DB2 – WS 2017/2018

B*-Baum – 29.11.2017 und 06.12.2017

Dr. Karsten Tolle

Graphtheorie …

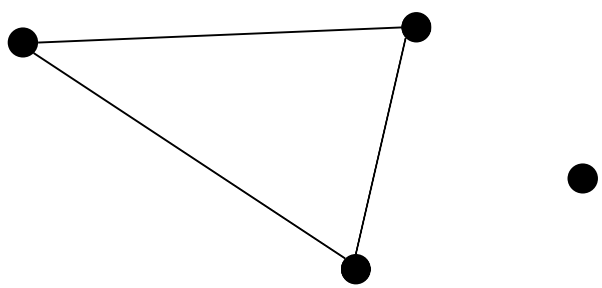
Bäume

Definitionen manchmal merkwürdig – man findet den Zusatz „Teil der Graphtheorie“.
… oder es wird nicht Baum, sondern eine spezielle Baumart definiert, z.B.:

- Binärbaum
- gerichteter Baum
- gewurzelter Baum
- geordneter Baum
- …

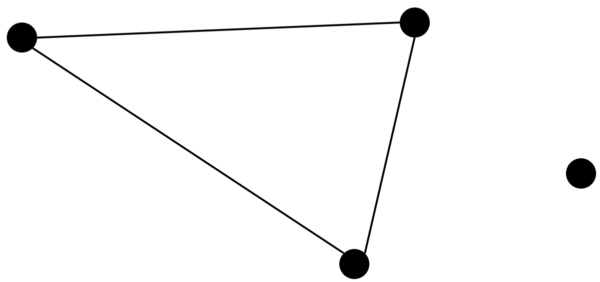
Graphtheorie …

Ein **Graph G** ist ein Paar G=(V,E) aus einer (endlichen) Menge von **Knoten** V und einer Menge von **Kanten** E ($E \subseteq V \times V$).

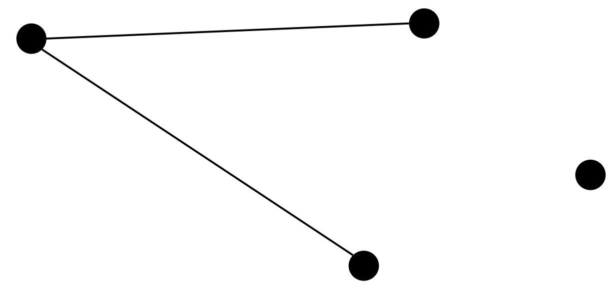


Basiert auf: Dieter Jungnickel „Graphen, Netzwerke und Algorithmen“
3. Auflage von 1994

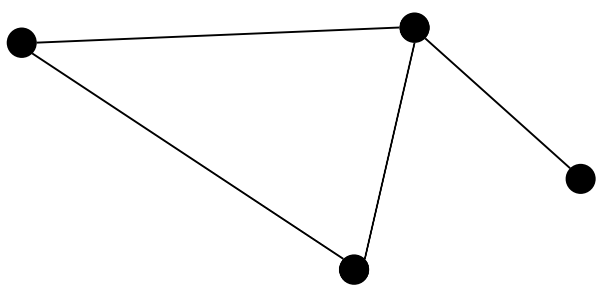
kreisfrei / zusammenhängend



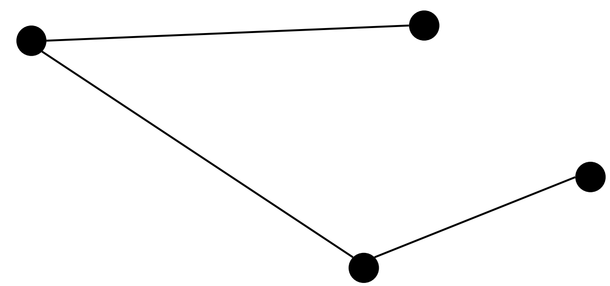
nicht zusammenhängend
nicht kreisfrei



kreisfrei (zyklenfrei)
nicht zusammenhängend



zusammenhängend gdw. alle Knoten verbunden
nicht kreisfrei



zusammenhängend und kreisfrei

Baum und Blatt

G sei ein Graph auf n Knoten. Dann implizieren je zwei der folgenden drei Bedingungen die dritte:

a) G ist zusammenhängend.

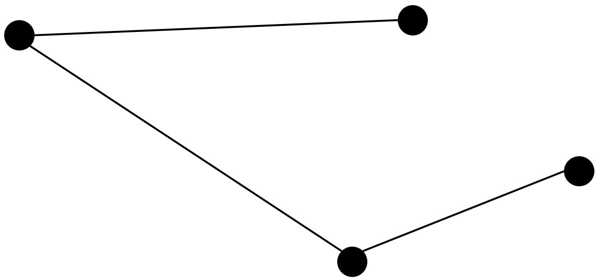
b) G ist kreisfrei.

c) G hat n-1 Kanten.

