# Bachelorarbeit

Andreas Windorfer

3. Mai 2020

Zusammenfassung

# Inhaltsverzeichnis

# 1 Fazit

## 1.1 Rot-Schwarz-Baum

Der Rot-Schwarz-Baum gehört zur Gruppe der **balancierten BST** und erfüllt alle Eigenschaften um ihn als Hilfsstruktur im Tango-Baum zu verwenden. Bei balancierten BST gilt für die Höhe $h = O(n)$, mit $n =$ Anzahl der Knoten. Jeder Knoten benötigt ein zusätzliches Attribut, um eine Farbinformation zu speichern. Der Name der Datenstruktur kommt daher, dass die beiden durch das zusätzliche Attribut unterschiedenen Zustände als *Rot* und *Schwarz* bezeichnet werden. Die Farbe ist also eine Eigenschaft der Knoten und im folgenden wird einfach von roten bzw. schwarzen Knoten gesprochen. Innerhalb mancher Operationen wird von einem Knoten aus direkt auf dessen Vater zugegriffen, so dass man sich im Baum auch nach oben hin bewegen kann. In Implementierungen wird das so umgesetzt, dass es zusätzlich zu den beiden Zeigern auf die Kinder noch einen zum Vater gibt. Als Blätter werden schwarze Sonderknoten verwendet, deren Schlüssel auf *null* gesetzt wird, um sie eindeutig erkennen zu können. Fehlende Kinder von Knoten mit gewöhnlichem Schlüssel werden durch solche Blätter ersetzt.
Folgende zusätzliche Eigenschaften müssen bei einem Rot-Schwarz-Baum erfüllt sein.

1. Jeder Knoten ist entweder rot oder schwarz.
2. Die Wurzel ist schwarz.
3. Jedes Blatt (Sonderknoten) ist schwarz.
4. Beide Kinder eines roten Knotens sind schwarz.
5. Für jeden Knoten gilt, dass alle Pfade, die an ihm starten und an einem Blatt (Sonderknoten) enden, die gleiche Anzahl an schwarzen Knoten enthalten.

Sei $(v_0, v_1, ..., v_n)$ ein Pfad von einem Knoten $v_0$ zu einem Blatt $v_n$. Die Anzahl der schwarzen Knoten innerhalb $(v_1, ..., v_n)$ wird als **Schwarz-Höhe** $bh(v_0)$ von Knoten $v_0$ bezeichnet. Die eigene Farbe des betrachteten Knotens bleibt dabei also außen vor. Dadurch hat ein Knoten die gleiche Schwarz-Höhe wie ein rotes Kind und eine um eins erhöhte Schwarz-Höhe gegenüber einem schwarzen Kind. Die Schwarz-Höhe eines Knoten $x$ ist genau dann eindeutig wenn er Eigenschaft 5 nicht verletzt. Hält $x$ Eigenschaft 5 ein, so gilt $bh(x) = i$ wenn $x$ rot ist und $bh(x) = i - 1$ wenn $x$ schwarz ist. Ist $bh(x)$ eindeutig, so enthält jeder Pfad der mit $x$ startet und an einem Blatt endet

$bh(x) + 1$ schwarze Knoten, wenn $x$ schwarz ist und $bh(x)$ wenn $x$ rot ist. Im folgenden wird **RBT** (Red-Black-Tree) als Abkürzung für Rot-Schwarz-Baum verwendet.

