

Praktikum BKSP:

Blatt 2

Dr. David Sabel

SoSe 2012

Zeichenbasierte Komprimierung mit Codebäumen

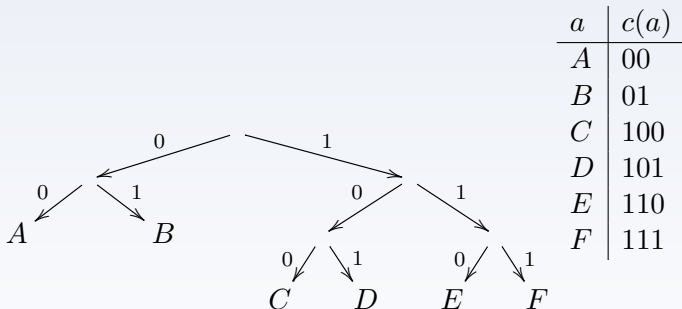
Idee:

- Kodiere jedes Zeichen durch eine Bitfolge
- Repräsentiere die Kodierung durch einen Codebaum

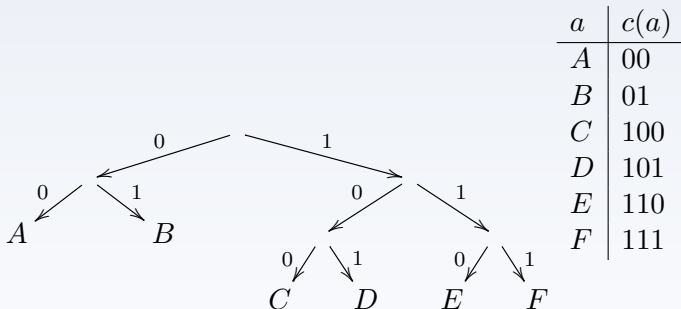
Codebäume:

- \mathcal{A} Alphabet
- Binärer Baum
- Pro $a \in \mathcal{A}$ gibt es ein mit a markiertes Blatt
- innerer Knoten hat zwei Kinder:
 - linke Kante markiert mit 0
 - rechte Kante markiert mit 1
- Codewort $c(a)$ für $a \in \mathcal{A}$:
Beschriftung des Pfades von der Wurzel zu a .

Beispiel



Beispiel



ABCDEFFA ergibt 00011001011101111100.

Aufgabe

- Modul `Codebaum`
- Datentyp `Codebaum a` (polymorph über Alphabeten)
- `mkGetCode :: Codebaum a -> a -> [Int]`, erhält Codebaum und erstellt Funktion, die $c(a)$ berechnet. `[Int]` soll Liste aus 0en und 1en darstellen
- `mkGetCode` möglichst effizient (Baum jedes mal durchsuchen ist ineffizient)
- `kodiere :: (Codebaum a) -> [a] -> [Int]` erhält Codebaum und Wort, berechnet Code

Huffman-Kodierung

- Entropiekodierung
- Zeichenbasiert

Idee

- Verwende kurze Bitfolgen für häufige Zeichen, längere für seltene Zeichen

Huffman-Kodierung

Sei \mathcal{T} ein Text über dem Alphabet \mathcal{A}

- 1 Berechne relative Häufigkeiten $r(a, \mathcal{T})$ für alle $a \in \mathcal{A}$

$$r(a, \mathcal{T}) = \frac{\text{Anzahl der Vorkommen von } a \text{ in } \mathcal{T}}{\text{Anzahl der Zeichen von } \mathcal{T}}$$

Huffman-Kodierung

Sei \mathcal{T} ein Text über dem Alphabet \mathcal{A}

- 1 Berechne relative Häufigkeiten $r(a, \mathcal{T})$ für alle $a \in \mathcal{A}$

$$r(a, \mathcal{T}) = \frac{\text{Anzahl der Vorkommen von } a \text{ in } \mathcal{T}}{\text{Anzahl der Zeichen von } \mathcal{T}}$$

- 2 Erstelle optimalen Codebaum:

Huffman-Kodierung

Sei \mathcal{T} ein Text über dem Alphabet \mathcal{A}

- 1 Berechne relative Häufigkeiten $r(a, \mathcal{T})$ für alle $a \in \mathcal{A}$

$$r(a, \mathcal{T}) = \frac{\text{Anzahl der Vorkommen von } a \text{ in } \mathcal{T}}{\text{Anzahl der Zeichen von } \mathcal{T}}$$

- 2 Erstelle optimalen Codebaum:

- 1 Starte mit Wald von Codebäumen:

Für jedes $a \in \mathcal{A}$ eine Wurzel

Bäume sind mit relativen Häufigkeiten markiert

Huffman-Kodierung

Sei \mathcal{T} ein Text über dem Alphabet \mathcal{A}

- 1 Berechne relative Häufigkeiten $r(a, \mathcal{T})$ für alle $a \in \mathcal{A}$

$$r(a, \mathcal{T}) = \frac{\text{Anzahl der Vorkommen von } a \text{ in } \mathcal{T}}{\text{Anzahl der Zeichen von } \mathcal{T}}$$

- 2 Erstelle optimalen Codebaum:

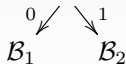
- 1 Starte mit Wald von Codebäumen:

Für jedes $a \in \mathcal{A}$ eine Wurzel

Bäume sind mit relativen Häufigkeiten markiert

- 2 Iteriere den Schritt bis ein Baum entstanden ist:

Wähle Bäume $\mathcal{B}_1, \mathcal{B}_2$ mit kleinsten rel. Häufigkeiten und verschmelze diese:



Neue Häufigkeit: Summe $r(\mathcal{B}_1) + r(\mathcal{B}_2)$

Beispiel

$$\mathcal{T} = \text{AECACEFAABDDBEAEAECEEEEDF.}$$

a	Anzahl a im Text \mathcal{T}	$r(a, \mathcal{T})$
A	6	$6/25 = 0,24 = 24\%$
B	2	$2/25 = 0,08 = 8\%$
C	3	$3/25 = 0,12 = 12\%$
D	3	$3/25 = 0,12 = 12\%$
E	9	$9/25 = 0,36 = 36\%$
F	2	$2/25 = 0,08 = 8\%$

