tupel speichern. Fertig sind wir, wenn nur noch eine Kante übrig ist. Übrigens: Ein Knoten mit Grad d kommt $(d-1)$ -mal im Prüfer-Code vor.

8.4 Prim und Kruskal

Die Algorithmen von Prim und Kruskal finden für einen gewichteten zusammenhängenden Graphen G mit n Knoten einen minimalen Spannbaum von G .

Kruskal's Algorithmus

1. Wähle eine billigste unmarkierte Kante.
2. Markiere sie, falls sie keinen Kreis mit anderen markierten Kanten schließt.
3. Wiederhole, bis $n - 1$ Kanten markiert sind.

Prim's Algorithmus

1. Wähle einen Startknoten.
2. Markiere die billigste vom bereits konstruierten Baum ausgehende Kante, die keinen Kreis mit anderen markierten Kanten schließt.
3. Wiederhole Schritt 2 bis $n - 1$ Kanten markiert sind.

9 Kürzeste Pfade

Nachfolgend geht es um das Kürzeste-Pfade-Problem. Ein Beispiel: Wir haben eine Straßenkarte, auf der der Abstand zwischen jedem Paar von benachbarten Kreuzungen eingezeichnet ist, eine Startkreuzung s und eine Zielkreuzung t . Wie finden wir nun algorithmisch den kürzesten Weg von s nach t ?

9.1 Notation

Wir arbeiten mit einem kantengewichteten Graphen $G = (V, E, W)$. Ein kürzester Pfad vom Knoten $s \in V$ zum Knoten $v \in V$ ist ein Pfad von s nach v mit minimalem Gewicht.

Wir definieren $\delta : (V \times V) \rightarrow (\mathbb{R} \cup \{\infty\})$. Es ist $\delta(s, v)$ das Gewicht des kürzesten Pfades von s nach v und $\delta(s, v) = \infty$, falls v von s nicht erreichbar ist.

9.2 Problemvarianten

Es gibt viele verschiedene Varianten des Kürzeste-Pfade-Problems:

Kürzeste Pfade von einem Startknoten s zu allen anderen Knoten sind Single-Source Shortest Paths (SSSP). Kürzeste Pfade von allen Knoten zu einem Zielknoten t lässt sich auf SSSP zurückführen. Für den kürzesten Pfad für ein festes Knotenpaar (u, v) ist kein Algorithmus bekannt, der asymptotisch schneller ist als der beste SSSP-Algorithmus.

Die kürzesten Pfade für alle Knotenpaare ermitteln wir mit All-Pairs Shortest Paths (APSP).

9.3 Bellman-Ford-Algorithmus

Der Bellman-Ford-Algorithmus berechnet die kürzesten Pfade von einem einzigen Startpunkt aus, ist also ein SSSP-Algorithmus. Dabei erlaubt der Algorithmus auch negative Kantengewichte und kann negative Zyklen feststellen, die vom Startknoten aus erreichbar sind (diese sind problematisch, da man durch unendliches Durchlaufen durch diese Knoten immer kürzere Strecken erreichen und somit keine kürzeste Strecken finden kann) - dann gibt es keine Lösung und der Algorithmus terminiert, sobald ein negativer Kreis gefunden wurde.

Initial ist für den Startknoten (s) $dist[s] = 0$ und für alle anderen Knoten ($v \in V \setminus \{s\}$) $dist[v] = \infty$. Wenn es eine Lösung gibt (also kein negativer Kreis vorhanden ist), verbessert der Algorithmus für jeden Knoten v eine obere Grenze $dist[v]$ für $\delta(s, v)$, bis das Minimum gefunden wurde. Dazu wird für alle Kanten $(v, w) \in E$ iterativ geprüft, ob das bisher bekannte Gewicht $dist[w]$ größer als $dist[v] + W(v, w)$ ist - ist dies der Fall, wird $dist[w]$ auf den neuen Wert verbessert.

Wir sind fertig, wenn keine Verbesserungen mehr möglich sind. Gibt es nach $|V| - 1$ Wiederholungen noch Verbesserungen, so existiert ein negativer Kreis. Ansonsten gilt $dist[v] = \delta(s, v)$ für alle Knoten $v \in V$.

Der Algorithmus speichert zusätzlich die Vorgängerknoten entlang eines kürzesten Pfades im Array *prev*, sodass sich so nach Terminierung die kürzesten Pfade rekonstruieren lassen.

Die Zeitkomplexität des Algorithmus ist gegeben durch $\mathcal{O}(|V| \cdot |E|)$.