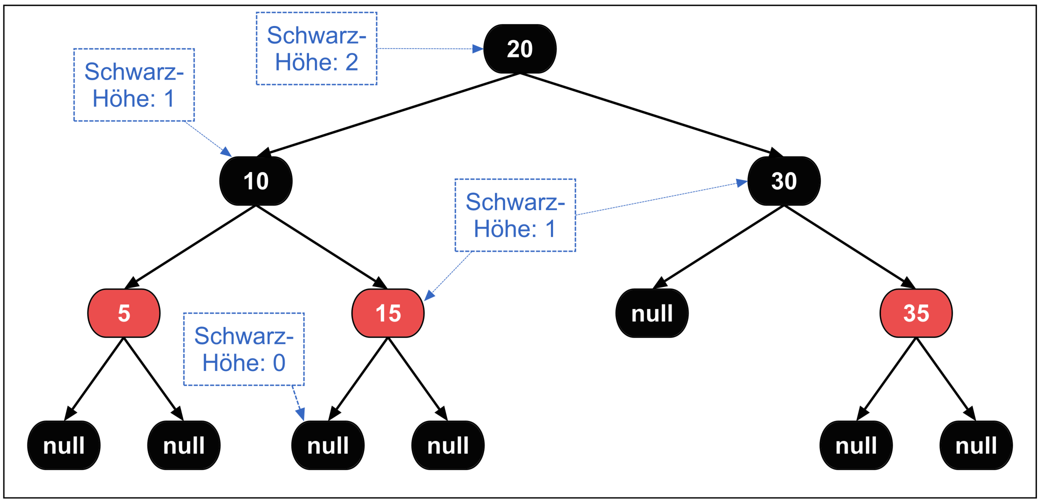


**Abbildung 1:** Rot-Schwarz-Baum ohne Verletzung von Eigenschaften.

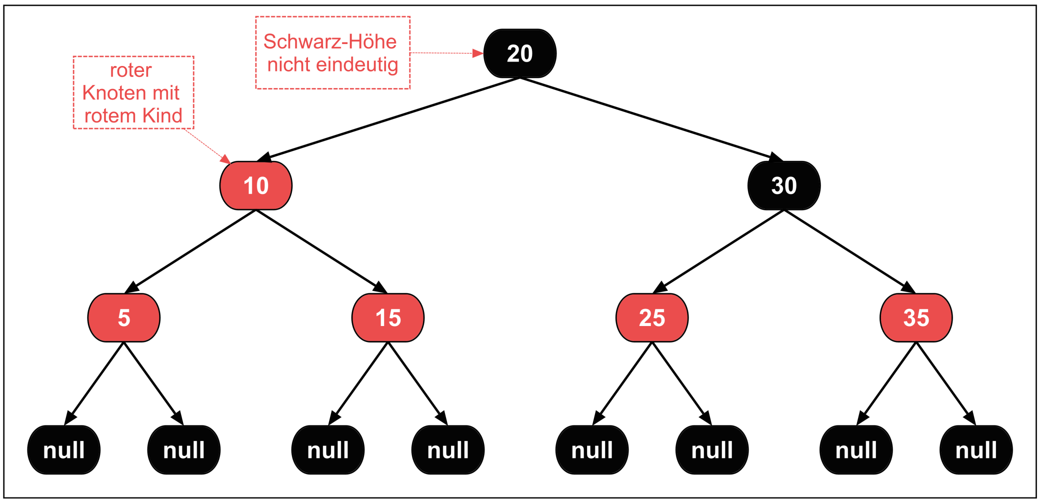


**Abbildung 2:** Rot-Schwarz-Baum bei dem Eigenschaft vier und fünf verletzt sind.

### 1.1.1 Suchen im Rot-Schwarz-Baum

Die Suche unterscheidet sich nur in einem Punkt von der in **??** vorgestellten. Wird nach einem Schlüssel gesucht, der im RBT nicht vorhanden ist, so wird