Wald am Anfang:

$$W_0 = \{(A, 24\%), (B, 8\%), (C, 12\%), (D, 12\%), (E, 36\%), (F, 8\%)\}$$

Beispiel (2)

$$W_0 = \{(A, 24\%), (B, 8\%), (C, 12\%), (D, 12\%), (E, 36\%), (F, 8\%)\}$$

$$W_1 = \{(A, 24\%), (\begin{array}{c} 0 \swarrow \\ B \end{array} \begin{array}{c} \searrow 1 \\ F \end{array}, 16\%), (C, 12\%), (D, 12\%), (E, 36\%)\}$$

Beispiel (2)

$$W_0 = \{(A, 24\%), (B, 8\%), (C, 12\%), (D, 12\%), (E, 36\%), (F, 8\%)\}$$

$$W_1 = \{(A, 24\%), (\begin{array}{c} 0 \swarrow \searrow 1 \\ B \quad F \end{array}, 16\%), (C, 12\%), (D, 12\%), (E, 36\%)\}$$

$$W_2 = \{(A, 24\%), (\begin{array}{c} 0 \swarrow \searrow 1 \\ B \quad F \end{array}, 16\%), (\begin{array}{c} 0 \swarrow \searrow 1 \\ C \quad D \end{array}, 24\%), (E, 36\%)\}$$

Beispiel (2)

$$W_0 = \{(A, 24\%), (B, 8\%), (C, 12\%), (D, 12\%), (E, 36\%), (F, 8\%)\}$$

$$W_1 = \{(A, 24\%), (\begin{array}{c} \swarrow \\ B \\ \searrow \\ \\ F \end{array}, 16\%), (C, 12\%), (D, 12\%), (E, 36\%)\}$$

$$W_2 = \{(A, 24\%), (\begin{array}{c} \swarrow \\ B \\ \searrow \\ \\ F \end{array}, 16\%), (\begin{array}{c} \swarrow \\ C \\ \searrow \\ \\ D \end{array}, 24\%), (E, 36\%)\}$$

$$W_3 = \{(\begin{array}{c} \swarrow \\ A \\ \searrow \\ \\ \swarrow \\ \\ B \\ \searrow \\ \\ F \end{array}, 40\%), (\begin{array}{c} \swarrow \\ C \\ \searrow \\ \\ D \end{array}, 24\%), (E, 36\%)\}$$

Beispiel (2)

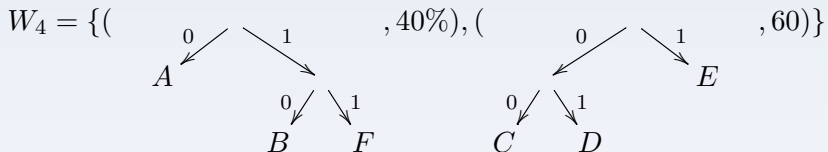
$$W_0 = \{(A, 24\%), (B, 8\%), (C, 12\%), (D, 12\%), (E, 36\%), (F, 8\%)\}$$

$$W_1 = \{(A, 24\%), (\begin{array}{c} \swarrow \\ B \\ \searrow \\ \\ \\ F \end{array}, 16\%), (C, 12\%), (D, 12\%), (E, 36\%)\}$$

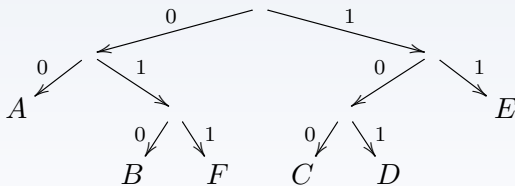
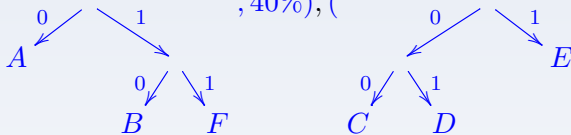
$$W_2 = \{(A, 24\%), (\begin{array}{c} \swarrow \\ B \\ \searrow \\ \\ \\ F \end{array}, 16\%), (\begin{array}{c} \swarrow \\ C \\ \searrow \\ \\ \\ D \end{array}, 24\%), (E, 36\%)\}$$

$$W_3 = \{(\begin{array}{c} \swarrow \\ A \\ \searrow \\ \\ \\ \\ \\ \\ \\ B \\ \\ \\ F \end{array}, 40\%), (\begin{array}{c} \swarrow \\ C \\ \searrow \\ \\ \\ D \end{array}, 24\%), (E, 36\%)\}$$

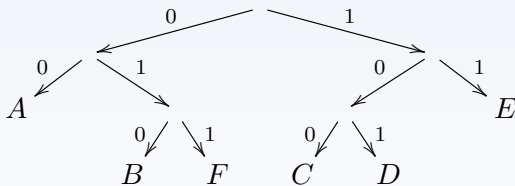
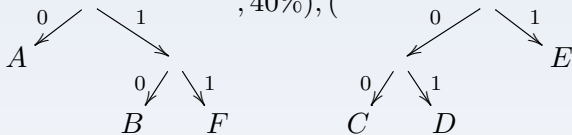
Beispiel (3)



Beispiel (3)

$$W_4 = \{ (\text{Tree 1}, 40\%), (\text{Tree 2}, 60\%) \}$$


Beispiel (3)

$$W_4 = \{ (\text{Tree 1}, 40\%), (\text{Tree 2}, 60\%) \}$$


Kodierung von $\mathcal{T} = \text{AECACEFAABDDBEAEAECEEEEDF}$ ergibt
 0011110000100111011100000101011101010110011001111100111111101011
 Je 8-Bit zu einem ASCII-Zeichen ergibt Wort der Länge 8
 ($|\mathcal{T}| = 25$)

Aufgabe

- Modul `Huffman`
- Funktion `text2Huffman :: String -> (Codebaum Char, [Int])`, die einen Text in Bitfolge mit der Huffmankodierung kodiert
- Konsolen-Programm, das einen Text Huffmankodieren und dekodieren kann und in Datei speichert
- Am Ende muss 0-1-Folge in ASCII-Buchstaben konvertiert werden
- Zum Dekodieren muss der Huffman-Baum sinnvoll in Datei gespeichert werden.

Stringersatzverfahren

Ideen:

- Huffman-Kodierung ist **Zeichenbasiert**
- Stringersatzverfahren versuchen **Wiederholungen von Substrings** zusammenzufassen
- Z.B. in abcabc kommt der Substring abc (zweimal) wiederholt vor.