Ein Graph T, der diesen Bedingungen genügt, heißt ein **Baum** („tree“).

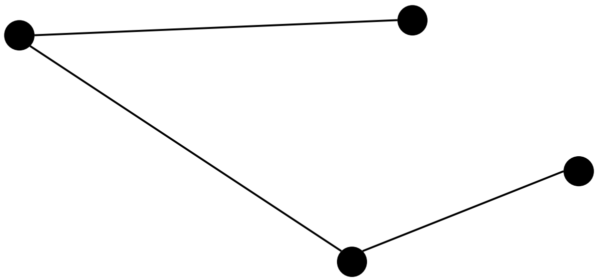
Jeder Knoten vom Grad 1 (Anzahl der Kanten, die den Knoten enthalten) heißt **Blatt** („leaf“).

Baum → Wurzel?



zusammenhängend und kreisfrei
→ Baum

Wurzel?



zusammenhängend und kreisfrei
→ Baum ohne Wurzel!

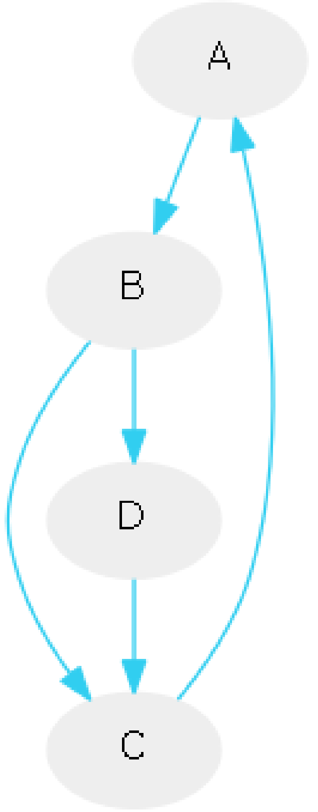
<u>Graphentheorie:</u>

In einem **gerichteten Graphen** heißt ein Knoten, von dem aus alle Knoten erreichbar sind, eine **Wurzel** („root“).

Basiert auf: Dieter Jungnickel „Graphen, Netzwerke und Algorithmen“
3. Auflage von 1994

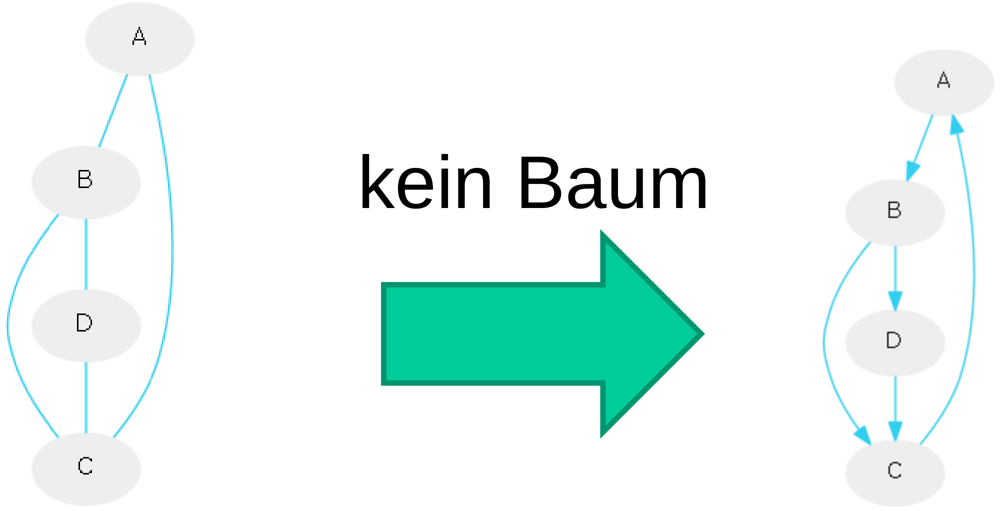
gerichteter Graph (digraph)