9.4 Dijkstra-Algorithmus

Dijkstra arbeitet im Gegensatz zu Bellman und Ford nur mit nicht-negativen Kantengewichten. Es gilt also $\forall (v, w) \in E : W(v, w) \geq 0$. Kürzeste Pfade enthalten damit keine Zyklen.

Wir beginnen mit unserem Startknoten s , dieser hat $dist[s] = 0$. Alle anderen Knoten $v \in V \setminus \{s\}$ haben $dist[v] = \infty$. Der Startknoten wird schwarz gefärbt, er ist Teil unseres (Hilfs-)Baums. Wir betrachten nun alle Randkanten $(s, v) \in E$ und suchen die Kante mit dem geringsten Kantengewicht. Sodann fügen wir den neuen Zielknoten v zum Baum hinzu, indem wir ihn ebenfalls schwarz färben. Wir betrachten nun iterativ die Randkanten unseres Baums und suchen immer wieder die billigste

Kante (dabei muss auf die Summe geachtet werden, es dürfen ferner keine Zyklen entstehen). Entsprechend füllen wir $dist[v]$ für $v \in V$, bis alle Knoten im Baum sind.

Der Dijkstra-Algorithmus findet die kürzesten Wege mit zunehmenden Abstand zur Quelle s , die Implementierung ist ähnlich zum Spannbaum-Algorithmus von Prim (8.4). Die Zeitkomplexität im Worst-Case ist $\in \Theta(|V|^2)$, ferner ist die untere Schranke $\Omega(|E|)$. Wir haben eine Platzkomplexität von $\mathcal{O}(|V|)$.

9.5 All-Pairs Shortest Paths

Wir betrachten gewichtete, gerichtete Graphen $G = (V, E, W)$. Dabei sind negative Gewichte zwar erlaubt, aber keine Kreise mit negativen Gewicht. Nicht vorhandene Kanten haben das Gewicht $W(-, -) = \infty$. Wir möchten nun für jedes Paar (i, j) das Gewicht $D[i, j]$ des kürzesten Pfades berechnen.

Man kann nun einen SSSP-Algorithmus (z.B. Bellman-Ford) $|V|$ -mal anwenden, das führt allerdings zu einer Worst-Case-Zeitkomplexität von $\mathcal{O}(|V|^4)$. Im Folgenden werden wir effizientere Algorithmen für das APSP-Problem kennenlernen.

9.6 Binäre Relationen

Eine (binäre) Relation R über einer Menge S ist eine Teilmenge $R \subseteq S \times S = S^2$.

Aus der Vorlesung *Diskrete Strukturen* kennen wir die Reflexivität ($\forall u \in S : (u, u) \in R$) und die Transitivität ($((u, v) \in R \wedge (v, w) \in R \Rightarrow (u, w) \in R)$). Die transitive Hülle R^* einer Relation R ist die kleinste Erweiterung (Obermenge) $R \subseteq R^* \subseteq S^2$, sodass R^* reflexiv und transitiv ist.

Für einen Graphen mit der Menge $S = V$ und der Relation $R = E$ gilt: Es ist $(u, v) \in R^*$ gdw. ein Pfad von u nach v existiert.

9.7 Warshall-Algorithmus

Der Warshall-Algorithmus ist die Basis für den Floyd-Algorithmus. Er berechnet die transitive Hülle eines Graphen. Dazu wird zuerst die Reflexivität erstellt - im Graphen wird so bei jedem Knoten eine Schleife hinzugefügt (sofern noch keine vorhanden ist), in der Adjazenzmatrix wird die Diagonale mit Einsen gefüllt.

Dann schauen wir für jeden Knoten einzeln, welche Knoten direkt von diesem aus erreichbar sind. Das machen wir rekursiv für die erreichbaren Knoten und fügen die entsprechenden direkten Kanten vom ursprünglichen Knoten zum über den neuen Knoten erreichbaren Knoten hinzu bzw. schreiben die entsprechende Eins in die Matrix.

Nachdem wir das für alle Knoten gemacht haben (in der Regel machen wir das in aufsteigender Reihenfolge der Knotennummerierung), haben wir die transitive Hülle als Graph bzw. ihre Adjazenzmatrix.

Die Zeitkomplexität des Warshall-Algorithmus ist $\Theta(|V|^3)$, die Platzkomplexität ist $\Theta(|V|^2)$.

9.8 Floyd-Algorithmus

Es gelten die Voraussetzungen aus 9.5. Der Floyd-Algorithmus greift auf den Warshall-Algorithmus zurück, allerdings werden hier die Distanzen berücksichtigt (statt der Berechnung der linearen Hülle).