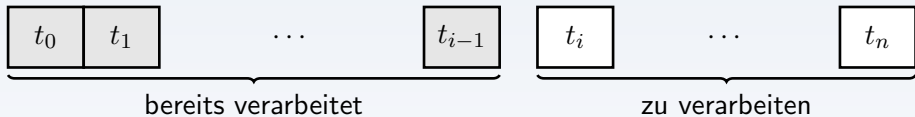
Run Length Encoding

- Einfaches Verfahren
- Ersetze Folgen von gleichen Zeichen durch Zeichen mit Zähler
- `aaaaabaabaaacccccddaaaa` wird zu `5ab2ab3a5c2d4a`
(einzelne Zeichen werden nicht durch Zähler versehen).
- Wenn Ziffern in der Eingabe vorkommen braucht man Trennsymbole, z.B `#5:ab#2:ab#3:a#5:c#2:d#4:a`

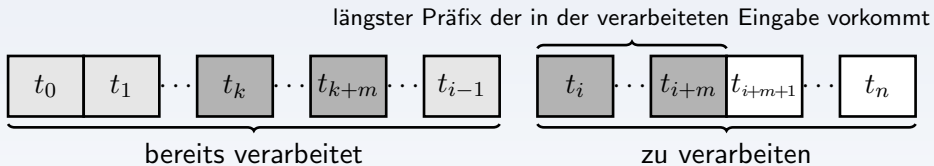
Aufgabe:

- Modul RLE, das RLE durchführt
- Verknüpfung Huffman und RLE.

Lempel-Ziv-Komprimierung



Lempel-Ziv-Komprimierung



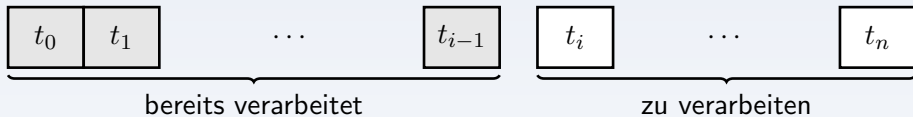
Ausgabe: $(k, m + 1)$

Lempel-Ziv-Komprimierung



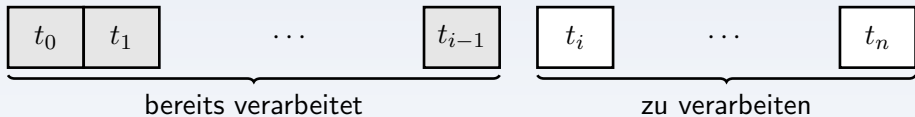
Ausgabe: $(k, m + 1)$

Lempel-Ziv-Komprimierung (2)



Fall: Gar kein echter Präfix kommt in der verarbeiteten Eingabe vor

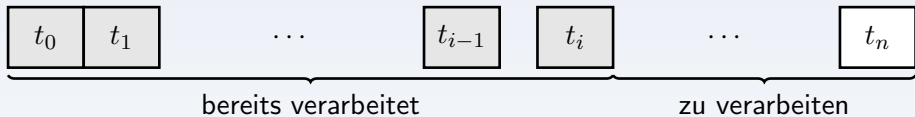
Lempel-Ziv-Komprimierung (2)



Fall: Gar kein echter Präfix kommt in der verarbeiteten Eingabe vor

Ausgabe: t_i

Lempel-Ziv-Komprimierung (2)



Fall: Gar kein echter Präfix kommt in der verarbeiteten Eingabe vor

Ausgabe: t_i

Beispiel

Text aabaaabaabababaaababbbaabbbbccaaa

Verarbeitete Eingabe

Resteingabe

Ausgabe

aabaaabaabababaaababbbaabbbbccaaa

Beispiel

Text aabaaabaabababaaababbbaabbbbccaaa

Verarbeitete Eingabe

a

Resteingabe

aabaaabaabababaaababbbaabbbbccaaa

abaaabaabababaaababbbaabbbbccaaa

Ausgabe

a

Beispiel

Text aabaaabaabababaaababbbaabbbbccaaa

Verarbeitete Eingabe

a
aa

Resteingabe

aabaaabaabababaaababbbaabbbbccaaa
abaaabaabababaaababbbaabbbbccaaa
baaabaabababaaababbbaabbbbccaaa

Ausgabe

a
(0,1)

Beispiel

Text aabaaabaabababaaababbbaabbbbccaaa

Verarbeitete Eingabe

a
aa
aab

Resteingabe

aabaaabaabababaaababbbaabbbbccaaa
abaaabaabababaaababbbaabbbbccaaa
baaabaabababaaababbbaabbbbccaaa
aaabaabababaaababbbaabbbbccaaa

Ausgabe

a
(0,1)
b

Beispiel

Text aabaaabaabababaaababbbaabbbbccaaa

Verarbeitete Eingabe

a
aa
aab
aabaa

Resteingabe

aabaaabaabababaaababbbaabbbbccaaa
abaaabaabababaaababbbaabbbbccaaa
baaabaabababaaababbbaabbbbccaaa
aabaaabaabababaaababbbaabbbbccaaa
abaabababaaababbbaabbbbccaaa

Ausgabe

a
(0,1)
b
(0,2)

Beispiel

Text aabaaabaabababaaababbbaabbbbccaaa

Verarbeitete Eingabe

a
aa
aab
aabaa
aabaaabaa

Resteingabe

aabaaabaabababaaababbbaabbbbccaaa
abaaabaabababaaababbbaabbbbccaaa
baaabaabababaaababbbaabbbbccaaa
aaabaabababaaababbbaabbbbccaaa
aabaaabababaaababbbaabbbbccaaa
babaaababbbaabbbbccaaa

Ausgabe

a
(0,1)
b
(0,2)
(1,4)

Beispiel

Text aabaaabaabababaaababbbaabbbbccaaa

Verarbeitete Eingabe

a
aa
aab
aabaa
aab**a**abaa
aabaaabaaba

Resteingabe

aabaaabaabababaaababbbaabbbbccaaa
abaaabaabababaaababbbaabbbbccaaa
baaabaabababaaababbbaabbbbccaaa
aaabaabababaaababbbaabbbbccaaa
abaabababaaababbbaabbbbccaaa
babaaababbbaabbbbccaaa
baaababbbaabbbbccaaa

Ausgabe

a
(0,1)
b
(0,2)
(1,4)
(2,2)