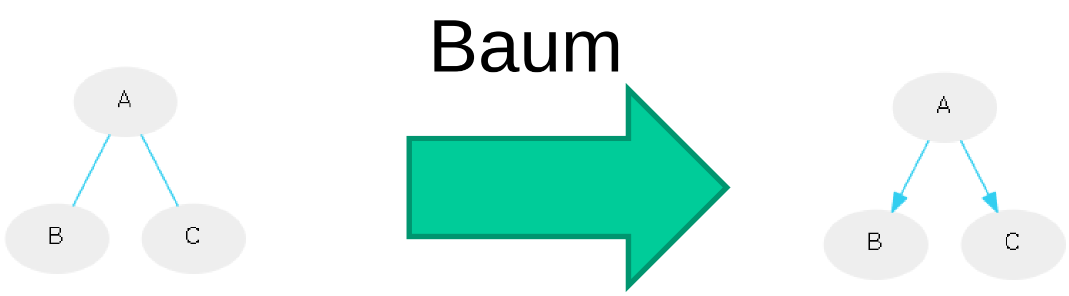
„In *gerichteten* oder auch *orientierten Graphen* werden Kanten statt durch Linien durch Pfeile gekennzeichnet, wobei der Pfeil vom ersten zum zweiten Knoten zeigt. Dies verdeutlicht, dass jede Kante des Graphen nur in eine Richtung durchlaufen werden kann.“ (Wikipedia)



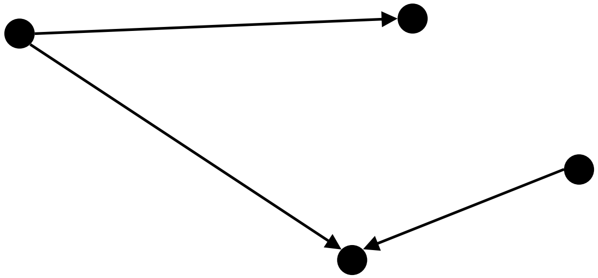
gerichteter Graph (digraph) -> Baum

Ein gerichteter Graph ist ein Baum gdw. der entsprechende ungerichtete Graph ein Baum ist.