Deshalb wird er auch als Floyd-Warshall-Algorithmus bezeichnet.

Das Vorgehen ist wie bei Warshall, allerdings nutzen wir folgende Rekursionsgleichung für die Distanzen: Es ist $d_{ij}^{(0)} = W(i, j)$ und $d_{ij}^{(k)} = \min\{d_{ij}^{(k-1)}, d_{ik}^{(k-1)} + d_{kj}^{(k-1)}\}$ für $k > 0$.

10 Matching

10.1 Definition

Es sei $G = (V, E)$ ein ungerichteter Graph. Eine Kantenmenge $M \subseteq E$ heißt **Matching**, falls je zwei verschiedene Kanten in M disjunkt sind (für alle $e, e' \in M$ mit $e \neq e'$ gilt $e \cap e' = \emptyset$). Ein Knoten $v \in V$ wird von einem Matching M überdeckt, falls $v \in e$ für eine Kante $e \in M$ gilt.

Wir bezeichnen einen Graphen $G = (V, E)$ als **bipartit**, falls es eine Partition $V = X \cup Y$ der Knotenmenge gibt, sodass $E \subseteq X \times Y$.

Für einen ungerichteten Graphen $G = (V, E)$ suchen wir oft ein **Matching** $M \subseteq E$ mit maximaler Kardinalität. Ist der Graph G bipartit, bezeichnen wir das Matching auch als **bipartites Matching**.

10.2 Algorithmus für bipartites Matching

In der Vorlesung wurde ein Algorithmus für bipartites Matching entwickelt. Dabei beginnen wir mit dem leeren Matching $M = \emptyset$. Das Matching lassen wir Schritt für Schritt wachsen, wobei jeder Schritt die Kardinalität $|M|$ um 1 erhöht. Wenn wir stecken bleiben, ist M ein Matching maximaler Kardinalität.

Wir können ein Matching mit maximaler Kardinalität in $\mathcal{O}(|V| \cdot |E|)$ Zeit berechnen, denn ein einzelner augmentierender Pfad kostet $\mathcal{O}(|V| + |E|)$ Zeit und es muss höchstens $\frac{|V|}{2}$ -mal augmentiert werden.

10.3 Satz von Philip Hall

Ein Matching M in einem Graphen $G = (V, E)$ heißt **perfekt**, falls M alle Knoten in V überdeckt. Für bipartite Graphen $G = (X \cup Y, E)$ lautet die Hall-Bedingung: $\forall S \subseteq X : |N(S)| \geq |S|$. $N(S)$ ist die Menge aller Nachbarn der Knoten $v \in S$.

Hat ein Graph G ein perfektes Matching, erfüllt er die Hall-Bedingung.

Ein Graph G mit $|X| = |Y|$ hat ein perfektes Matching gdw. die Hall-Bedingung erfüllt ist.

10.4 Satz von Dénes König

Es sei $G = (V, E)$ ein ungerichteter Graph. Eine Knotenmenge $C \subseteq V$ heißt **Vertex Cover**, falls jede Kante $e \in E$ zu mindestens einem Knoten in C inzident ist.

Bezeichne $\nu(G)$ die Kardinalität des größten Matchings in G und $\tau(G)$ die Kardinalität des kleinsten Vertex Covers in G . Jeder Graph erfüllt $\nu(G) \leq \tau(G)$.

Nach König gilt in einem bipartiten Graphen $G = (X \cup Y, E)$ stets $\nu(G) = \tau(G)$.

10.5 Algorithmus für gewichtetes Matching

Es sei $G = (V, E)$ ein bipartiter Graph mit Kantengewichten $w : E \rightarrow \mathbb{R}$. Wir definieren für jedes Matching $M \subseteq E$: $w(M) = \sum_{e \in M} w(e)$. Ein Matching $M \subseteq E$ heißt **k-extrem**, wenn $|M| = k$ und $w(M') \leq w(M)$ für alle Matchings $M' \subseteq E$ mit $|M'| = k$.

Es ist $M_0 = \emptyset$ stets ein 0-extremes Matching. Wir beginnen mit M_0 und konstruieren für $k \geq 1$ ein k -extremes Matching M_k aus M_{k-1} . Dazu definieren wir für jede Kante $e \in E$ eine reelle Länge mit $l(e) = w(e)$ wenn $e \in M_{k-1}$ und $l(e) = -w(e)$ wenn $e \notin M_{k-1}$.