Beispiel

Text aabaaabaabababaaababbbaabbbbccaaa

Verarbeitete Eingabe

a
aa
aab
aabaa
aabaaabaa
aab**aaaba**aba
aabaaabaababaaaba

Resteingabe

aabaaabaabababaaababbbaabbbbccaaa
abaaabaabababaaababbbaabbbbccaaa
baaabaabababaaababbbaabbbbccaaa
aaabaabababaaababbbaabbbbccaaa
abaabababaaababbbaabbbbccaaa
babaaababbbaabbbbccaaa
baaababbbaabbbbccaaa
bbbaabbbbccaaa

Ausgabe

a
(0,1)
b
(0,2)
(1,4)
(2,2)
(2,6)

Beispiel

Text aabaaabaabababaaababbbaabbbbccaaa

Verarbeitete Eingabe

a
aa
aab
aabaa
aabaaabaa
aabaaabaaba
aab**a**aaabaababaaaba
aabaaabaababaaabab

Resteingabe

aabaaabaabababaaababbbaabbbbccaaa
abaaabaabababaaababbbaabbbbccaaa
baaabaabababaaababbbaabbbbccaaa
aaabaabababaaababbbaabbbbccaaa
abaabababaaababbbaabbbbccaaa
babaaababbbaabbbbccaaa
baaababbbaabbbbccaaa
bbaabbbbccaaa
bbaabbbbccaaa

Ausgabe

a
(0,1)
b
(0,2)
(1,4)
(2,2)
(2,6)
(2,1)

Beispiel

Text aabaaabaabababaaababbbaabbbbccaaa

Verarbeitete Eingabe

a
aa
aab
aabaa
aabaaabaa
aabaaabaaba
aabaaabaababaaaba
aabaaabaababaaabab
aabaaabaababaaababb

Resteingabe

aabaaabaabababaaababbbaabbbbccaaa
abaaabaabababaaababbbaabbbbccaaa
baaabaabababaaababbbaabbbbccaaa
aaabaabababaaababbbaabbbbccaaa
abaabababaaababbbaabbbbccaaa
babaaababbbaabbbbccaaa
baaababbbaabbbbccaaa
bbbaabbbbccaaa
bbaabbbbccaaa
baabbbbccaaa

Ausgabe

a
(0,1)
b
(0,2)
(1,4)
(2,2)
(2,6)
(2,1)
(2,1)

Beispiel

Text aabaaabaabababaaababbbaabbbbccaaa

Verarbeitete Eingabe

a
aa
aab
aabaa
aabaaabaa
aabaaabaaba
aabaaabaababaaaba
aabaaabaababaaabab
aabaaa**baab**abaaababb
aabaaabaababaaababbbaab

Resteingabe

aabaaabaabababaaababbbaabbbbccaaa
abaaabaabababaaababbbaabbbbccaaa
baaabaabababaaababbbaabbbbccaaa
aaabaabababaaababbbaabbbbccaaa
abaabababaaababbbaabbbbccaaa
babaaababbbaabbbbccaaa
baaababbbaabbbbccaaa
bbbaabbbbccaaa
bbaabbbbccaaa
baabbbbccaaa
bbbccaaa

Ausgabe

a
(0,1)
b
(0,2)
(1,4)
(2,2)
(2,6)
(2,1)
(2,1)
(6,4)

Beispiel

Text aabaaabaabababaaababbbaabbbbccaaa

Verarbeitete Eingabe

a
aa
aab
aabaa
aabaaabaa
aabaaabaaba
aabaaabaababaaaba
aabaaabaababaaabab
aabaaabaababaaababb
aabaaabaababaaaba**bbb**aab
aabaaabaababaaababbbaabbbb

Resteingabe

aabaaabaabababaaababbbaabbbbccaaa
abaaabaabababaaababbbaabbbbccaaa
baaabaabababaaababbbaabbbbccaaa
aaabaabababaaababbbaabbbbccaaa
abaabababaaababbbaabbbbccaaa
babaaababbbaabbbbccaaa
baaababbbaabbbbccaaa
bbbaabbbbccaaa
bbaabbbbccaaa
baabbbbccaaa
bbbccaaa
ccaaa

Ausgabe

a
(0,1)
b
(0,2)
(1,4)
(2,2)
(2,6)
(2,1)
(2,1)
(6,4)
(17,3)

Beispiel

Text aabaaabaabababaaababbbaabbbbccaaa

Verarbeitete Eingabe

a
aa
aab
aabaa
aabaaabaa
aabaaabaaba
aabaaabaababaaaba
aabaaabaababaaabab
aabaaabaababaaababb
aabaaabaababaaababbbaab
aabaaabaababaaababbbaabbbb
aabaaabaababaaababbbaabbbbc

Resteingabe

aabaaabaabababaaababbbaabbbbccaaa
abaaabaabababaaababbbaabbbbccaaa
baaabaabababaaababbbaabbbbccaaa
aaabaabababaaababbbaabbbbccaaa
abaabababaaababbbaabbbbccaaa
babaaababbbaabbbbccaaa
baaababbbaabbbbccaaa
bbbaabbbbccaaa
bbaabbbbccaaa
baabbbbccaaa
bbbccaaa
ccaaa
caaa

Ausgabe

a
(0,1)
b
(0,2)
(1,4)
(2,2)
(2,6)
(2,1)
(2,1)
(6,4)
(17,3)
c

Beispiel

Text aabaaabaabababaaababbbaabbbbccaaa

Verarbeitete Eingabe

a
aa
aab
aabaa
aabaaabaa
aabaaabaaba
aabaaabaababaaaba
aabaaabaababaaabab
aabaaabaababaaababb
aabaaabaababaaababbbaab
aabaaabaababaaababbbaabbbb
aabaaabaababaaababbbaabbbb**c**
aabaaabaababaaababbbaabbbbcc

Resteingabe

aabaaabaabababaaababbbaabbbbccaaa
abaaabaabababaaababbbaabbbbccaaa
baaabaabababaaababbbaabbbbccaaa
aaabaabababaaababbbaabbbbccaaa
abaabababaaababbbaabbbbccaaa
babaaababbbaabbbbccaaa
baaababbbaabbbbccaaa
bbbaabbbbccaaa
bbaabbbbccaaa
baabbbbccaaa
bbbccaaa
ccaaa
caaa
aaa

Ausgabe

a
(0,1)
b
(0,2)
(1,4)
(2,2)
(2,6)
(2,1)
(2,1)
(6,4)
(17,3)
c
(26,1)