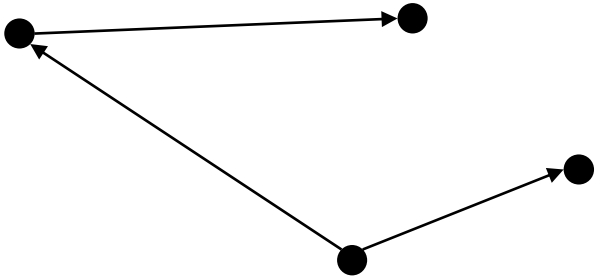




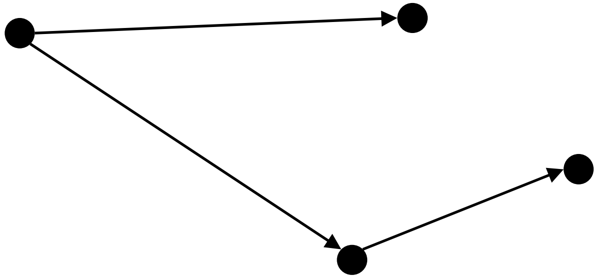
gerichtete Graphen



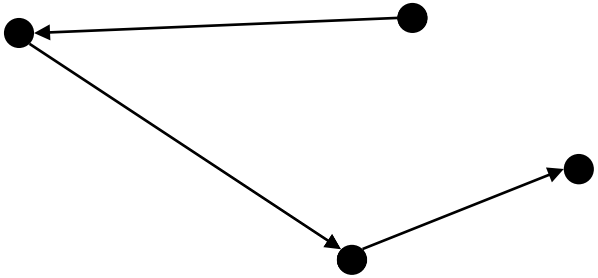
zusammenhängend und kreisfrei
→ Baum



zusammenhängend und kreisfrei
→ Baum

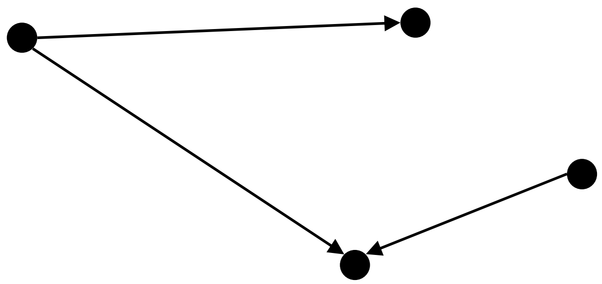


zusammenhängend und kreisfrei
→ Baum

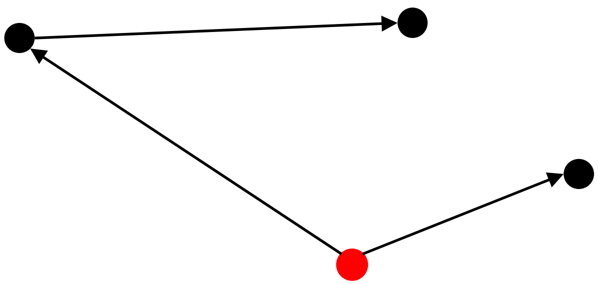


zusammenhängend und kreisfrei
→ Baum

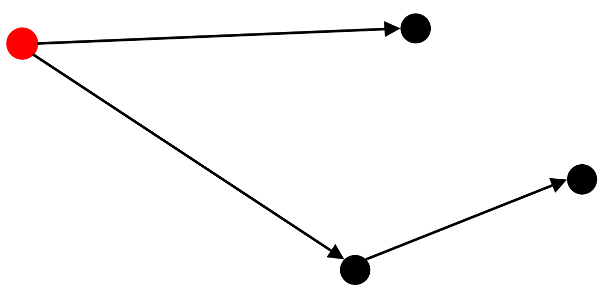
gerichtete Graphen (Wurzel rot)



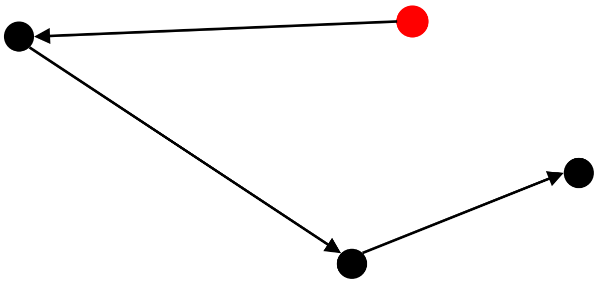
zusammenhängend und kreisfrei
→ Baum ohne Wurzel



zusammenhängend und kreisfrei
→ Baum

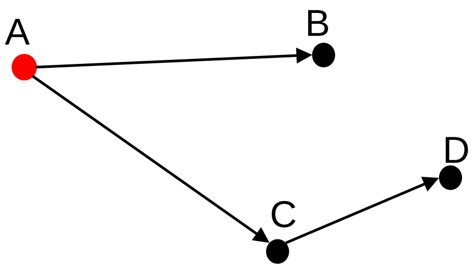


zusammenhängend und kreisfrei
→ Baum

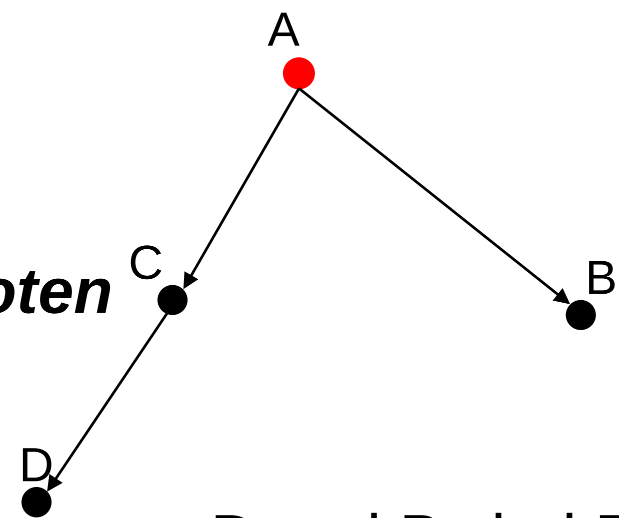


zusammenhängend und kreisfrei
→ Baum

gewurzelter Baum (Baum mit Wurzel)