In einem kantengewichteten bipartiten Graphen $G = (V, E)$ kann ein Matching mit maximalem Gewicht in $\mathcal{O}(|V|^2 \cdot |E|)$ Zeit berechnet werden, da maximal $|V|$ -mal augmentiert werden muss und eine Augmentierung $\mathcal{O}(|V| \cdot |E|)$ Zeit kostet (mit Bellman-Ford).

11 Maximaler Fluss

Wir haben ein Netzwerk mit einer Quelle und einer Senke, die über mehrere Wege miteinander verbunden sind. Diese Wege haben unterschiedliche Kapazitäten (wie zum Beispiel Straßen). Uns interessiert der maximale Fluss, also die maximale Rate, mit der Material von der Quelle bis zur Senke transportiert werden kann, ohne die Kapazitätsbeschränkungen der Straßen zu verletzen.

11.1 Flussnetzwerke

Zur Formalisierung des Problems verwenden wir ein Flussnetzwerk. Das ist ein gerichteter Graph $G = (V, E, c)$ mit der Kapazitätsfunktion $c : V \times V \rightarrow \mathbb{R}_{\geq 0}$, einer Quelle s und einer Senke t . Dabei liegt jeder Knoten $v \in V$ auf einem Pfad von der Quelle s zur Senke t .

Ein Fluss ist eine Funktion $f : V \times V \rightarrow \mathbb{R}$, dabei ist $f(u, v)$ der Fluss vom Knoten u zum Knoten v . Es ist für $u, v \in V$ immer $f(u, v) \leq c(u, v)$ und $f(u, v) = -f(v, u)$. Ferner gilt für $u \in V \setminus \{s, t\}$ immer $\sum_{v \in V} f(u, v) = 0$.

Der Wert $|f|$ eines Flusses f ist der Gesamtfluss aus der Quelle s .

Ein Flussnetzwerk mit mehreren Quellen und Senken lässt sich auf ein Flussnetzwerk mit einer Quelle und einer Senke zurückführen, indem wir eine „Superquelle“ und eine „Supersenke“ einführen, die wir mit den einzelnen Quellen bzw. Senken verbinden.

Der Fluss zwischen zwei Knoten lässt sich leicht auf den Fluss zwischen einem Knoten und einer Knotenmenge oder sogar zwischen zwei Knotenmengen erweitern. Es ist für $x \in V, Y \subseteq V$ stets $f(x, Y) = \sum_{y \in Y} f(x, y)$ und für $X, Y \subseteq V$ stets $f(X, Y) = \sum_{x \in X} \sum_{y \in Y} f(x, y)$. Für $X \subseteq V$ gilt übrigens $f(X, X) = 0$.

Wir nennen ein Flussproblem **integral**, wenn alle Kapazitäten ganzzahlig sind.

11.2 Restnetzwerke

Ein Restnetzwerk beschreibt, welche Kapazitäten in einem gegebenen Flussnetzwerk mit einem bestimmten Fluss noch vorhanden sind. Sei das Flussnetzwerk $G = (V, E, c)$ und der Fluss f gegeben. Dann ist $G_f = (V, E_f, c_f)$ das Restnetzwerk (auch Residualnetzwerk) zu G und f mit $c_f(u, v) = c(u, v) - f(u, v)$ und $E_f = \{(u, v) \in V \times V \mid c_f(u, v) > 0\}$.

Damit ist $c_f(u, v)$ die Restkapazität von (u, v) in G zum Fluss f und E_f sind alle Kanten, die noch mehr Fluss aufnehmen können.

Es muss unbedingt beachtet werden, dass die Restkapazität in beide Richtungen zwischen den Knoten beachtet wird (negativer Fluss!). Daher gilt stets $|E_f| \leq 2 \cdot |E|$.

Ein s - t -Pfad p im Restnetzwerk G_f heißt augmentierender Pfad (vergrößernder Pfad), da über diesen der Fluss im Flussnetzwerk noch erhöht werden kann. Die Restkapazität des Pfades p ist die kleinste Kapazität der Kanten in diesem Pfad.

11.3 Ford-Fulkerson-Theorem

Wir haben einen Fluss f in G und einen Fluss f' im Restnetzwerk G_f . Die Hauptidee der Ford-Fulkerson-Theorem ist es, zu versuchen, den Fluss f um den Fluss f' zu ergänzen.

Dazu definieren wir zunächst die **Flusssumme**. Es ist $f_1 + f_2$ trivial definiert durch $(f_1 + f_2)(u, v) = f_1(u, v) + f_2(u, v)$.