Beispiel

Text aabaaabaabababaaababbbaabbbbccaaa

Verarbeitete Eingabe

a
aa
aab
aabaa
aabaaabaa
aabaaabaaba
aabaaabaababaaaba
aabaaabaababaaabab
aabaaabaababaaababb
aabaaabaababaaababbbaab
aabaaabaababaaababbbaabbbb
aabaaabaababaaababbbaabbbbc
aab**aaa**baababaaababbbaabbbbcc
aabaaabaababaaababbbaabbbbccaaa

Resteingabe

aabaaabaabababaaababbbaabbbbccaaa
abaaabaabababaaababbbaabbbbccaaa
baaabaabababaaababbbaabbbbccaaa
aaabaabababaaababbbaabbbbccaaa
abaabababaaababbbaabbbbccaaa
babaaababbbaabbbbccaaa
baaababbbaabbbbccaaa
bbbaabbbbccaaa
bbaabbbbccaaa
baabbbbccaaa
bbbccaaa
ccaaa
caaa
aaa

Ausgabe

a
(0,1)
b
(0,2)
(1,4)
(2,2)
(2,6)
(2,1)
(2,1)
(6,4)
(17,3)
c
(26,1)
(3,3)

Beispiel

Text aabaaabaabababaaababbbaabbbbccaaa

Verarbeitete Eingabe

a
aa
aab
aabaa
aabaaabaa
aabaaabaaba
aabaaabaababaaaba
aabaaabaababaaabab
aabaaabaababaaababb
aabaaabaababaaababbbaab
aabaaabaababaaababbbaabbbb
aabaaabaababaaababbbaabbbbc
aabaaabaababaaababbbaabbbbcc
aabaaabaababaaababbbaabbbbccaaa

Resteingabe

aabaaabaabababaaababbbaabbbbccaaa
abaaabaabababaaababbbaabbbbccaaa
baaabaabababaaababbbaabbbbccaaa
aaabaabababaaababbbaabbbbccaaa
abaabababaaababbbaabbbbccaaa
babaaababbbaabbbbccaaa
baaababbbaabbbbccaaa
bbbaabbbbccaaa
bbaabbbbccaaa
baabbbbccaaa
bbbccaaa
ccaaa
caaa
aaa

Ausgabe

a
(0,1)
b
(0,2)
(1,4)
(2,2)
(2,6)
(2,1)
(2,1)
(6,4)
(17,3)
c
(26,1)
(3,3)

Ausgabe:

a(0,1)b(0,2)(1,4)(2,2)(2,6)(2,1)(2,1)(6,4)(17,3)c(26,1)(3,3)

Optimierungen

- Zu kurze gefundene Präfixe durch den String nicht durch die Referenz darstellen.
- Zumindest für einzelne Zeichen
- Statt
a(0,1)b(0,2)(1,4)(2,2)(2,6)(2,1)(2,1)(6,4)(17,3)c(26,1)(3,3)
aab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

Dekomprimierung

Komprimierter Reststring

aab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

Ausgabe

ε

Dekomprimierung

Komprimierter Reststring

aab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

ab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

Ausgabe

ϵ

a

Dekomprimierung

Komprimierter Reststring

aab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

ab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

b(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

Ausgabe

ϵ

a

aa

Dekomprimierung

Komprimierter Reststring

aab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

ab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

b(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

Ausgabe

ϵ

a

aa

aab

Dekomprimierung

Komprimierter Reststring

aab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

ab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

b(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

Ausgabe

ϵ

a

aa

aab

aabaa

Dekomprimierung

Komprimierter Reststring

aab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
ab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
b(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
(2,2)(2,6)bb(6,4)(17,3)cc(3,3)

Ausgabe

ϵ
a
aa
aab
aabaa
aabaaabaa

Dekomprimierung

Komprimierter Reststring

aab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
ab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
b(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
(2,6)bb(6,4)(17,3)cc(3,3)

Ausgabe

ϵ
a
aa
aab
aabaa
aab**aa**baa
aab**aa**a**ba**ba

Dekomprimierung

Komprimierter Reststring

aab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 ab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 b(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (2,6)bb(6,4)(17,3)cc(3,3)
 bb(6,4)(17,3)cc(3,3)

Ausgabe

ε
 a
 aa
 aab
 aabaa
 aabaabaa
 aabaaabaaba
 aabaaabaababaaaba

Dekomprimierung

Komprimierter Reststring

aab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 ab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 b(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (2,6)bb(6,4)(17,3)cc(3,3)
 bb(6,4)(17,3)cc(3,3)
 b(6,4)(17,3)cc(3,3)

Ausgabe

ε
 a
 aa
 aab
 aabaa
 aabaaabaa
 aabaaabaaba
 aabaaabaababaaaba
 aabaaabaababaaabab

Dekomprimierung

Komprimierter Reststring

aab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 ab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 b(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (2,6)bb(6,4)(17,3)cc(3,3)
 bb(6,4)(17,3)cc(3,3)
 b(6,4)(17,3)cc(3,3)
 (6,4)(17,3)cc(3,3)

Ausgabe

ε
 a
 aa
 aab
 aabaa
 aabaaabaa
 aabaaabaaba
 aabaaabaababaaaba
 aabaaabaababaaabab
 aabaaaabababaaababb

Dekomprimierung

Komprimierter Reststring

aab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 ab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 b(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (2,6)bb(6,4)(17,3)cc(3,3)
 bb(6,4)(17,3)cc(3,3)
 b(6,4)(17,3)cc(3,3)
 (6,4)(17,3)cc(3,3)
 (17,3)cc(3,3)

Ausgabe

ϵ
 a
 aa
 aab
 aabaa
 aabaabaa
 aabaaabaaba
 aabaaabaababaaaba
 aabaaabaababaaabab
 aabaaa**ba**ababaaababb
 aabaaabaababaaaba**bbb**aab

Dekomprimierung

Komprimierter Reststring

aab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 ab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 b(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (2,6)bb(6,4)(17,3)cc(3,3)
 bb(6,4)(17,3)cc(3,3)
 b(6,4)(17,3)cc(3,3)
 (6,4)(17,3)cc(3,3)
 (17,3)cc(3,3)
 cc(3,3)

Ausgabe

ε
 a
 aa
 aab
 aabaa
 aabaabaa
 aabaaabaaba
 aabaaabaababaaaba
 aabaaabaababaaabab
 aabaaa**ba**ababaaababb
 aabaaabaababaaaba**bbb**aab
 aabaaabaababaaababbbaabbbb