Darstellung: Wurzel oben



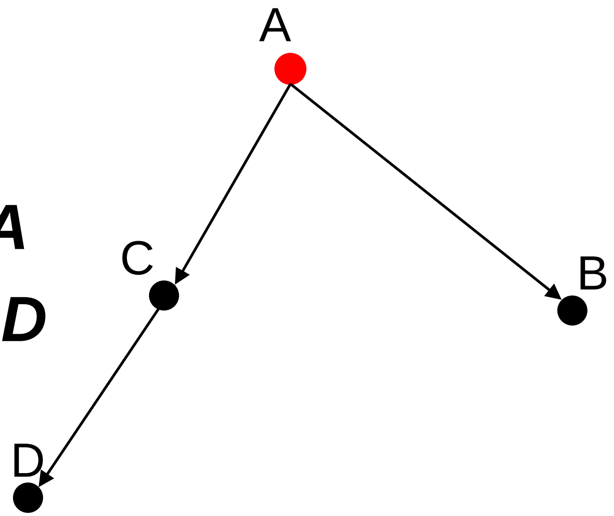
C ist ***innerer Knoten***

D und B sind Blätter

Vater, Mutter, Kind

C ist ***Kind von A***

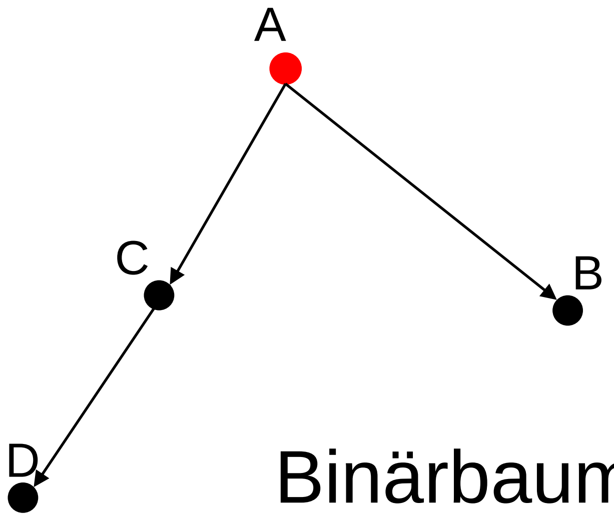
C ist ***Vater von D***



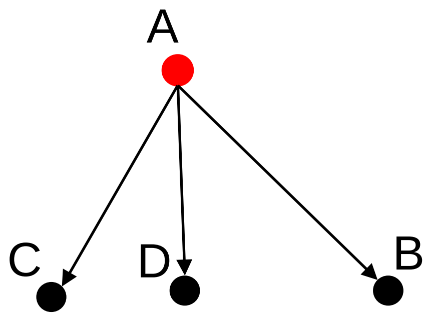
Jeder Knoten hat genau
einen Vater (außer die Wurzel)!
(in einem gewurzeltem Baum)

Binärbaum

Definition: Ein *Binärbaum* ist ein gewurzelter Baum, bei dem jeder Knoten höchstens zwei Kindknoten besitzt.



