Isomorphie von Bäumen

Alexandra Weinberger

23. Dezember 2011

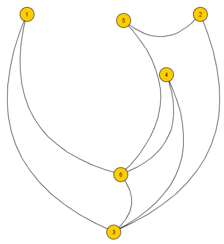
Inhaltsverzeichnis

1 Einige Grundlagen und Definitionen

Ehe wir uns mit dem Eigentlichen, der Isomorphie von Bäumen, beschäfftigen, sollten wir uns einige, (großteil schon bekannte) Definitionen ansehen.

Definition 1. *Ein geordnetes Paar (V, E) nennt sich* `Graph`*. Dabei ist V eine Menge, deren Elemente Knoten genannt werden und E eine Menge, deren Elemente Kanten genannt werden. Jedes Element aus E ist selbst eine Menge, die aus zwei Knoten aus V besteht. (In einer Zeichnung sind die Knoten Punkte und wenn x und y Knoten sind, wäre $\{x, y\} \in E$ eine Verbindungslinies zwischen den beiden.)*

Beispiel 1. *In diesem Graphen ist $V = \{1, 2, ..., 6\}$ und $E = \{\{1, 3\}, \{1, 6\}, \{3, 4\}, \{6, 4\}, \{3, 2\}, \{2, 5\}, \{5, 6\}\}$*



Definition 2. *Ein* `Kreis` *ist ein Graph*

$$G = (V, E) : V = \{1, 2, ..., n\} \text{ und } E = \{\{i, i + 1\} : i = 1, 2, ..., n - 1 \cup \{1, n\}\}$$

Einen Graphen, der keinen Kreis enthält, nennt man `kreisfrei`*.*

Der Graph aus Bsp.1 enthält mehrere Kreise (z.B. V'=$\{1, 3, 6\}$ und E'=$\{\{1, 3\}, \{3, 6\}, \{6, 1\}\}$) und ist demnach nicht kreisfrei.

Definition 3. *Ein* `Weg` *ist ein Grah*

$$G = (V, E) : V = \{0, 1, ..., n\} \text{ und } E = \{\{i - 1, i\} : i = 1, 2, ..., n\}$$

Definition 4. *Sei $G = (V, E)$ ein Graph. Wenn für je zwei Knoten $x, y \in V$ gilt, dass es einen Weg von x nach y gibt, dann ist der Graph* `zusammenhängend`*. Die zusammenhängenden Teilgraphen, mit maximaler Knotenanzahl (sodass sie noch zusammenhängend bleiben), nennen wir* `Zusammenhangskomponenten`*.*

Der Graph aus Bsp.1 ist zusammenhängend.

Definition 5. *Sei v ein Knoten des Graphen G. Die Anzahl von Kanten, die v enthalten, wird* `Grad` *von v genannt. Wir schreiben $deg_G(v)$ für den Grad des Knotens v im Graphen G.*

Der Grad des Knoten 1 aus dem Graphen in Bsp.1 ist 2. (Die Kanten $\{1, 6\}$ und $\{1, 3\}$ enthalten 1.)

[Matoušek, Nešetřil: Diskrete Mathematik] Seiten 120-131

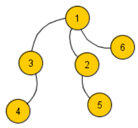
1.1 Bäume

Definition 6. *Ein* **Baum** *ist ein zusammenhängender, kreisfreier Graph.*

Allerdings gibt es mehrere äquivalente Definitionen. Eine davon lautet:

Definition 7. *Ein Baum ist ein minimal zusammenhängender Graph. Das heißt: Der Bau, $T = (V, E)$ ist zusammenhängend, aber wenn man eine beliebe Kante $\in E$ entfernt, ist er das nicht mehr. (Er hat dann zwei Zusammenhangskomponenten.)*

Beispiel 2.