Dekomprimierung

Komprimierter Reststring

aab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 ab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 b(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (2,6)bb(6,4)(17,3)cc(3,3)
 bb(6,4)(17,3)cc(3,3)
 b(6,4)(17,3)cc(3,3)
 (6,4)(17,3)cc(3,3)
 (17,3)cc(3,3)
 cc(3,3)
 c(3,3)

Ausgabe

ε
 a
 aa
 aab
 aabaa
 aabaaabaa
 aabaaabaaba
 aabaaabaababaaaba
 aabaaabaababaaabab
 aabaaa**ba**ababaaababb
 aabaaabaababaaaba**bbb**aab
 aabaaabaababaaababbbaabbbb
 aabaaabaababaaababbbaabbbbc

Dekomprimierung

Komprimierter Reststring

aab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 ab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 b(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (2,6)bb(6,4)(17,3)cc(3,3)
 bb(6,4)(17,3)cc(3,3)
 b(6,4)(17,3)cc(3,3)
 (6,4)(17,3)cc(3,3)
 (17,3)cc(3,3)
 cc(3,3)
 c(3,3)
 (3,3)

Ausgabe

ε
 a
 aa
 aab
 aabaa
 aabaabaa
 aabaaabaaba
 aabaaabaababaaaba
 aabaaabaababaaabab
 aabaaa**ba**ababaaababb
 aabaaabaababaaaba**bbb**aab
 aabaaabaababaaababbbaabbbb
 aabaaabaababaaababbbaabbbbcc
 aab**aaa**baababaaababbbaabbbbcc

Dekomprimierung

Komprimierter Reststring

aab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 ab(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 b(0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (0,2)(1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (1,4)(2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (2,2)(2,6)bb(6,4)(17,3)cc(3,3)
 (2,6)bb(6,4)(17,3)cc(3,3)
 bb(6,4)(17,3)cc(3,3)
 b(6,4)(17,3)cc(3,3)
 (6,4)(17,3)cc(3,3)
 (17,3)cc(3,3)
 cc(3,3)
 c(3,3)
 (3,3)
 ε

Ausgabe

ε
 a
 aa
 aab
 aabaa
 aabaaabaa
 aabaaabaaba
 aabaaabaababaaaba
 aabaaabaababaaabab
 aabaaa**ba**ababaaababb
 aabaaabaababaaaba**bbb**aab
 aabaaabaababaaababbbbaabbbb
 aabaaabaababaaababbbbaabbbbc
 aab**aaa**baababaaababbbbaabbbbcc
 aabaaabaababaaababbbbaabbbbccaaa

Aufgabe

- Modul LZ
- `lz77Compress`, `lz77Decompress` zur Implementierung der LZ-Kompression / Dekompression
- Konsolenprogramm, das Dateien komprimiert / dekomprimiert
- Verknüpfung Huffman und LZ77
- Optimieren: Wenn Stringreferenz länger als String, dann String speichern

Lempel-Ziv-Welch Algorithmus

- Ähnlich zu LZ77
- Baut ein Wörterbuch bekannter Strings auf
- Benutzt Wörterbuch zum Komprimieren der Resteingabe

Sei

- D das Wörterbuch
- $t_i \dots t_n$ die Resteingabe.
- Am Anfang alle **Zeichen** schon im Wörterbuch

Vorgehen

- Berechne längsten Präfix von t_i, \dots, t_n der in D vorkommt.
- Sei dies t_i, \dots, t_m mit Schlüssel k in D
- Ersetze t_i, \dots, t_m durch k für die Ausgabe
- Füge ins Wörterbuch ein: t_i, \dots, t_{m+1} (falls $m \neq n$)
Beachte: Es genügt k und t_{m+1} zu speichern.
- mache mit t_{m+1}, \dots, t_n weiter

Beispiel

Schlüssel	Wörterbuch			Resteingabe	Ausgabe
	Präfix	Suffix	Wort		
1		a	a	aabaaabaabababaaababbbbaabbbbccaaa	1
2		b	b		
3		c	c		

Beispiel

Wörterbuch					
Schlüssel	Präfix	Suffix	Wort		
1		a	a	Resteingabe aabaaabaabababaaababbbbaabbbbccaaa	Ausgabe 1
2		b	b		
3		c	c		

Beispiel

Schlüssel	Wörterbuch			Resteingabe	Ausgabe
	Präfix	Suffix	Wort		
1		a	a	aabaaabaababaaababbbaabbbbccaaa	1
2		b	b		
3		c	c		
4	1	a	aa	aabaaabaababaaababbbaabbbbccaaa	1

Beispiel

Wörterbuch					
Schlüssel	Präfix	Suffix	Wort		
1		a	a		
2		b	b		
3		c	c		
4	1	a	aa	Resteingabe	Ausgabe
5	1	b	ab	aabaaabaabababaaababbbbaabbbbccaaa	1
				abaaabaabababaaababbbbaabbbbccaaa	1
				baaabaabababaaababbbbaabbbbccaaa	2

Beispiel

Schlüssel	Wörterbuch			Resteingabe	Ausgabe
	Präfix	Suffix	Wort		
1		a	a	aabaaabaababaaaababbbbaabbbbccaaa	1
2		b	b		
3		c	c		
4	1	a	aa	abaaabaababaaaababbbbaabbbbccaaa	1
5	1	b	ab	baaabaababaaaababbbbaabbbbccaaa	2
6	2	a	ba	aaabaababaaaababbbbaabbbbccaaa	4

Beispiel

Schlüssel	Wörterbuch			Resteingabe	Ausgabe
	Präfix	Suffix	Wort		
1		a	a	aabaaabaababaaaababbbbaabbbbccaaa	1
2		b	b		
3		c	c		
4	1	a	aa	abaaabaababaaaababbbbaabbbbccaaa	1
5	1	b	ab	baaabaababaaaababbbbaabbbbccaaa	2
6	2	a	ba	aaabaababaaaababbbbaabbbbccaaa	4
7	4	a	aaa	abaababaaaababbbbaabbbbccaaa	5

