

Lernaufgabe 1a

Beim Kofferpacken können wir durch Druck die Luft herauslassen und bringen somit mehr Kleidungsstücke in den Koffer. Im Unterricht haben wir vom Huffman-Code gehört und wissen, dass er nach dem gleichen Prinzip funktioniert.

Aufgabe:

Wie muss der Eingabetext beschaffen sein, damit er sich gut komprimieren lässt?

Vorgehen:

- Verwenden Sie zum Lösen dieser Aufgabe das Applet unter der URL <http://www.swisseduc.ch/informatik/interaktiv/kompression/applet.html>
Eine Kurzanleitung finden Sie auf dem Beiblatt.
- Komprimieren Sie mehrere Wörter, um so auf die Lösung zu kommen. Zum Beispiel mit:
 - können
 - Schiffahrtsgesellschaft
 - Griffbrett
 - Maus
 - AAAAAABBCCED
 - ABCDEFGHIJKL
- Formulieren Sie Ihre Antwort in höchstens drei ganzen Sätzen.
- Sollten Sie die Lösung innert 10 Minuten nicht finden, so fragen Sie Ihre Kolleginnen und Kollegen

Richtige Lösung:

Die Häufigkeit spielt eine zentrale Rolle beim Huffman-Code. Es ist wichtig, dass einige Buchstaben sehr häufig auftreten und andere nur selten.

Lernaufgabe 1b

In der letzten Lernaufgabe haben wir gesehen, dass die Häufigkeit der Buchstaben eine zentrale Rolle spielt und wollen dies jetzt genauer untersuchen.

Aufgabe:

Vervollständigen Sie folgende Sätze. Hinter den Lücken finden Sie in Klammern mehrere Vorschläge.

- Häufig auftretende Buchstaben haben einen _____ (langen/kurzen) Code und seltene Buchstaben einen _____ (langen/kurzen) Code.
- Buchstaben mit einer kleinen Häufigkeit finden wir _____ (unten/oben/gar nicht) im Baum.
- Je weiter unten im Baum die Buchstaben angeordnet sind, desto _____ (länger/kürzer) ist ihr Code.

Vorgehen:

- Verwenden Sie zum Lösen dieser Aufgabe das Applet unter der URL <http://www.swisseduc.ch/informatik/interaktiv/kompression/applet.html>
Eine Kurzanleitung finden Sie auf dem Beiblatt
- Versuchen Sie die Lösung anhand der Komprimierung von AAAAAABBCCED herauszufinden.
- Sollten Sie die Lösung innert 5 Minuten nicht finden, so fragen Sie Ihre Kolleginnen und Kollegen

Richtige Lösung:

- Häufig auftretende Buchstaben haben einen **kurzen** Code und seltene Buchstaben einen **langen** Code.
- Buchstaben mit einer kleinen Häufigkeit finden wir **unten** im Baum.
- Je weiter unten im Baum die Buchstaben angeordnet sind, desto **länger** ist ihr Code.

Lernaufgabe 2a

Wir haben in den vorhergehenden Aufgaben gesehen, dass Buchstaben mit grosser Häufigkeit einen kurzen Code und solche mit kleiner Häufigkeit einen langen Code bekommen. Den Code der einzelnen Buchstaben können wir im Huffman-Baum ablesen. Wir wissen, dass dieser die Buchstaben optimal anordnet, aber noch nicht wie dieser konstruiert wird.

Aufgabe:

Beschreiben Sie in Worten, wie der Huffman-Baum aufgebaut wird.

Vorgehen:

- Verwenden Sie zum Lösen dieser Aufgabe das Applet unter der URL <http://www.swisseduc.ch/informatik/interaktiv/kompression/applet.html>
Eine Kurzanleitung finden Sie auf dem Beiblatt
- Versuchen Sie die Lösung anhand des "Wortes" AAAAAABBCCED herauszufinden.
- Sollten Sie die Lösung innert 15 Minuten nicht finden, so fragen Sie Ihre Kolleginnen und Kollegen.

Richtige Lösung:

- Die Buchstaben werden nach der relativen Häufigkeit ihres Auftretens sortiert.
- Solange bis nur noch ein Baum vorhanden ist, werden die beiden Teilbäume mit der kleinsten relativen Häufigkeit verschmolzen. Die Häufigkeit der neuen Wurzel erhalten wir durch das Summieren der Wahrscheinlichkeit der beiden verschmolzenen Teilbäumen.

Lernaufgabe 2b

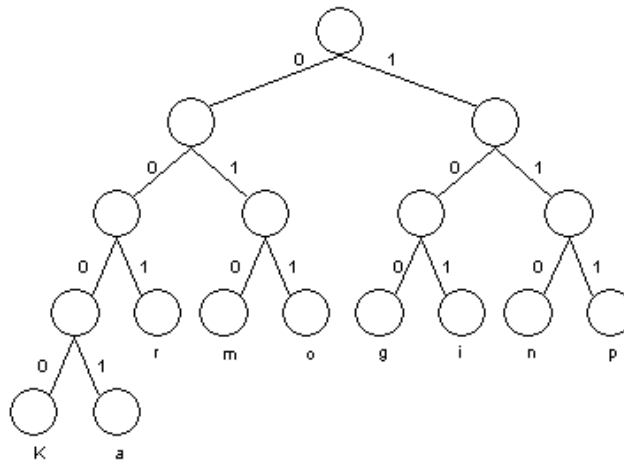
Der Huffman-Code erlaubt es, eine codierte Folge von Nullen und Einsen sehr einfach wieder in den ursprünglichen Text zu verwandeln. Dazu braucht man den Codebaum, der bei der Komprimierung verwendet wurde.

Aufgaben:

- Versuchen Sie, den folgenden komprimierten String:

00001011100111110010111000010001010010

zu dekomprimieren. Benutzen Sie dazu den folgenden Huffman-Baum:



Was war der ursprüngliche Text?

- Beschreiben in drei bis vier Sätzen, wie Sie bei der Dekomprimierung vorgegangen sind.

Vorgehen:

- Sollten Sie die Lösung innert 15 Minuten nicht finden, so fragen Sie Ihre Kolleginnen und Kollegen.

Richtige Lösung:

- Kinoprogramm
- Starten Sie bei der Wurzel im Code-Baum. Ist die erste Ziffer im komprimierten String eine "0", so steigen Sie in den linken Teilbaum ab, sonst in den rechten. Wiederholen Sie diesen Vorgang, bis Sie nicht weiter absteigen können. Somit haben Sie den ersten Buchstaben dekomprimiert. Wiederholen Sie dieses Prozedere, bis der ganze String abgearbeitet ist.

Anmerkung:

Der Huffman-Baum stellt sicher, dass es nur eine Möglichkeit gibt, den String zu dekomprimieren! Etwas technischer ausgedrückt, garantiert der Huffman-Baum, dass der daraus resultierende Code *präfixfrei* ist. Das heisst, kein Codewort ist der Anfang eines anderen Codewortes. Wäre dem nicht so, könnte man nicht so einfach wie oben beschrieben decodieren: man wüsste nicht, ob man in einem Knoten oder schon in einem Blatt ist!