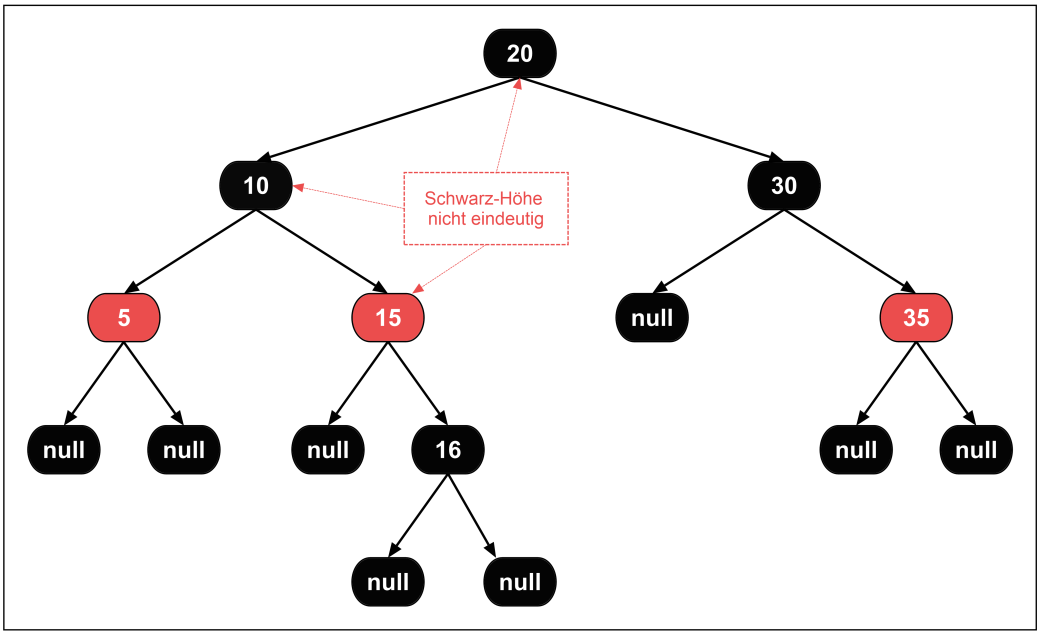


**Abbildung 3:** Rot-Schwarz-Baum bei dem Eigenschaft fünf verletzt ist.

einer der Sonderknoten erreicht. In diesem Fall wird die Suche abgebrochen ohne einen Knoten zurückzugeben. Die Operation verändert den RBT nicht.

### 1.1.2 Einfügen in den Rot-Schwarz-Baum

Sei $k$ der einzufügende Schlüssel. Zunächst wird wie beim Suchen vorgegangen. Wird $k$ gefunden wird der RBT nicht verändert. Ansonsten wird ein Sonderknoten $b$ erreicht. Ein neu erzeugter roter Knoten mit Schlüssel $k$ nimmt den Platz von $b$ ein. $k$ werden Sonderknoten als linkes und rechtes Kind angefügt. $k$ ist nun im Baum enthalten, es muss jedoch auf mögliche Verletzungen der fünf Eigenschaften geachtet werden. Welche können betroffen sein ?

1. Es ist immer noch jeder Knoten entweder rot oder schwarz.
2. Wurde in den leeren Baum eingefügt, so ist der neu eingefügte rote Knoten die Wurzel, was eine Verletzung darstellt. Waren bereits Knoten im Baum vorhanden bleibt die Wurzel unverändert.
3. Aufgrund der Sonderknoten sind die Blätter immer noch schwarz.
4. Der Baum wird nur direkt an der Einfügestelle verändert. Der neue Knoten hat schwarze Kindknoten, er könnte jedoch einen roten Vater haben, so dass diese Eigenschaft verletzt wäre.

5. Da der neue Knoten rot ist, ändern sich keine Schwarz-Höhen von bereits enthaltenen Knoten. Die Schwarz-Höhe des neuen Knoten ist immer 1. Eigenschaft fünf bleibt also erhalten.

Es können also die Eigenschaften zwei und vier betroffen sein. Jedoch nur eine von ihnen, denn Eigenschaft zwei wird genau dann verletzt wenn der neue Knoten der Einzige im Baum ist. Dann kann er aber keinen roten Vorgänger haben.
Zur Korrektur wird eine zusätzliche Operation, **einfügen-fixup** eingesetzt. Diese Operation arbeitet sich von der Einfügestelle solange nach oben in einer Schleife durch, bis alle Eigenschaften wieder erfüllt sind. Die Schleifenbedingung ist, dass eine Verletzung vorliegt. Dazu muss geprüft werden ob der betrachtete Knoten die rote Wurzel des Gesamtbaumes ist, oder ob er und sein Vater beide rot sind. Beim ersten Durchlauf wird der neu eingefügte Knoten übergeben. Innerhalb der Schleife werden sechs Fälle unterschieden. Im folgenden wird auf vier Fälle detailliert eingegangen. Die restlichen zwei verhalten sich symmetrisch zu einem solchen. Jeder der Fälle verantwortet, dass zum Start der nächsten Iteration wieder nur maximal eine der beiden Eigenschaften zwei oder vier verletzt sein kann und Eigenschaft vier höchstens einmal verletzt wird. Die Fallauswertung geschieht in aufsteigender Reihenfolge. Deshalb kann man innerhalb einer Fallbehandlung verwenden, dass die vorherigen Fallbedingungen nicht erfüllt sind. Eigenschaft eins bleibt in der Beschreibung außen vor, da es während der gesamten Laufzeit der Operation nur Knoten gibt, die entweder rot oder schwarz sind.