Beispiel

Schlüssel	Wörterbuch			Resteingabe	Ausgabe	
	Präfix	Suffix	Wort			
1		a	a	aabaaabaababaaababbbbaabbbbccaaa	1	
2		b	b		abaaabaababaaababbbbaabbbbccaaa	1
3		c	c			baaabaababaaababbbbaabbbbccaaa
4	1	a	aa	aaabaababaaababbbbaabbbbccaaa	4	
5	1	b	ab	abaababaaababbbbaabbbbccaaa	5	
6	2	a	ba	aabababaaababbbbaabbbbccaaa	4	
7	4	a	aaa		aaabaababaaababbbbaabbbbccaaa	5
8	5	a	aba	aabababaaababbbbaabbbbccaaa	4	

Beispiel

Schlüssel	Wörterbuch		Wort
	Präfix	Suffix	
1		a	a
2		b	b
3		c	c
4	1	a	aa
5	1	b	ab
6	2	a	ba
7	4	a	aaa
8	5	a	aba
9	4	b	aab

Resteingabe

aabaaabaababaaaababbbbaabbbbccaaa
 abaaabaababaaaababbbbaabbbbccaaa
 baaabaababaaaababbbbaabbbbccaaa
 aaabaababaaaababbbbaabbbbccaaa
 abaababaaaababbbbaabbbbccaaa
 aababaaaababbbbaabbbbccaaa
 baaabaababbbbaabbbbccaaa

Ausgabe

1
 1
 2
 4
 5
 4
 6

Beispiel

Wörterbuch					
Schlüssel	Präfix	Suffix	Wort		
1		a	a		
2		b	b		
3		c	c		
				Resteingabe	Ausgabe
4	1	a	aa	aabaaabaabababaaababbbbaabbbbccaaa	1
5	1	b	ab	abaaabaabababaaababbbbaabbbbccaaa	1
6	2	a	ba	baaabaabababaaababbbbaabbbbccaaa	2
7	4	a	aaa	aaabaabababaaababbbbaabbbbccaaa	4
8	5	a	aba	abaabababaaababbbbaabbbbccaaa	5
9	4	b	aab	aabababaaababbbbaabbbbccaaa	4
10	6	b	bab	babaaababbbbaabbbbccaaa	6
				baaababbbbaabbbbccaaa	6

Beispiel

Wörterbuch					
Schlüssel	Präfix	Suffix	Wort		
1		a	a		
2		b	b		
3		c	c		
				Resteingabe	Ausgabe
4	1	a	aa	aabaaabaababaaaababbbbaabbbbccaaa	1
5	1	b	ab	abaaabaababaaaababbbbaabbbbccaaa	1
6	2	a	ba	baaabaababaaaababbbbaabbbbccaaa	2
7	4	a	aaa	aaabaababaaaababbbbaabbbbccaaa	4
8	5	a	aba	abaababaaaababbbbaabbbbccaaa	5
9	4	b	aab	aababaaaababbbbaabbbbccaaa	4
10	6	b	bab	babaaaababbbbaabbbbccaaa	6
11	6	a	baa	baaababbbbaabbbbccaaa	6
				aababbbbaabbbbccaaa	9

Beispiel

Wörterbuch					
Schlüssel	Präfix	Suffix	Wort		
1		a	a		
2		b	b		
3		c	c		
4	1	a	aa		
5	1	b	ab		
6	2	a	ba		
7	4	a	aaa		
8	5	a	aba		
9	4	b	aab		
10	6	b	bab		
11	6	a	baa		
12	9	a	aaba		
				Resteingabe	Ausgabe
				aabaaabaababaaaababbbbaabbbbccaaa	1
				abaaabaababaaaababbbbaabbbbccaaa	1
				baaabaababaaaababbbbaabbbbccaaa	2
				aaabaababaaaababbbbaabbbbccaaa	4
				abaababaaaababbbbaabbbbccaaa	5
				aababaaaababbbbaabbbbccaaa	4
				babaaaababbbbaabbbbccaaa	6
				baaababbbbaabbbbccaaa	6
				aababbbbaabbbbccaaa	9
				abbbbaabbbbccaaa	5

Beispiel

Wörterbuch					
Schlüssel	Präfix	Suffix	Wort		
1		a	a		
2		b	b		
3		c	c		
				Resteingabe	Ausgabe
4	1	a	aa	aabaaabaabababaaababbbbaabbbbccaaa	1
5	1	b	ab	abaaabaabababaaababbbbaabbbbccaaa	1
6	2	a	ba	baaabaabababaaababbbbaabbbbccaaa	2
7	4	a	aaa	aaabaabababaaababbbbaabbbbccaaa	4
8	5	a	aba	abaabababaaababbbbaabbbbccaaa	5
9	4	b	aab	aabababaaababbbbaabbbbccaaa	4
10	6	b	bab	babaaababbbbaabbbbccaaa	6
11	6	a	baa	baaababbbbaabbbbccaaa	6
12	9	a	aaba	aababbbbaabbbbccaaa	9
13	5	b	abb	abbbaabbbbccaaa	5
				b baabbbbccaaa	2

Beispiel

Wörterbuch					
Schlüssel	Präfix	Suffix	Wort		
1		a	a		
2		b	b		
3		c	c		
				Resteingabe	Ausgabe
4	1	a	aa	aabaaabaabababaaababbbbaabbbbccaaa	1
5	1	b	ab	abaaabaabababaaababbbbaabbbbccaaa	1
6	2	a	ba	baaabaabababaaababbbbaabbbbccaaa	2
7	4	a	aaa	aaabaabababaaababbbbaabbbbccaaa	4
8	5	a	aba	abaabababaaababbbbaabbbbccaaa	5
9	4	b	aab	aabababaaababbbbaabbbbccaaa	4
10	6	b	bab	babaaaababbbbaabbbbccaaa	6
11	6	a	baa	baaababbbbaabbbbccaaa	6
12	9	a	aaba	aababbbbaabbbbccaaa	9
13	5	b	abb	abbbaabbbbccaaa	5
14	2	b	bb	bbaabbbbccaaa	2
				baabbbbccaaa	11

Beispiel

Schlüssel	Wörterbuch		Wort
	Präfix	Suffix	
1		a	a
2		b	b
3		c	c
4	1	a	aa
5	1	b	ab
6	2	a	ba
7	4	a	aaa
8	5	a	aba
9	4	b	aab
10	6	b	bab
11	6	a	baa
12	9	a	aaba
13	5	b	abb
14	2	b	bb
15	11	b	baab