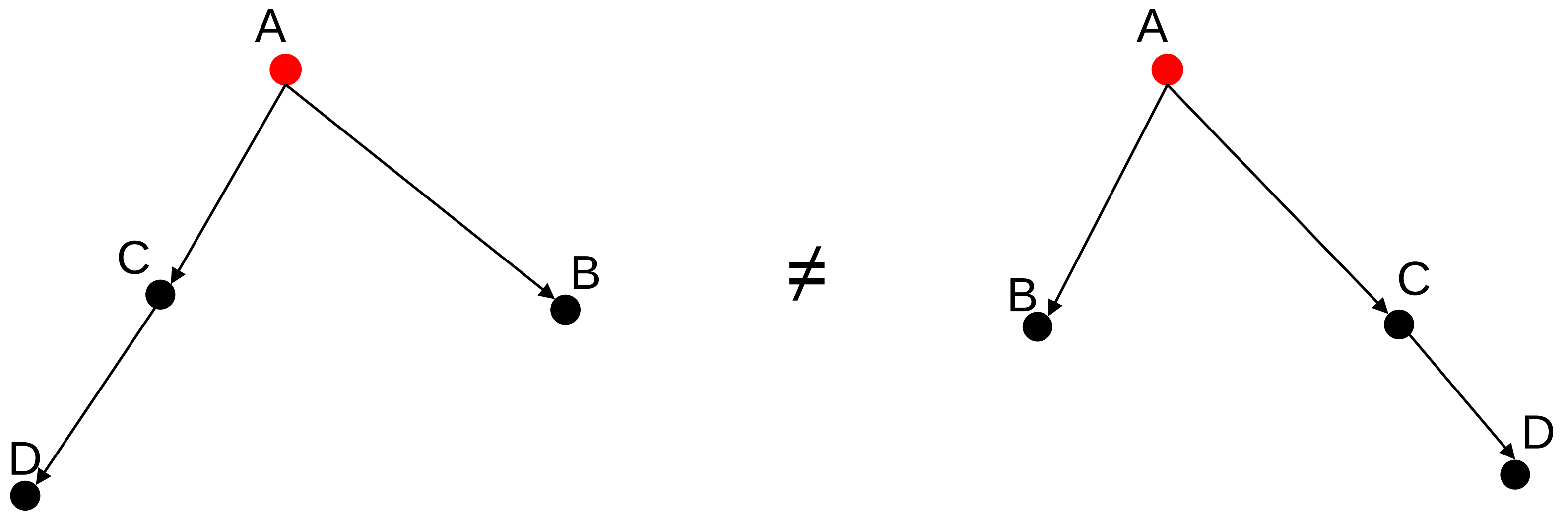
Binärbaum



kein Binärbaum

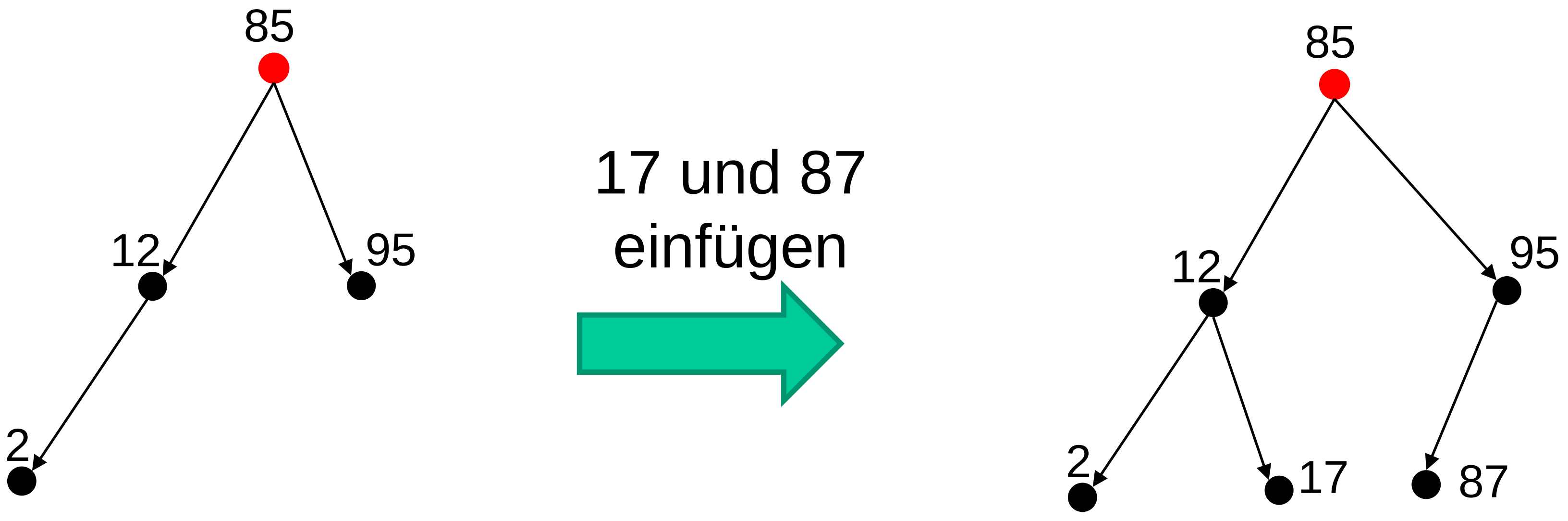
geordneter Baum

Bei einem geordneten Baum ist die Reihenfolge der Teilbäume eines Knotens relevant.



Binärer Suchbaum

Ist ein geordneter Binärbaum, bei dem durch die links-rechts Aufteilung eine Ordnung erstellt wird. … z.B. alle kleiner nach links und alle größer oder gleich nach rechts.



B*-Baum (bei uns) mit k und k*

Balancierter Baum (alle Blätter auf einer Ebene)

k gibt die minimal erlaubte Anzahl von Kindknoten an (innerer Knoten) → maximal 2k-1