Nach dem Ford-Fulkerson-Theorem ist $f + f'$ ein Fluss in G mit dem Wert $|f + f'|$.

11.4 Ford-Fulkerson-Methode

Bei der Ford-Fulkerson-Methode suchen wir zunächst einen beliebigen Pfad p von s nach t . Nun setzen wir den Fluss der Kanten in p auf die minimale Kapazität in p (wenn eine Kante eine Kapazität von 5 setzen wir den Fluss aller Kanten in p auf 5). Sodann suchen wir uns einen neuen (wieder beliebigen) Pfad p' von s nach t aus Kanten mit **freier** Kapazität (die Kante, die soeben das Minimum gestellt hat, wird also nicht mehr verwendet). Es wird dann der Fluss entsprechend der minimalen Restkapazität auf diesem Pfad angepasst. Das wiederholen wir, bis es keinen Pfad mit freien Kapazitäten mehr gibt.

Das entspricht einer Anwendung des Ford-Fulkerson-Theorems: Solange es einen augmentierenden Pfad p gibt, augmentieren wir f entlang p .

Sei f^* der durch die Ford-Fulkerson-Methode bestimmte Fluss zu einem integralen Flussproblem, so benötigt die Methode $|f^*|$ Iterationen und es ergibt sich eine Laufzeit von $\mathcal{O}(|E| \cdot |f^*|)$.

11.5 Schnitte in Flussnetzwerken

Ein Schnitt (S, T) in einem Flussnetzwerk $G = (V, E, c)$ ist eine Partition $S \cup T = V$, $S \cap T = \emptyset$ mit $s \in S$ und $t \in T$. Wenn f ein Fluss in G ist, dann ist $f(S, T)$ der Fluss über (S, T) . Die Kapazität von (S, T) ist $c(S, T)$. Ein minimaler Schnitt ist ein Schnitt mit minimaler Kapazität. Für den Fluss über einen Schnitt gilt stets $f(S, T) = |f|$.

11.6 Max-flow Min-cut Theorem

Das Max-flow min-cut Theorem besagt für einen Fluss f im Flussnetzwerk $G = (V, E, c)$, dass folgende Sätze äquivalent sind:

1. f ist ein maximaler Fluss.
2. Im Restnetzwerk G_f gibt es keinen augmentierenden Pfad.

3. $|f| = c(S, T)$ für einen Schnitt (S, T) , d.h. (S, T) ist minimal.

Damit ist die Kapazität eines minimalen Schnittes (min cut) gleich dem Wert eines maximalen Flusses (max flow). Mit diesem Theorem lässt sich auch die Korrektheit der Ford-Fulkerson-Methode zeigen.

11.7 Edmonds-Karp-Algorithmus

Der Edmonds-Karp-Algorithmus übernimmt die Ford-Fulkerson-Methode, wählt den augmentierenden Pfad aber nicht beliebig, sondern nutzt zur Bestimmung der Pfade eine Breitensuche. Damit präferiert der Algorithmus kürzere Pfade gegenüber längeren - er erweitert den Fluss stets entlang kürzester Pfade. Die Laufzeit liegt in $\mathcal{O}(|V| \cdot |E|^2)$.

12 Dynamische Programmierung

12.1 Memoization

Der Hauptansatz von dynamischer Programmierung ist, bereits berechnete Werte nicht noch einmal zu berechnen. Besonders wichtig ist das bei Rekursionsgleichungen (wie etwa bei Fibonacci), weil man hier oft den gleichen Wert sehr oft verwenden muss.

Deshalb speichern wir einmal berechneten Werte. Das kostet Speicher, kann aber zu einer enormen Reduzierung der Zeitkomplexität führen. Bei jedem Funktionsaufruf prüfen wir nun, ob das Ergebnis bereits berechnet wurde (im Cache ist) - ist dies der Fall, holen wir den Wert aus dem Speicher. Ansonsten müssen wir den Wert berechnen, zusätzlich speichern wir das Ergebnis im Cache.

Memoization hilft, wenn die Teilprobleme überlappen. Bei den Fibonacci-Zahlen sinkt die Zeitkomplexität von $\Theta(2^n)$ auf $\Theta(n)$, was natürlich enorm ist. Dafür wächst der Platzbedarf aber auf $\Theta(n)$. Eine Verbesserung ergibt sich bei den Fibonacci-Zahlen, indem man erkennt, dass nur jeweils die zwei letzten Werte benötigt werden (in-place).