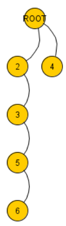


Wurzelbäume:

Definition 8. *Ein Baum T mit einem ausgezeichneten Knoten $r \in V(T)$ ist ein* **Wurzelbaum***. Wir schreiben ihn als Paar (T, r). r wird* **Wurzel** *genannt.*

Definition 9. *Ist $\{x, y\} \in E(T)$ eine Kante, wird der Knoten x* **Vater** *von y genannt, wenn sie auf dem Weg von y zur Wurzel ist. y wird* **Sohn** *von x genannt.*

Beispiel 3.



Dieser Graph ist ein Wurzelbaum, der Knoten 3 ist Vater des Knotens 4 und Sohn des Knotens 2.

Gepflanzte Bäume:

Definition 10. *Ein Wurzelbaum, bei dem für jeden Knoten $v \in V(T)$ die Ordnung $p(v)$ der Söhne vorgegeben ist, wird mit* **gepflanzter Baum** *bezeichnet. Wir schreiben ihn als Tripel (T, r, p).*

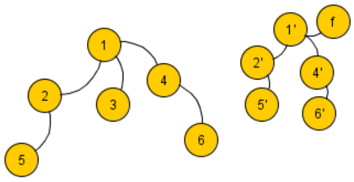
[Matoušek, Nešetřil: Diskrete Mathematik]

1.2 Isomorphie

Definition 11. *Gibt es zwischen zwei Graphen $G = (V, E)$ und $G' = (V', E')$ eine bijektive Abbildung $f : V \to V'$ und ist $\{x, y\} \in E$ genau dann, wenn $\{x', y'\} \in E'$, sind die beiden Graphen* **`isomorph`**. *f wird* **`Isomorphismus`** *zwischen G und G' genannt. Wir schreiben $G \cong G'$, wenn die beiden Graphen isomorph sind.*

[Matoušek, Nešetřil: Diskrete Mathematik] Seiten 123-124
Diese Definition gilt auch für Bäume.

Beispiel 4.



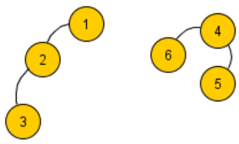
Die beiden Graphen $G = (V, E)$ mit $V = \{1, ..., 6\}$ und $G' = (V', E')$ mit $V' = \{1', ..., 6'\}$ sind isomorphe Bäume.

Definition 12. *Zwei Wurzelbäume (T, r) und (T', r') sind isomorph, wenn die beiden Bäume T und T' isomorph sind und außerdem gilt: $f(r) = r'$.*

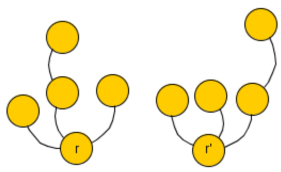
Definition 13. *Zwei gepflanzten Bäumen (T, r, p) und (T', r', p') sind isomorph, wenn die beiden Wurzelbäume(T, r) und (T', r') isomorph sind und außerdem die Ordnung der Söhne erhalten bleibt.*

Der Isomorphiebegriff der geplanzten Bäume ist stärker als der, der Wurzelbäume und beide Begriffe sind stärker als der von Bäumen ohne Wurzel.

Beispiel 5.



Die beiden Bäume $T = (V, E)$ mit $V = \{1, 2, 3\}$ und $T' = (V', E')$ mit $V' = \{4, 5, 6\}$ sind isomorph. Die beiden Wurzelbäume $(T, 1)$ und $(T', 4)$ sind aber nicht isomorph.



Die beiden Graphen sind isomorph, wenn es Wurzelbäume mit den Wurzeln r und r' sind, als geplanzte Bäume sind sie aber nicht isomorph.

1.3 Bestimmung der Isomorphie von Graphen

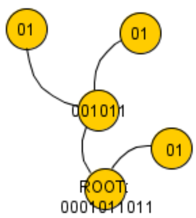
Es existiert kein Algorithmus, der in endlicher Zeit überprüfen kann, ob zwei Graphen isomorph sind. Isomorphie von Graphen ist sogar ein Prblem in NP. Allerdings gibt es für manche Klassen von Graphen effiziente Algorithmen. Zu diesen Klassen gehören auch die Bäume.