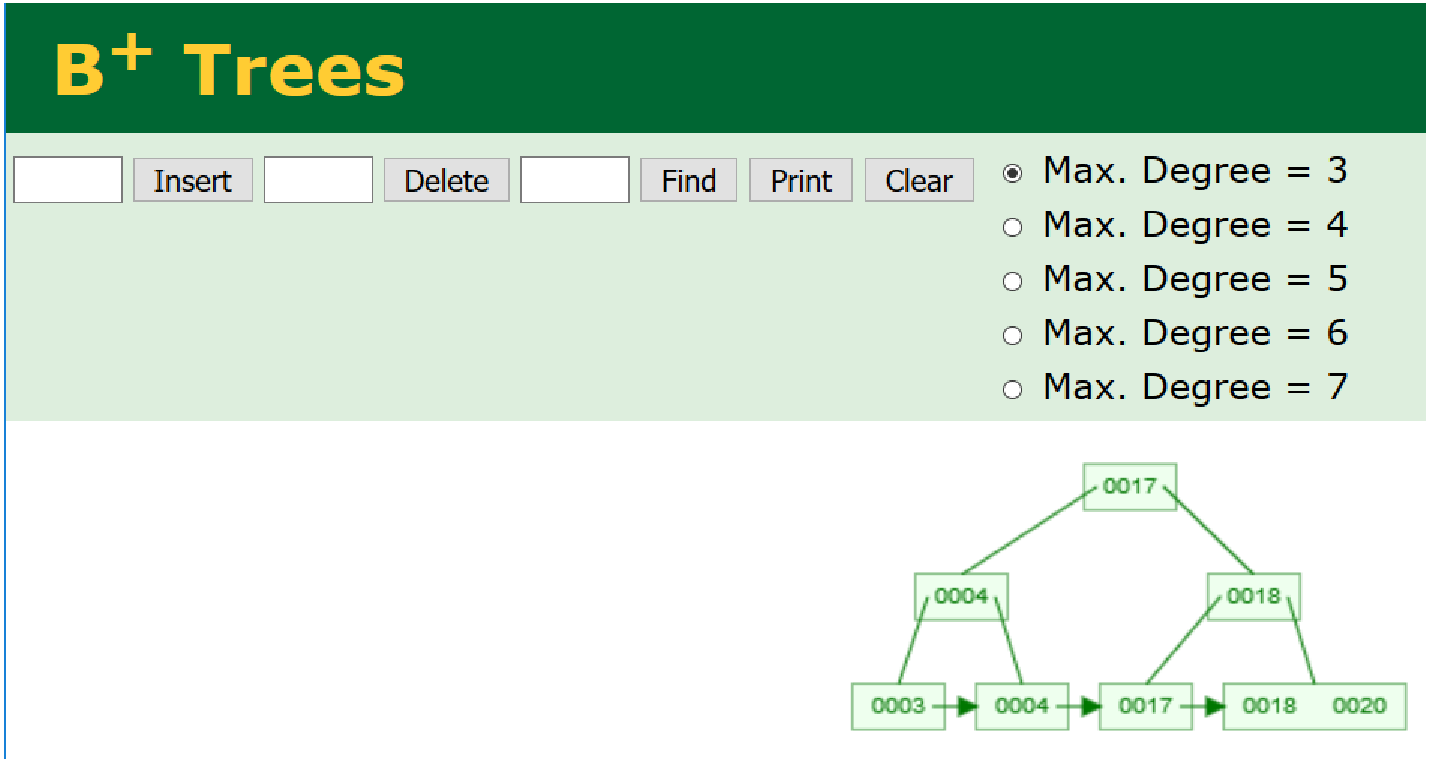
k* gibt die minimal erlaubte Anzahl von Einträgen in Blättern an → maximal 2k*-1

… die Wurzel ist nur nach oben begrenzt!