**Fall 1: Die Wurzel ist rot:** Dieser Fall wird behandelt in dem man die Wurzel schwarz färbt. Man muss noch zeigen, dass es durch das Umfärben zu keiner anderen Verletzung gekommen ist.

Betrachtung der Eigenschaften:

1. -

2. Die Wurzel wurde schwarz gefärbt.

3. Die Blätter sind unverändert.

4. Es wurden weder rote Knoten hinzugefügt, noch wurden Zeiger umgesetzt.

5. Das Umfärben der Wurzel kann hierauf keinen Einfluss haben, da sie in der Berechnung der Schwarz-Höhe jedes Knotens außen vor ist.

Es wird also keine Eigenschaft mehr verletzt und die Schleife wird keine weitere Iteration durchführen.
Die Fälle 2 - 6 behandeln nun die Situation zweier aufeinanderfolgender roter Knoten. Der untere rote Knoten wird als $x$ bezeichnet, der obere als $y$. Da Eigenschaft fünf nach jeder Iteration erfüllt ist muss $y$ einen Bruder haben. Denn da $y$ rot ist und Fall 1 nicht ausgewählt wurde, kann es nicht die Wurzel sein. Also muss auch $y$ einen Vorgänger $z$ haben. Da $z$ kein Blatt(Sonderknoten) ist, müssen beide Kinder vorhanden sein.

**Fall 2: $y$ hat einen roten Bruder:** Diesen Fall veranschaulicht Abbildung 4. Da $y$ rot ist, muss $z$ schwarz sein, ansonsten wäre Eigenschaft vier mehrfach verletzt gewesen. Nun wird $z$ rot gefärbt und beide Kinder von $z$, also $y$ und dessen Bruder, schwarz. Somit ist der Vater von $x$ nun einen schwarz und die Verletzung der Eigenschaft vier wurde an dieser Stelle behoben. Wie sieht es aber mit den Verletzungen insgesamt aus ?

Betrachtung der Eigenschaften:

1. -
2. Wenn $z$ die Wurzel des Baumes ist, wurde sie rot gefärbt und eine Verletzung liegt vor.
3. Der rot umgefärbte Knoten $z'$ hat zwei Kinder, somit wurde kein Blatt rot gefärbt.
4. Wenn der rot gefärbte Knoten $z'$ nicht die Wurzel ist, könnte er einen roten Vater haben und Eigenschaft vier ist weiterhin Verletzt. Das Problem liegt nun aber zwei Baumebenen höher.
5. Die Schwarz-Höhen der Vorfahren von $z'$ bleiben unverändert, da jeder Pfad von ihnen zu einem Blatt auch entweder $y'$ oder dessen Bruder enthält. $z'$ Schwarz-Höhe steigt um eins gegenüber $z$, bleibt aber eindeutig. An keinem anderen Knoten ändert sich die Schwarz-Höhe.

Es kann also wieder nur entweder Eigenschaft zwei oder vier verletzt sein. Wenn das Problem noch nicht an der Wurzel ist, liegt es zumindest zwei Ebenen näher daran.
**Fall 3: $y$ ist ein linkes Kind. $x$ ist ein linkes Kind:**
Abbildung 5 zeigt eine entsprechende Situation. Da in dieser Situation, die Wurzel schwarz sein muss, gibt es den Vater $z$ von $y$. Da es nur eine Stelle im Baum geben kann an der Eigenschaft vier verletzt wird, muss $z$ schwarz

