Modul 114

Thema 7/11

Speicherplatz als rares Gut - Kompression

Agenda

Thema	Inhalte
1	Zahlensysteme BIN - DEZ - HEX
2	Arithmetische und logische Grundoperationen im Binärsystem
3	**Die Logik und den Prozessor verstehen**
4	Grosse Zahlen in kleinen Variablen ablegen
5	Fehler in der Datenübertragung finden und korrigieren
6	Speicherplatz als rares Gut - Dateien und ihr Platzbedarf
7	Speicherplatz als rares Gut - Kompression
8	Speicherplatz als rares Gut - Reduktion
9	Vektorgrafiken - Eine Alternative zu den Pixeln
10	Verschlüsselung - Geschichte und Grundsätzliches
11	Verschlüsselung – Moderne Verfahren

Tagesziele

Ich kann...

- den Unterschied zwischen verlustfreier und verlustbehafteter Kompression erklären.
- die Huffman-Codierung auf einfache Texte anwenden.
- den Kompressionsfaktor und die Kompressionsrate berechnen.

Kompression allgemein

Kompression: verlustfrei…

...oder verlustbehaftet

Grundsätzliches

Wikipedia: Die Datenkompression oder Datenkomprimierung ist ein Vorgang, bei dem die Menge digitaler Daten reduziert wird.

Ziele:
- Speicherplatz wird eingespart
- Übertragungszeit wir kürzer

Grundsatz: Es wird versucht, überflüssige Informationen zu vermeiden oder zu entfernen.

Zwei Arten:

Verlustfreie Kompression
Es werden Redundanzen ausgenutzt, es geht keine Information verloren.

Verlustbehaftete Kompression
Es wird «irrelevante» Information weggelassen
(Bild-, Ton- und Videoübertragung)

Verlustfreie Kompression

Lösungsansätze für Binärwerte

Lauflängencodierung

Gibt an, wie viele Nullen oder Einsen hintereinander kommen und versucht so, Platz zu sparen. Diese Methode bewährt sich nur in Spezialfällen!

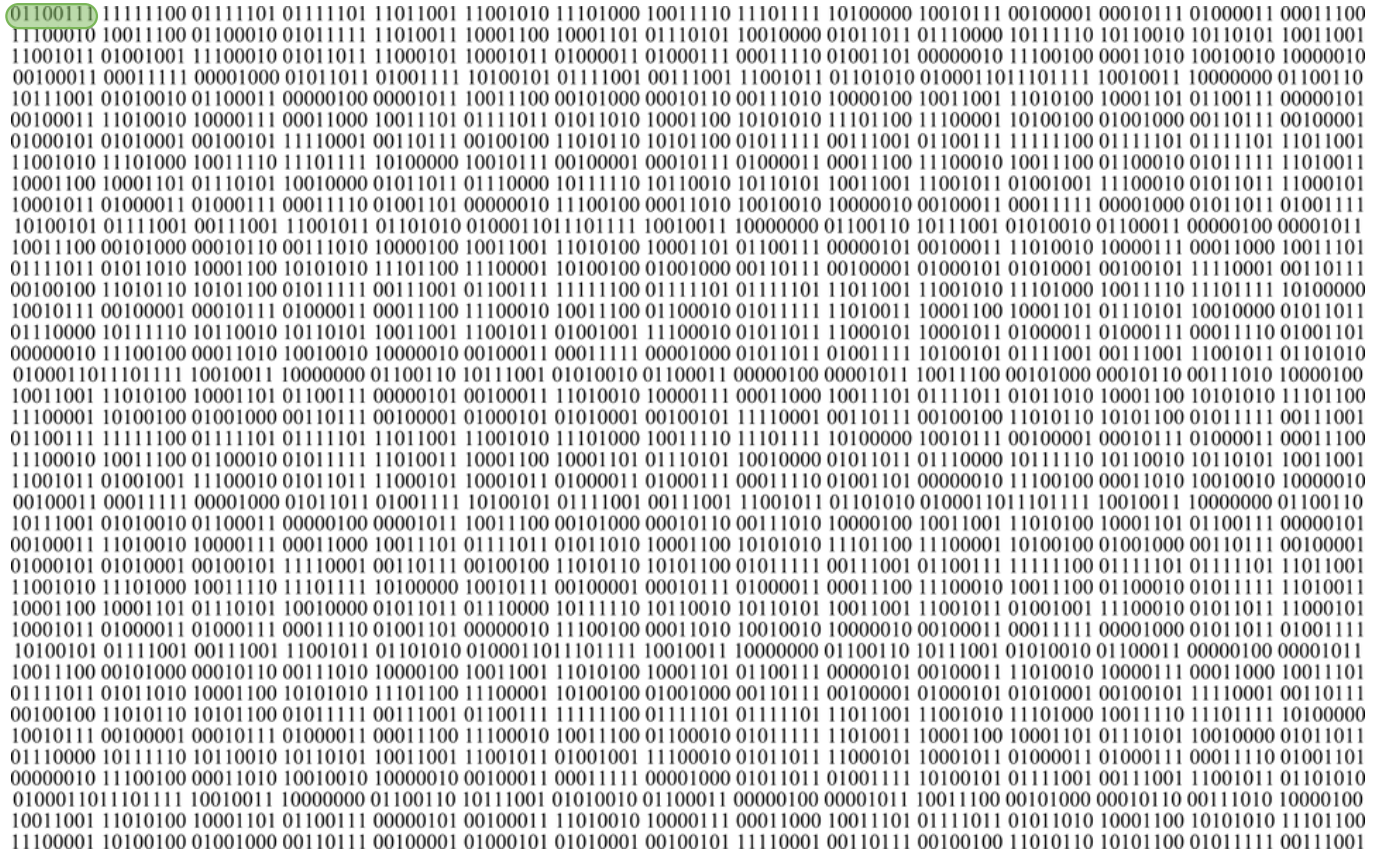
Wörterbuchverfahren

Häufig auftretende Muster werden identifiziert und durch einen kurzen Code ersetzt. Bei gleichartigen Mustern sehr effizient.