2 Isomorphie von gepflanzten Bäumen - Algorithmus

2.1 Code-Darstellung

Um die Isomorphie zweier Bäume festzustellen, wird zunächst eine Folge von Nullen und Einsen mit dem Algorithmus berechnet. Die Folge hat die Länge $2n$, wobei n die Anzahl der Knoten des Baumes ist. Wir nennen sie *Code* des Baumes.
Berechnung des Codes:
`Schritt 1:` Jedes Blatt, außer der Wurzel, bekommt den Code 01.
`Schritt 2:` Sei A_i die Codefolge des Knotens v_i. Ein Knoten mit den Söhnen v_1,v_2,...,v_t (von links nach rechts) bekommt den Code $0A_1A_2...A_t1$.

Beispiel 6.



2.2 Bestimmung der Isomorphie mithilfe der Code-Darstellung

Satz 1. *Gepflantze Bäum sind genau dann isomorph, wenn sie denselben Code haben.*

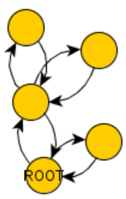
Beweis. Da wir nur Eigenschaften verwendet haben, die bei Isomorphie der Bäume gleich sind, haben isomorphe geplanzte Bäume denselben Code.
Zu zeigen bleibt, dass nicht isomorphe Bäume verschiedene Codes haben. Dazu beweisen wir, dass man den ursprünglichen Baum aus dem Code rekonstruieren können. Denn dann ist ein Baum eindeutig durch seinen Code bestimmt.
`Vollständige Induktion nach der Länge des Codes:` Induktionsbasis: Der kürzest mögliche Code lautet 01. Er entsteht nur, wenn die Wurzel keine Söhne hat, denn ansonsten werden im Schritt 2 zwischen 0 und 1 die Codes der Söhne eingefügt.

Induktionsannahme: Einem Code der Länge $\leq 2n$ lässt sich eindeutig ein gepflantzer Baum zuordnen.
Induktionsschritt: $n \to (n+1)$: Ein Code der Länge $2*(n+1)$ hat die Form $0A1$, mit $A{=}A_1A_2...A_t$. In der Zusammensetzung von A (also dem Code des Graphen, ohne die führende 0 und die abschließende 1) von links nach rechts gelesen, ist A_1 der kürzeste Codeteil, der gleich viele Nullen wie Einsen enthält. Nach A_1 schließt A_2 an, der ebenso der kürzeste Codeteil ist, der gleich viele Nullen wie Einsen enthält, usw. So erhälten wir den Code jedes Sohnes der Wurzel. Für jedes A_i, deren Code länger als zwei ist, wiederholen wir den Vorgang, sodass wir auch den Code der Söhne dieser Knoten bekommen. So erhalten wir schließlich den gesamten Baum bis zu den Blättern. Der Code ist demnach eindeutig. □

2.3 Einfache Decodieren eines gepflanzten Baumes

Will man einen so codierten Graphen zeichnen, gibt es eine alternative Methode zur Decodierung. Wir ersetzen alle Nullen des Codes durch einen Pfeil nach oben und alle Einsen durch einen Pfeil nach unten. Der erste Pfeil zeigt sozusagen auf die Wurzel und stellt noch keine Kante da, für die restlichen Pfeile nach oben zeichnen wir eine Kante nach oben (vom aktuellen Punkt aus) und für jeden Pfeil nach unten eine Kante nach unten (bzw. wir gehen im Graphen die Kante zurück). Bei dem letzten Pfeil sind wir allerdings schon wieder bei der Wurzel angekommen und zeichnen keine neue Kante nach unten. Gibt es vom aktuellen Punkt aus bereits eine Kante nach oben, so zeichnen wir die neue etwas weiter rechts.