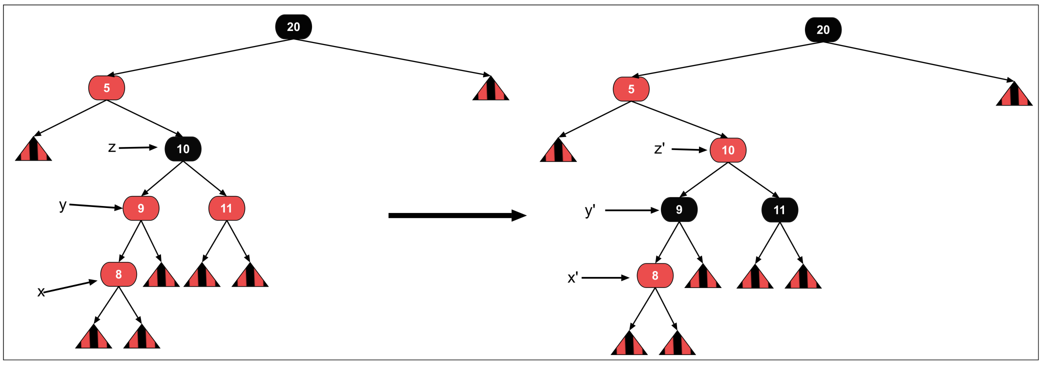


**Abbildung 4:** einfügen-fixup. Dargestellt ist Fall 2

sein. Es wird nun eine Rechtsrotation auf $y$ ausgeführt. Anschließend wird $z$ rot gefärbt und $y$ schwarz.

Betrachtung der Eigenschaften:
Dazu werden vier weitere Variablen auf Knoten verwendet. Es zeigt auf $xl$ das linke Kind von $x$, $xr$ entsprechend das rechte Kind. $yr$ und $zr$ bezeichnen die rechten Kinder von $y$ bzw. $z$. Nachfolgend wird verwendet, dass die Teilbäume mit den Wurzeln $xl$, $xr$, $yr$ und $zr$ durch die Ausführung unverändert bleiben.

1. -

2. Wenn $z$ zu Beginn nicht die Wurzel des Gesamtbaumes war, bleibt diese unverändert. Ansonsten wurde durch die Rotation $y'$ zur neuen Wurzel und $y'$ wurde schwarz gefärbt.

3. Alle vier Plätze in der zweiten Ebene unter $z'$ werden von den unveränderten Teilbäumen mit den Wurzeln $xl$, $xr$, $yr$ oder $zr$ besetzt. An den Blättern verändert sich also durch die Ausführung nichts.

4. Knoten $x'$ ist linkes Kind des schwarzen $y'$. $x'$ Teilbäume blieben unverändert. Der linke Teilbaum von $y'$ enthält somit keine aufeinanderfolgenden roten Knoten. Das rechte Kind von $y'$ ist der rote $z'$. Rechts an $z'$ hängt nun ein unveränderter Teilbaum, dessen Wurzel zuvor Bruder von $y$ war. Dieser ist nach Fallunterscheidung ein schwarzer Knoten. Links hängt ebenfalls ein unveränderter Teilbaum, dessen Wurzel zuvor rechter Nachfolger von $y$ war. Der rechte Nachfolger von $y$ muss schwarz sein, ansonsten wäre Eigenschaft vier an zwei Stellen verletzt gewesen. Im Teilbaum mit Wurzel $y$ gibt es also keine aufeinanderfolgenden roten Knoten. Da $y'$ schwarz gefärbt wurde, kann auch außerhalb des Teilbaumes mit $y'$ keine neue Verletzung entstanden sein.

5. Es gilt $bh(xl) = bh(xr) = bh(yr) = bh(zr) = bh(z) - 1$. Wie oben bereits erwähnt wird die zweite Ebene unter der Wurzel von den Knoten $xl'$, $xr'$, $yr'$ und $zr'$ gebildet. Es müssen also lediglich die Knoten $x'$, $y'$ und $z'$ betrachtet werden. An $x'$ und an $z'$ folgen schwarze Knoten mit der Schwarz-Höhe $bh(z) - 1$. Die Schwarz-Höhen von $x'$ und $z'$ sind also eindeutig und es gilt $bh(x') = bh(z') = bh(z)$. $y'$ Kinder sind die roten Knoten $x'$ und $z'$. Da beide Kinder rot sind gilt $bh(y') = mathitbh(x') = bh(z)$. Somit sind alle Schwarz-Höhen im betrachteten Teilbaum eindeutig. Die neue Wurzel der Teilbaumes $y'$ hat die gleiche Schwarz-Höhe und die gleiche Farbe wie die vorherige Wurzel $z$. Damit kann es auch im Gesamtbaum zu keiner Verletzung der Eigenschaft gekommen sein.