12.2 The principle of optimality

Es sei L die optimale Lösung für ein Problem P . Es sei Q ein Teilproblem von P . Dann ist die Restriktion von L auf Q eine optimale Lösung für das Teilproblem Q .

12.3 Ketten von Matrixmultiplikation

Bei einer einfachen Matrixmultiplikation benötigen wir $i \cdot j \cdot k$ Fließkomma-Multiplikationen (Zeilen erste Matrix * Spalten erste Matrix * Spalten zweite Matrix). Wir betrachten nun die Multiplikation mehrerer Matrizen $M = A_1 \cdot A_2 \cdot \dots \cdot A_n$ (wir gehen davon aus, dass die Matrizen jeweils miteinander kompatibel bzgl. der Matrixmultiplikation sind). Solche Ketten lassen sich wegen der Assoziativität der Matrixmultiplikation in beliebiger Reihenfolge berechnen bzw. klammern - darin liegt der Trick. Wenn wir beispielsweise drei Matrizen multiplizieren sollen, können wir $(A_1 \cdot A_2) \cdot A_3$ oder $A_1 \cdot (A_2 \cdot A_3)$ berechnen. Das kann ein gravierender Unterschied sein: Sei $A_1 \in \mathbb{R}^{10 \times 100}$, $A_2 \in \mathbb{R}^{100 \times 5}$ und $A_3 \in \mathbb{R}^{5 \times 50}$. Dann müssen wir bei der ersten Klammerung 7.500 Multiplikationen durchführen - bei der zweiten Klammerung benötigen wir allerdings enorme 75.000 Multiplikationen.

Die Problemstellung ist nun also, für eine Kette von Matrizen A_1, A_2, \dots, A_n mit den Dimensionen $d_0 \times d_1, d_1 \times d_2, \dots, d_{n-1} \times d_n$ eine Klammerung so zu finden, dass die Anzahl der Fließkomma-Multiplikationen minimal ist - wir möchten die Berechnung ja möglichst effizient durchführen.

Es sei nun $P(n)$ die Anzahl der möglichen Klammerungen für n Matrizen. Wir können nun bestimmen, dass $P(n) = \sum_{k=1}^{n-1} P(k) \cdot P(n-k) \in \Omega(2^n)$ und $P(1) = 1$ gilt. Allerdings benötigen wir eine Rekursionsgleichung für die minimale Anzahl an Multiplikationen.

Sei nun $m[i, j]$ die minimale Anzahl Multiplikationen für die Teilkette $A_i \dots A_j$. Offenbar ist $m[i, i] = 0$ für alle $0 < i \leq n$. Die Dimension einer Teilkette ist $d_{i-1} \times d_j$. Durch das Teilen der Kette an der Position k ergibt sich: $m[i, j] = m[i, k] + d_{i-1} \cdot d_k \cdot d_j + m[k+1, j]$. Wir suchen dabei das optimale k , also ist $m[i, j] = 0$ für $i = j$ und für $i < j$ ist $\min_{i \leq k < j} (m[i, k] + m[k+1, j] + d_{i-1} \cdot d_k \cdot d_j)$. Wie bei Fibonacci werden Teilprobleme wieder mehrfach verwendet - z.B. benötigen wir $m[0, 1]$ für $m[0, 2], m[0, 3], \dots, m[0, n]$. Deswegen kann DP zum Einsatz kommen. Wir berechnen die Matrix von unten nach oben und von links nach rechts.

Die Zeitkomplexität war exponentiell, nun ist sie $\Theta(n^3)$. Für die naive Variante wurde kein Platz benötigt, nun haben wir einen Platzbedarf von $\Theta(n^2)$.

12.4 Das Rucksackproblem

Gegeben sei ein Rucksack mit maximaler Tragkraft M sowie n Gegenstände, die sowohl ein Gewicht als auch einen Wert haben. Wir wollen nun möglichst viel Wert mitnehmen, ohne den Rucksack zu überladen. Ein erster Ansatz ist, die Gegenstände nach dem Gewicht-Wert-Verhältnis absteigend zu sortieren und dann solange Gegenstände einzupacken, bis die maximale Tragkraft überschritten werden würde. Allerdings ist dieser Algorithmus nicht optimal, weil eventuell noch Gegenstände reinpassen würden.