Beispiel 7. *Wir haben den Code* 0001011011*, das entspricht also von der Wurzel ausgehend:* ↑↑↑↓↑↓↓↑↓↓.



3 Isomorphie von Wurzelbäumen - Algorithmus

3.1 Code-Darstellung

Auch Wurzelbäume lassen sich als Code darstellen. Der 1. Schritt bleibt derselbe (alle Blätter außer der Wurzel bekommen den Code 01), nur der 2. Schritt wird leicht verändert:
`Schritt 2'`: Haben schon alle Söhne w_i eines Knotens v einen Code $A(w_i)$, dann ordnen wir die Söhne so, dass $A(w_1) \leq A(w_2) \leq ... \leq A(w_t)$. Der Knoten v bekommt dann den Code $0A_1A_2...A_t1$, wobei $A_i :\equiv A(w_i)$.
Die Relation $A \leq B$ bedeutet hier, dass die Folge A in einer Ordnung vor B kommt. Die Ordnung muss vorher festgelegt werden und für alle endlichen Folgen von Nullen und Einsen gelten. Dafür gibt es verschiedene Möglichkeiten.

Wir verwenden die "lexikographische Ordnung".
Die Folgen $A = (a_1, a_2, ..., a_n)$ und $B = (b_1, b_2, ..., b_m)$ werden dabei so verglichen: Ist A ein Anfangsstück von B (z.B.: A =01,B = 011) dann gilt $A < B$. Wenn B ein Anfangsstück von A ist gilt $B < A$. Ist beides nicht der Fall betrachten wir den kleinsten Index i für den gilt: $a_i \neq b_i$. Ist $a_i < b_i$ dann gilt: $A < B$, ist $b_i < a_i$ dann gilt: $B < A$.

3.2 Bestimmung der Isomorphie mithilfe der Code-Darstellung

[Matoušek, Nešetřil: Diskrete Mathematik] Seiten 176-180

Satz 2. *Wurzelbäume sind genau dann isomorph, wenn sie denselben Code haben.*

Beweis. Wir können zeigen dass, wie auch bei gepflanzten Bäumen, ein Wurzelbaum durch seinen Code eindeutig bestimmt ist. Der Beweis geht äquivalent. Auch hier haben wir bei der Codierung nur Eigenschaften verwendet, die bei isomorphen Bäumen gleich bleiben (wir müssen nur immer dieselbe Ordnung verwenden, dann bleibt die Reihung der Söhne bei isomorphen Bäumen gleich). Die Decodierung verläuft ebenso, wie bei gepflanzten Bäumen und ist daher eindeutig (es ließe sich wieder mit Induktion zeigen). Der einzig mögliche Unterschied zwischen dem ursprünglichen Baum und dem decodierten Baum ist die Reihenfolge der Söhne (da wir künstlich eine Ordnung eingeführt haben). Diese ist aber unerheblich, da bei Wurzelbäumen die Reihenfolge der Söhne nicht vorgegeben ist (und sie keine Ordnung haben). Der decodierte Baum ist also sicher isomorph zum ursprünglichen. □

4 Isomorphie von Bäumen ohne Wurzel- Algorithmus

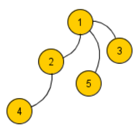
4.1 Weitere Grundlagen des Algorithmus

Bevor wir nun den Algorithmus für Bäume ohne Wurzel betrachten, definieren wir einige nützliche Begriffe.