Es ist keine der Eigenschaften verletzt, daher wird es zu keiner Iteration mehr kommen.
**Fall 4: $y$ ist ein linkes Kind. $x$ ist ein rechtes Kind:**
Dieser in Abbildung 6 gezeigte Fall wird so umgeformt, dass eine Situation entsteht bei der Fall drei angewendet werden kann. Dazu wird eine Linksrotation an Knoten $x$ durchgeführt.

Betrachtung der Eigenschaften:
Zu Veränderungen kommt es durch die Rotation lediglich im linken Teilbaum von $z$. Es sei $xl$ das linke Kind von $x$, und $xr$ das rechte Kind von $x$. $yl$ ist das linke Kind von $y$. $xl$, $xr$ und $yl$ müssen schwarz sein, ansonsten wäre Eigenschaft vier mehrfach verletzt gewesen.

1. -

2. Die Wurzel bleibt unverändert.

3. Die Teilbäume mit den Wurzeln $xl$, $xr$ und $yl$ enthalten alle Blätter innerhalb des linken Teilbaumes von $z$. Die Teilbäume $xl$, $xr$ und $yl$ bleiben durch die Rotation unverändert und $xl'$, $xr'$ und $yl'$ enthalten auch alle Blätter des linken Teilbaumes von $z'$.

4. Da $x$ und $y$ rot sind müssen $z$, $xl$ und $xr$ schwarz sein. Nach der Rotation ist $y'$ linkes Kind von $x'$. $x'$ ist Kind vom schwarzen $z'$. Alle verbleibenden Kinder von $x'$ und $y'$ werden durch die unveränderten Teilbäume $xl'$, $xr'$ und $yl'$ gebildet. Deren Wurzeln müssen schwarz sein, ansonsten hätte es in ursprünglichen Baum an mehr als einer Stelle eine Verletzung von Eigenschaft vier gegeben. Durch die Rotation verbleibt es also bei einer Verletzung der Eigenschaft vier in der gleiche Baumebene. Die beiden beteiligten roten Knoten sind nun aber jeweils linke Kinder.

5. $bh(yl) = bh(xl) = bh(xr) = bh(yl') = bh(xl') = bh(xr')$. Die Schwarz-Höhen von $x$ und $y$ bleiben unverändert. Damit kommt es auch bei $z$ zu keiner Veränderung bei der Schwarz-Höhe.

Es sind also weiterhin zwei rote aufeinanderfolgende rote Knoten in den gleichen Baumebenen vorhanden. Diese sind nun aber beides linke Kinder. Der Bruder des oberen roten Knotens ist der selbe schwarze Knoten wie vor der Ausführung von Fall 4. Damit kann direkt mit dem bearbeiten von Fall 3 begonnen werden.

**Fall 5: $y$ ist ein rechtes Kind. $x$ ist rechtes Kind:**
Links/Rechts-Symmetrische Ausgangssituation zu Fall 3.
**Fall 6: $y$ ist ein rechtes Kind. $x$ ist linkes Kind:**
Links/Rechts-Symmetrische Ausgangssituation zu Fall 4.

Sei $h$ die Höhe des Gesamtbaumes vor Aufruf von einfügen-Fixup. Fall 2 kann maximal $\frac{h}{2}$ mal ausgewählt werden, bevor $x$ oder $y$ an der Wurzel liegt. Nach einer Iteration bei der nicht Fall 2 ausgewählt wird, terminiert einfügen-fixup.

