

Lauf längencodierung

Bei der Lauf längencodierung (Run Length Encoding, abgekürzt RLE) handelt es sich um ein einfaches Verfahren zur Datenkomprimierung. Dabei werden sich unmittelbar wiederholende Zeichen nur einmal abgespeichert, und die Anzahl an Wiederholungen wird mit abgespeichert.

Die Einsparung ist der Quotient aus der Anzahl an durch die Codierung eingesparten Zeichen und der Anzahl der Zeichen des Klartextes.

Beispiel:

Der Klartext `IIIIIIINFOOOOOOOORRRMAAAAAAATIKKKKKK`

wird codiert zum RLE-Code `I(7)N(1)F(1)O(9)R(3)M(1)A(8)T(1)I(1)K(6)`

wobei jede Zahl und jedes Zeichen in einem Byte gespeichert wird und die Klammern lediglich der besseren Lesbarkeit dienen und nicht mit abgespeichert werden:

`I7N1F1O9R3M1A8T1I1K6`

Die Einsparung beträgt $\frac{38-20}{38} \approx 0,474 = 47,4\%$.

Aufgaben:

- Erstelle auf Papier den RLE-Code des Klartextes `MISSISSIPPI`. 2 Punkte
- Decodiere auf Papier den RLE-Code `P(7)A(1)B(4)C(2)D(1)` und berechne die Einsparung des RLE-Codes. 2 Punkte
- Nimm Stellung zu Peters Vorschlag: „`AAAABBBAAACCAA` enthält 9x das A, 3x das B und 2x das C. Folglich lässt es sich platzsparend als `A(9)B(3)C(2)`kodieren“. 2 Punkte
- Erläutere, wieso ein kein Problem darstellt, wenn der zu komprimierende Text auch Ziffern enthält, z.B. „`PASSWORD123`“ 2 Punkte
- Erstelle eine Computerlösung zum Decodieren eines RLE-Codes (ohne Berücksichtigung der Klammern). Der Einfachheit halber kannst du zudem davon ausgehen, dass ein Zeichen maximal neunmal wiederholt wird. 4 Punkte
- Ergänze deine Computerlösung um die Möglichkeit, für einen gegebenen Klartext den RLE-Code zu ermitteln. 5 Punkte
- Gib einen Text an, der sich durch RLE um 50% komprimieren lässt. Gib einen Text an, dessen Einsparung -75% beträgt. Gib die Bedingung an, die ein Text erfüllen muss, damit die Einsparung positiv ist, sprich der RLE-Code kürzer als der Klartext ist. Gib an, in welchem Zahlenbereich die Einsparung liegt. 6 Punkte
- Ergänze deine Computerlösung dahingehend, dass sowohl beim Codieren als auch beim Decodieren die Einsparung ausgegeben wird. Falls sie negativ ist, soll eine Warnung ausgegeben werden. 3 Punkte
- Gib an, was passiert, wenn sich ein Zeichen mehr als 255-mal wiederholt. Mache einen Vorschlag, wie sich dieses Problem sinnvoll lösen lässt und gehe auf die Vor- und Nachteile dieses Vorschlags ein. 4 Punkte