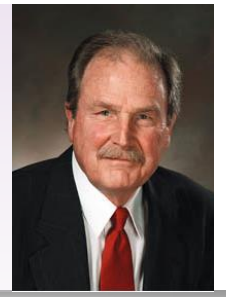
Huffman-Code

Es wird ein für die zu komprimierende Datei optimierter Code erstellt. Dieser Code codiert häufig vorkommende Zeichen möglichst kurz und kommt ohne Trennzeichen aus. Er hat sich stark durchgesetzt.

Welche Bytes kommen wie oft vor?



Huffman-Code



Huffman – Codierung

Mit dem Huffman-Algorithmus wird ein Code generiert, welcher für den vorliegenden Text optimal gestaltet ist:

- Häufige Zeichen → kürzester Code
- Seltene Zeichen → längerer Code
- Aus allen Zeichen-Codes wird ein eindeutiger Bitstrom (keine Trennzeichen), d.h. präfix-frei
- Liefert einen von meist mehreren optimalen Codes

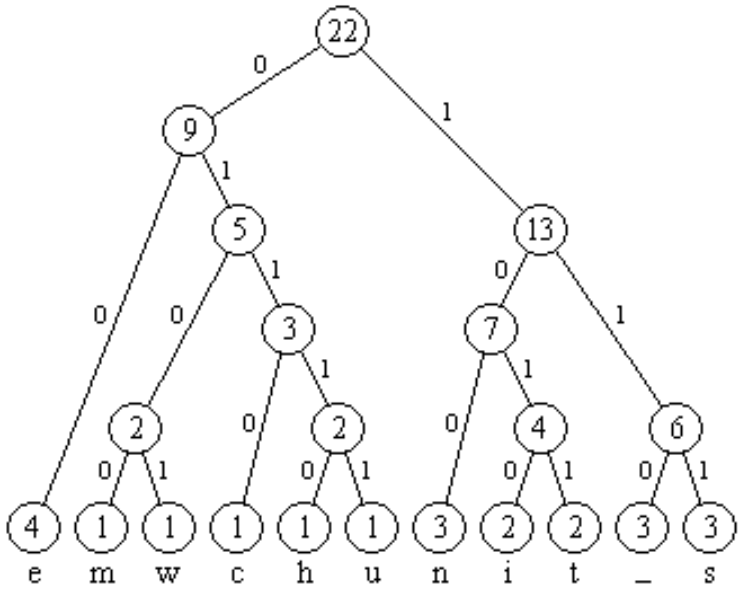
Huffman-Code

Anleitung zur Huffman – Codierung

Zu komprimierender Text: IM WESTEN NICHTS NEUES

1. Schreibe jedes vorkommende Zeichen unten auf ein Blatt.
2. Schreibe über jedes Zeichen seine Häufigkeit.
3. Verbinde immer die beiden tiefsten (freien) Häufigkeitswerte (oder Knoten nach oben zu einem Summenknoten.
4. Wiederhole Schritt 3 bis Du den Stammknoten gebildet hast (Totalsumme)
5. Male unter jeden Knoten links eine null und rechts eine eins.
6. Nun kannst Du den Binärcode jedes Zeichens vom Stammknoten her ablesen.

Huffman-Code



E	00
M	0100
W	0101
C	0110
H	01110
U	01111
N	100
I	1010
T	1011
[ ]	110
S	111

Codierter Bitstrom (73 Bit):

1010010011001010011111011001001101001010011001110101111111010000011110011 1

Kompressionsfaktor und Kompressionsrate

Kompressionsfaktor und -rate

Kompressionsfaktor:
Wie gross ist der Kompressionsfaktor, wenn wir das Resultat unserer Huffman-Codierung mit dem Originaltext in Unicode vergleichen?

Lösung:

$$KF = \frac{73\ Bit}{22 \bullet 16\ Bit} = 0.207 = 20.7\%$$

Merke: Die Kompressionsrate ist der Kehrwert des Kompressionsfaktors!

$$KF = \frac{komprimierte\ Grösse}{ursprüngliche\ Grösse} \qquad KR = \frac{ursprüngliche\ Grösse}{komprimierte\ Grösse}$$

Einsatzgebiete Huffman-Code

- JPEG
- ZIP (Implode)
- MPEG

Und viele weitere…

Übungsaufgaben

- **Das Gelernte können Sie mit Hilfe von AB 114-07 üben**

Ziel: Repetition und Vertiefung des Stoffes
SF: Einzelarbeit/Partnerarbeit
Zeit: 45 Minuten

Abschluss

- **Offene Punkte / Fragen**
- **Feedback**
- **Hausaufgaben**
  - Arbeitsblatt AB114-07 fertig lösen