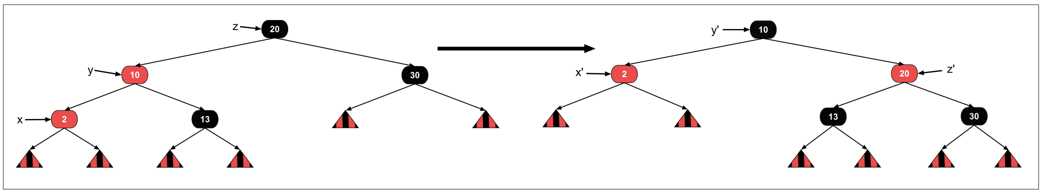


**Abbildung 5:** einfügen-fixup. Dargestellt ist Fall 3

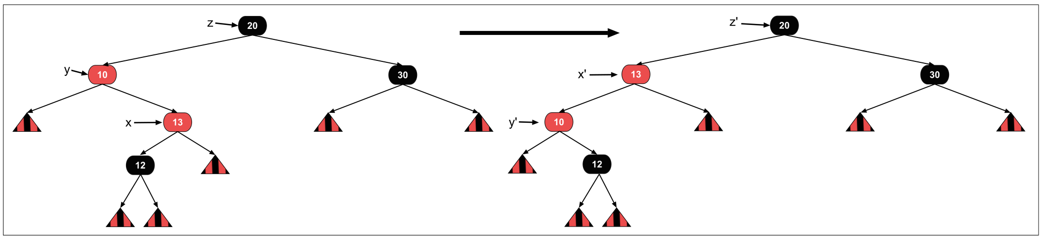


**Abbildung 6:** einfügen-fixup. Dargestellt ist Fall 4

### 1.1.3 Löschen aus dem Rot-Schwarz-Baum

Evtl. noch machen, für tango eigentlich nicht notwendig

### 1.1.4 Laufzeit der Grundoperationen

Zu Beginn des Kapitels wurde erwähnt, dass für die Höhe $h$ eines RBT mit $n$ Knoten $h = O(\log n)$ gilt. Das wird nun gezeigt.

**Lemma 1.1.** *Für die Höhe $h$ eines RBT $T$ mit $n$ Knoten gilt $h = O(\log n)$ .*

*Beweis.* Sei $w$ die Wurzel von $T$ und $m$ Zunächst wird gezeigt, dass T mindestens $2^{bh(w)} - 1$ innere Knoten enthält. Dies geschieht mit Induktion über $h$. Für $h = 0$ mit $2^0 - 1 = 0$ stimmt die Behauptung, denn der Baum ist leer. Für $h = 1$ mit $2^0 - 1 = 0$ stimmt die Behauptung ebenfalls, denn der Baum besteht aus einem einzigem Blatt. Induktionsschritt mit Höhe $h + 1$:
Sei $tl$ der linke Teilbaum von $w$ und $tr$ der rechte Teilbaum von $w$. Im Induktionsschritt kann nun verwendet werden, dass $h > 1$ gilt und $w$ ein innerer Knoten sein muss. $tl$ und $tr$ haben Schwarz-Höhe $bh(w) - 1$ wenn ihre Wurzel schwarz ist und Schwarz-Höhe $bh(w)$ wenn ihre Wurzel rot ist. Ihre Höhe ist $h - 1$ und somit enthalten sie nach Induktionsnahme mindestens $2^{bh(w)-1} - 1$ innere Knoten. Aufaddieren ergibt die Behauptung.

$$m \geq 2^{bh(w)-1} - 1 + 1 + 2^{bh(w)-1} - 1 = 2^{bh(w)} - 1$$

Für einen Knoten $x$ gilt folgender Zusammenhang, da höchstens jeder zweite Knoten in einem Pfad rot sein kann

$$h(x) \leq 2 \cdot bh(x) + 1$$

Es gilt also:

$$\begin{aligned} m &\geq 2^{bh(x)} - 1 \\ \Rightarrow 2 \cdot log_2(m+1) &\geq h(x) + 1 \end{aligned}$$

Damit gilt $h = O(\log m)$. Mit Blättern steigt die Höhe des Baumes um eins, damit ist auch das Lemma gezeigt.

□

# Literatur