Definition 14. *Sei v der Knoten eines Graphen G. Das Maximumum der Abstände von v zu allen anderen Knoten wird* ***Exzentrizität*** *genannt. Für die Exzentrizität von v schreiben wir $ex_G(v)$.*

Definition 15. *Die Menge aller Knoten des Graphen G mit minimaler Exzentrizität wird* ***Zentrum*** *genannt. Wir schreiben $C(G)$ für das Zentrum des Graphen G.*

Beispiel 8. *In diesem Graphen ist die minimale Exzentrizität 2 und das Zentrum besteht aus den Knoten* 1 *und* 2. *Die restlichen Knoten haben Exzentrizität 3.*



Satz 3. *In jedem Baum T enthält die Menge $C(T)$ maximal zwei Knoten. Wenn $C(T)$ zwei Knoten x und y enthält, so existiert eine Kante $\{x, y\}$ in T.*

Beweis. Sei $T = (V, E)$ ein Baum. Hat T zwei oder weniger Knoten, so ist das Zentrum gleich der (gesamten) Knotenmenge. Hat T zwei Knoten x und y, so gibt es eine Kante $\{x, y\}$.
Hat C(T)T mehr als zwei Knoten konsturieren wir einen Baum $T' = (V', E')$, indem wir die Blätter des Baumes T entfernen.

$$V' = x \in V : deg_T(x) > 1$$

und V' kann nicht die leere Menge sein, da bei mehr als zwei Knoten nicht jeder Knoten ein Blatt sein kann.

$$E' = x, y \in E : deg_T(x) > 1 \text{ } und \text{ } deg_T(y) > 1.$$

Da für alle Knoten v der am weitesten entfernte Knoten ein Blatt sein muss gilt:

$$ex_T(v) = ex_{T'}(v) + 1$$

Das heißt: $C(T) = C(T')$.
Diese Konstruktion wiederholen wir so lange, bis T' nicht mehr als 2 Knoten hat. □

Mit Hilfe des Zentrums lässt sich auch ein eindeutiger Code für einen Baum ohne Wurzel erstellen:

4.2 Code-Darstellung

Besteht das Zentrum des Baumes $T = (E, V)$ nur aus einem einzelnen Knoten v, dann ist der Code des Baumes derselbe wie der Code des Wurzelbaumes (T, v).
Besteht das Zentrum des Baumes T aus zwei Knoten x und y, die durch die Kante $e = \{x, y\}$ verbunden sind, dann betrachten wir den Graphen $G = (E', V)$ mit $E' = E$ *ohne* $\{x, y\}$. Da T ein Baum war muss G aus zwei Zusammenhangskomponenten bestehen, die wieder Bäume sind. Wir nennen sie T_x und T_y, wobei T_x die Komponente ist, die den Knoten x enthält und T_y demnach den Knoten y enthält.
Sei A der Code des Wurzelbaums (T_x, x) und B der Code des Wurzelbaumes (T_y, y). Ist in der lexikographischen Ordnung $A \leq B$, dann erhält T den Code des Wurzelbaumes (T, x), ansonsten bekommt T den Code des Wurzelbaumes (T, y).

4.3 Bestimmung der Isomorphie mithilfe der Code-Darstellung

Satz 4. *Bäume sind genau dann isomorph, wenn sie denselben Code haben.*

Beweis. Sind zwei Bäume isomorph, so ist auch ihr Zentrum isomorph. Also haben wir auch hier bei der Codierung nur Eigenschaften verwendet, die bei isomorphen Bäumen gleich bleiben. Die Decodierung verläuft genauso wie bei Wurzelbäumen und gepflanzten Bäumen und ist demnach eindeutig. □

[Matoušek, Nešetřil: Diskrete Mathematik] Seiten 178-181

Die Codierung zweier Bäume und der anschließende Vergleich ihrer Codes ist also der gesuchte Algorithmus, mit dem wir die Isomorphie zweier Bäume überprüfen können. Die Dauer der Ausführung hängt nur von der Anzahl der Knoten des Baumes ab und solange wir zwei endlichen Graphen überprüfen ist unser Algorithmus effizient.

Literatur

[Matoušek, Nešetřil: Diskrete Mathematik] Jiří Matoušek, Jaroslav Nešetřil. *Diskrete Mathematik - Eine Entdeckungsreise.* Springer-Lehrbuch, Berlin, Heidelberg 2005