Wir überlegen uns nun eine Rekursionsgleichung. Diese basiert auf der Überlegung, ob wir den n -ten Gegenstand mitnehmen (können) oder nicht. Es sei nun $C[i, j]$ der maximale Wert des Rucksacks mit Tragkraft j , wenn man nur die Gegenstände $\{0, \dots, i-1\}$ berücksichtigt. Daraus ergibt sich $C[i, j] = 0$ für $i = 0$, $C[i, j] = -\infty$ für $j < 0$ und ansonsten $C[i, j] = \max(C[i-1, j], c_{i-1} + C[i-1, j-w_{i-1}])$. Damit ist $c_{\max} = C[n, M]$.

Wir lösen die Rekursionsgleichung nun bottom-up und kommen auf die gleiche Zeit- und Platzkomplexität $\Theta(n \cdot M)$.

13 Greedy-Algorithmen

13.1 Grundlegendes

„Greedy“ lässt sich zu „gierig“ übersetzen. Ein Greedy-Algorithmus trifft in jedem Schritt eine Entscheidung, die bezüglich eines „kurzsichtigen“ Kriteriums optimal ist. Dieses Kriterium ist einfach und schnell auszuwerten. Wurde eine Wahl gemacht, kann sie nicht mehr rückgängig gemacht werden.

Greedy-Algorithmen finden nicht immer die beste Lösung, da die wiederholte Wahl eines lokalen Optimums nicht automatisch zum globalen Optimum führt. Greedy kann optimal, gut und beliebig schlecht sein.

So ist Greedy zum Beispiel optimal für Kruskal's Algorithmus, gut für die Bildung von Matchings mit maximaler Kardinalität (markiere neue Kante, die mit keiner markierten Kante kollidiert) und mit der in der Vorlesung präsentierten Implementierung schlecht für das Knotenfärbungsproblem.

13.2 Sequentielle Speicherung

In diesem Beispiel stellt sich das Problem, dass wir Dateien D_1, \dots, D_n mit den Längen L_1, \dots, L_n haben, die wir auf einem Magnetband speichern sollen. Dabei suchen wir eine Anordnung der Dateien auf dem Magnetband, die die durchschnittliche Lesezeit minimiert.

Wenn wir die Dateien in der Reihenfolge D_1, \dots, D_n abspeichern, ist die Lesezeit für D_k dann $\sum_{i=1}^k L[i]$. Die durchschnittliche Lesezeit ist dann $\frac{1}{n} \sum_{k=1}^n (n - k + 1) \cdot L[k]$. Diese durchschnittliche Lesezeit müssen wir mit einer Permutation π auf $L[1], \dots, L[n]$ minimieren.

Wir kommen darauf, dass wir die durchschnittliche Lesezeit minimieren können, indem wir Dateien nach monoton steigender Länge abspeichern. Ein Greedy-Algorithmus wählt also immer die kürzeste Datei und speichert sie an der nächsten freien Stelle des Bandes. Die resultierende Laufzeit ist $\mathcal{O}(n \cdot \log(n))$.

13.3 Stundenplanung

Wir haben Vorträge V_1, \dots, V_n mit Startzeiten $L[1], \dots, L[n]$ und Endzeiten $R[1], \dots, R[n]$. Dazu suchen wir eine möglichst große Teilmenge von Vorträgen, die zeitlich nicht überlappen.

Der Greedy-Algorithmus nimmt $R[1] \leq R[2] \leq \dots \leq R[n]$ an. Es gibt nun logischerweise eine optimale Lösung des Stundenplanungs-Problems, die Vortrag V_1 (mit kleinstem rechten Endpunkt $R[1]$) verwendet. Das löst der Greedy-Algorithmus iterativ.

13.4 Korrektheitsbeweis

Bei beiden vorgestellten Algorithmen kann man die Korrektheit beweisen, indem man annimmt, dass es eine optimale Lösung gibt, die sich von Greedy unterscheidet. Nun findet man den „ersten“ Unterschied zwischen den beiden Lösungen. Dann kann man argumentieren, dass man die optimale Entscheidung durch die Greedy-Entscheidung ersetzen kann, ohne die Qualität zu verschlechtern.

13.5 Huffman-Codes

Ein Binär-Code kodiert jeden Buchstaben eines Alphabets Σ durch einen String aus $\{0, 1\}$. Wir nennen einen Binär-Code präfixfrei, wenn kein Codewort der Präfix eines anderen Codeworts ist.

Einen präfix-freien Binär-Code können wir durch einen Binär-Baum darstellen: Die Buchstaben sind in den Blättern, das Codewort für Buchstaben entspricht dem Pfad von der Wurzel zum Blatt (links = 0, rechts = 1). Die Länge des Codeworts entspricht der Tiefe des Blattes im Baum.