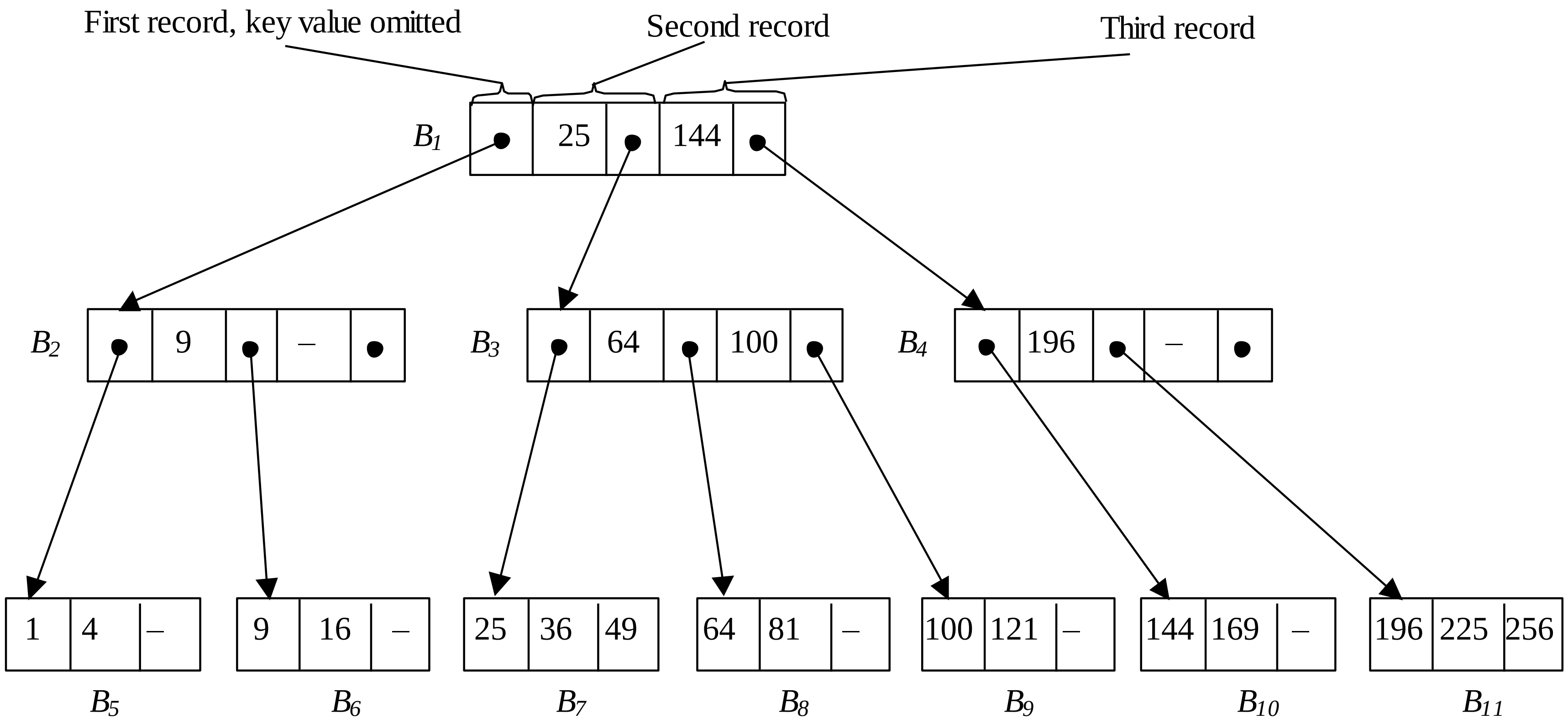
Visualisierungen

http://www.cs.usfca.edu/~galles/visualization/BPlusTree.html

… man achte auf die feinen Unterschiede zur Vorlesung (z.B. nur ein Wert!)



Aus den Folien: Wie erstellen?



Reihenfolge der Eingabe relevant?

Typische Aufgabentypen zu B*-Baum

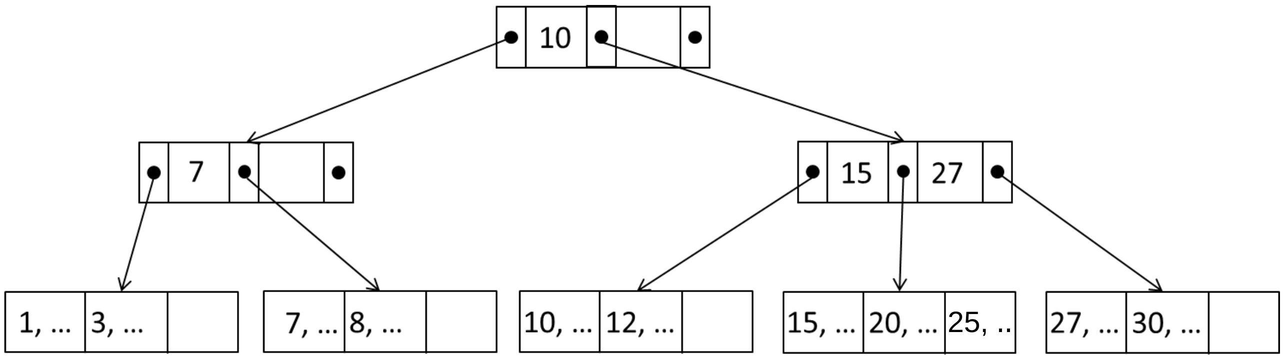
- Fragen zu Grundlagen (siehe oben)
- Erstellung/Zeichnen eines Baumes
- Beschreibung/Zeichnen der Operationen (Lookup, Modify, Insert, Delete)
- Berechnung von k und k* entsprechend der Vorgaben

Gegeben sei ein leerer B*-Baum mit $k = 3$, $k^* = 2$ und den folgenden Daten:

Daten	
ID	Wert
1	A
3	B
6	C
4	D
7	C
5	F
2	A

Fügen Sie die Daten aus der Tabelle in den leeren B*-Baum (über die ID) entsprechend der Reihenfolge der Tabelle (von oben nach unten) ein. Zeichnen Sie mindestens jeden Zwischenschritt, bei dem ein neuer Block für den Baum benötigt wird und den fertigen Baum.

Gegeben sei folgende Ausgangsbaum B*-Baum mit k=k*=2.
Ausgangsbaum:



Führen Sie folgende Operationen jeweils auf dem Ausgangsbaum aus und zeichnen Sie resultierenden Baum erneut auf.

Löschen der 10

Löschen der 7

Einfügen von 18

Berechnung von k und k* entsprechend der Vorgaben

a) Bestimmen Sie die Werte k und k* für einen B*-Baum bei gegebener Seitengröße p=4096 Bytes, Schlüsselgröße s=8 Bytes, Satzgröße r=500 Bytes und Verweisgröße v=8 Bytes. Sonstiger Verwaltungsoverhead inkl. Header kann vernachlässigt werden.

b) Wie viele Ebenen benötigt man mindestens mit diesen Parametern um 5000, 500000 und 50000000 Datensätze zu speichern?