Resteingabe

aabaaabaababaaababbbbaabbbbccaaa
 abaaabaababaaababbbbaabbbbccaaa
 baaabaababaaababbbbaabbbbccaaa
 aaabaababaaababbbbaabbbbccaaa
 abaababaaababbbbaabbbbccaaa
 aababaaababbbbaabbbbccaaa
 babaaababbbbaabbbbccaaa
 baaababbbbaabbbbccaaa
 aababbbbaabbbbccaaa
 abbbbaabbbbccaaa
 bbaabbbbccaaa
 baabbbbccaaa
 bbbccaaa

Ausgabe

1
 1
 2
 4
 5
 4
 6
 6
 9
 5
 2
 11
 14

Beispiel

Wörterbuch					
Schlüssel	Präfix	Suffix	Wort		
1		a	a		
2		b	b		
3		c	c		
				Resteingabe	Ausgabe
4	1	a	aa	aabaaabaababaaababbbaabbbbccaaa	1
5	1	b	ab	abaaabaababaaababbbaabbbbccaaa	1
6	2	a	ba	baaabaababaaababbbaabbbbccaaa	2
7	4	a	aaa	aaabaababaaababbbaabbbbccaaa	4
8	5	a	aba	abaababaaababbbaabbbbccaaa	5
9	4	b	aab	aababaaababbbaabbbbccaaa	4
10	6	b	bab	babaaababbbaabbbbccaaa	6
11	6	a	baa	baaababbbaabbbbccaaa	6
12	9	a	aaba	aababbbbaabbbbccaaa	9
13	5	b	abb	abbbaabbbbccaaa	5
14	2	b	bb	bbaabbbbccaaa	2
15	11	b	baab	baabbbbccaaa	11
16	14	b	bbb	bbbccaaa	14
				bbccaaa	14

Beispiel

Wörterbuch					
Schlüssel	Präfix	Suffix	Wort		
1		a	a		
2		b	b		
3		c	c	Resteingabe	Ausgabe
4	1	a	aa	aabaaabaababaaababbbbaabbbbccaaa	1
5	1	b	ab	abaaabaababaaababbbbaabbbbccaaa	1
6	2	a	ba	baaabaababaaababbbbaabbbbccaaa	2
7	4	a	aaa	aaabaababaaababbbbaabbbbccaaa	4
8	5	a	aba	abaababaaababbbbaabbbbccaaa	5
9	4	b	aab	aababaaababbbbaabbbbccaaa	4
10	6	b	bab	babaaababbbbaabbbbccaaa	6
11	6	a	baa	baaababbbbaabbbbccaaa	6
12	9	a	aaba	aababbbbaabbbbccaaa	9
13	5	b	abb	abbbaabbbbccaaa	5
14	2	b	bb	bbaabbbbccaaa	2
15	11	b	baab	baabbbbccaaa	11
16	14	b	bbb	bbbbaabbbbccaaa	14
17	14	c	bbc	bbbaabbbbccaaa	14
				ccaaa	3

Beispiel

Wörterbuch					
Schlüssel	Präfix	Suffix	Wort		
1		a	a		
2		b	b		
3		c	c	Resteingabe	Ausgabe
4	1	a	aa	aabaaabaababaaababbbbaabbbbccaaa	1
5	1	b	ab	abaaabaababaaababbbbaabbbbccaaa	1
6	2	a	ba	baaabaababaaababbbbaabbbbccaaa	2
7	4	a	aaa	aaabaababaaababbbbaabbbbccaaa	4
8	5	a	aba	abaababaaababbbbaabbbbccaaa	5
9	4	b	aab	aababaaababbbbaabbbbccaaa	4
10	6	b	bab	babaaababbbbaabbbbccaaa	6
11	6	a	baa	baaababbbbaabbbbccaaa	6
12	9	a	aaba	aababbbbaabbbbccaaa	9
13	5	b	abb	abbbaabbbbccaaa	5
14	2	b	bb	bbaabbbbccaaa	2
15	11	b	baab	baabbbbccaaa	11
16	14	b	bbb	bbbbaabbbbccaaa	14
17	14	c	bbc	bbccaaa	14
18	3	c	cc	ccaaa	3
				caaa	3

Beispiel

Wörterbuch					
Schlüssel	Präfix	Suffix	Wort		Ausgabe
1		a	a		
2		b	b		
3		c	c		
				Resteingabe	
4	1	a	aa	aabaaabaababaaababbbbaabbbbccaaa	1
5	1	b	ab	abaaabaababaaababbbbaabbbbccaaa	1
6	2	a	ba	baaabaababaaababbbbaabbbbccaaa	2
7	4	a	aaa	aaabaababaaababbbbaabbbbccaaa	4
8	5	a	aba	abaababaaababbbbaabbbbccaaa	5
9	4	b	aab	aababaaababbbbaabbbbccaaa	4
10	6	b	bab	babaaababbbbaabbbbccaaa	6
11	6	a	baa	baaababbbbaabbbbccaaa	6
12	9	a	aaba	aababbbbaabbbbccaaa	9
13	5	b	abb	abbbaabbbbccaaa	5
14	2	b	bb	bbaabbbbccaaa	2
15	11	b	baab	baabbbbccaaa	11
16	14	b	bbb	bbbbaabbbbccaaa	14
17	14	c	bbc	bbccaaa	14
18	3	c	cc	ccaaa	3
19	3	a	ca	caaa	3
				aaa	7

Beispiel (2)

Ergebnis

- Code: 1, 1, 2, 4, 5, 6, 6, 9, 5, 2, 11, 14, 14, 3, 3, 7
- Wörterbuch:

$\{1 \mapsto a, 2 \mapsto b, 3 \mapsto c, 4 \mapsto 1a, 5 \mapsto 1b, 6 \mapsto 2a, 7 \mapsto 4a, 8 \mapsto 5a,$
 $9 \mapsto 4b, 10 \mapsto 6b, 11 \mapsto 6a, 12 \mapsto 9a, 13 \mapsto 5b, 14 \mapsto 2b,$
 $15 \mapsto 11b, 16 \mapsto 14b, 17 \mapsto 14c, 18 \mapsto 3c, 19 \mapsto 3c\}$

- Wörterbuch ist wie eine kontextfreie Grammatik
- Dekomprimieren entspricht Herleitung
- Code braucht nicht komplett gespeichert werden, wenn man die Reihenfolge im Wörterbuch beachtet

Aufgabe

- LZW-Algorithmus in Haskell implementieren
- Dabei effiziente Datenstrukturen benutzen
- LZW mit Huffmankodierung verknüpfen
- Konsolenprogramm zur LZW-Kompression von Dateien