Wir möchten nun unseren präfix-freien Binär-Code so optimieren, dass kodierte Botschaften mit der Texthäufigkeit h_1, \dots, h_n der Buchstaben so kurz wie möglich sind. Die Gesamtlänge unserer Botschaft ist $\sum_{i=1}^n h_i \cdot \text{Tiefe}[i]$.

Der Greedy-Algorithmus verschmilzt nun die beiden Buchstaben mit kleinster Häufigkeit h und h' , das Resultat hat die Häufigkeit $h + h'$. Das wird solange wiederholt, wie zwei oder mehr Buchstaben vorhanden sind. Wir implementieren das mit MinHeap (ExtractMin; Insert).

14 Algorithmische Geometrie

In der algorithmischen Geometrie geht es um geometrische Probleme im n -dimensionalen Raum \mathbb{R}^n , zum Beispiel die Frage, ob sich zwei Geraden schneiden. Wir beschränken uns auf Probleme im zweidimensionalen Raum ($n = 2$), dazu nutzen wir Konzepte der linearen Algebra.

Wir entwerfen Algorithmen, die trigonometrische Funktionen (Sinus, ...) und Division vermeiden, da diese zu Rundungsfehlern neigen. Wir erhoffen uns davon, dass unsere Algorithmen genauer sind.

14.1 Geometrische Interpretation

Der Winkel α zwischen zwei Vektoren, die im gleichen Punkt starten, ergibt sich aus $\vec{x} \cdot \vec{y} = |\vec{x}| \cdot |\vec{y}| \cdot \cos(\alpha)$. Die Fläche (allgemeiner: das Volumen) des durch \vec{x} und \vec{y} aufgespannten Parallelogramms ist der Absolutwert der Determinanten, $|\det(\vec{x}, \vec{y})|$.

14.2 Links oder rechts?

Das erste Problem, das wir betrachten, stellt die Frage, ob ein Vektor \vec{y} links oder rechts von einem gegebenen Vektor \vec{x} liegt. Dazu konstruieren wir den Vektor \vec{z} , der zu \vec{x} orthogonal (90 Grad) im mathematisch positiven Sinn (Gegenuhrzeigersinn) ist.

Dann projizieren wir \vec{y} auf \vec{z} . Ist $\vec{z} \cdot \vec{y}$ negativ, zeigt \vec{y} von \vec{z} weg und liegt damit rechts von \vec{x} .

Wie berechnen wir aber \vec{z} aus \vec{x} ? Nun, es ist für $\vec{x} = (x_1, x_2)$ gerade $\vec{z} = (-x_2, x_1)$. Damit ist insbesondere $\vec{z} \cdot \vec{y} = -x_2 y_1 + x_1 y_2 = \det(\vec{x}, \vec{y})$.

Also: Ist $\det(\vec{x}, \vec{y}) = 0$, dann sind \vec{x} und \vec{y} parallel bzw. antiparallel. Wenn $\det(\vec{x}, \vec{y}) > 0$ ist, dann liegt \vec{y} links von \vec{x} - ist hingegen $\det(\vec{x}, \vec{y}) < 0$, dann liegt \vec{y} rechts von \vec{x} .

14.3 Punkt, Strecke, Polygon

Es ist $p = (p_1, p_2)$ ein Punkt aus dem \mathbb{R}^2 . Der Punkt $(0, 0)$ heißt Ursprung. Mit dem Vektor $\vec{d}_{pq} = q - p$ kommen wir vom Punkt p nach q . Die (ungerichtete) Strecke \overline{pq} ist die Menge aller Punkte zwischen den beiden Endpunkten p und q gegeben durch: $\overline{pq} = \{(1 - \alpha) \cdot p + \alpha \cdot q \mid 0 \leq \alpha \leq 1\} = \{p + \alpha \cdot \vec{d}_{pq} \mid 0 \leq \alpha \leq 1\}$. Fassen wir \overline{pq} als gerichtete Strecke auf, so ist \vec{d}_{pq} die Richtung.

Ein Streckenzug ist eine Folge von Punkten (p_1, \dots, p_n) , die durch Strecken miteinander verbunden sind: $\overline{p_1 p_2}, \overline{p_2 p_3}, \dots, \overline{p_{n-1} p_n}$. Ein Polygon mit den Ecken p_1, \dots, p_n hat als Rand den geschlossenen Streckenzug (p_1, \dots, p_n, p_1) .

VL22 (12.08.2017)

15 Abkürzungsverzeichnis

12.08.2017