

## 2017-RU-05 Zerteile den Code

### Body

In einem speziellen Code für Texte wird jeder Buchstabe durch ein Codewort aus den Ziffern 0 bis 9 kodiert. Dabei gilt diese Regel: Kein Codewort darf mit dem Codewort eines anderen Buchstabens beginnen.

Der Buchstabe **X** wird beispielsweise durch 12 kodiert. Nun kann **Y** durch 2 kodiert werden, denn 12 beginnt nicht mit 2 (und 2 nicht mit 12). Jetzt kann **Z** durch 11 kodiert werden; denn weder 12 noch 2 beginnen mit 11 und 11 beginnt weder mit 12 noch mit 2. 21 wäre jedoch nicht als Codewort für **Z** erlaubt, weil es mit 2, also dem Codewort von **Y** beginnt.

Das Wort **MEMORY** wird durch die Ziffernfolge 12112233321 kodiert.

### Question/Challenge - for the brochures

*Teile die Ziffernfolge in die Codewörter der einzelnen Buchstaben!*

### Interactivity instruction - for the online challenge

Bewege den Mauszeiger in die Lücken zwischen den Ziffern. Es erscheint ein  $\langle - \rangle$ . Klicke um die Ziffernfolge dort zu trennen. Klicke auf «Von vorne» um alle Lücken zu entfernen. Wenn du fertig bist, klicke auf «Antwort speichern».

### Answer Options/Interactivity Description

Der Code wird angezeigt, mit ausreichend Abstand zwischen den Ziffern. Geht man mit der Maus in einen solchen Abstand, wird der Mauspointer als «Trennungssymbol» angezeigt. Ein Klick führt dann dazu, dass der Abstand zwischen den Ziffern vergrößert wird. Ein Klick auf einen vergrößerten Abstand macht daraus wieder einen normalen Abstand. Es gibt einen Knopf um wieder von vorne anzufangen.

### Answer Explanation

So ist es richtig: 1 21 1 22 33 321

Man beginnt links am Anfang der Ziffernfolge. Falls **M** durch 12 kodiert würde, hätte **E** zwangsläufig das Codewort 1, denn dahinter käme wieder 12 für das zweite **M**. Das würde jedoch der Regel widersprechen: Das Codewort für **M** würde dann mit 1 beginnen, dem Codewort für **E**. Längere Anfangsstücke der Ziffernfolge (121, 1211, 12112 etc.) können auch nicht das Codewort für **M** sein, weil dieses Codewort zwei Mal vorkommen muss, diese Stücke aber jeweils nur einmal in der Ziffernfolge enthalten sind. Folglich ist das Codewort für **M** die Ziffer 1.

Nun muss das Codewort für **E** folgen - und dahinter wieder das für **M**, also die 1. Somit kann das Codewort für **E** nur eine der folgenden Ziffernfolgen sein: 2, 21 oder 211223332. Es kann nicht 2 sein;

dann würde das Wort mit MEMM beginnen. Es kann nicht 211223332 sein; dann wäre das Wort insgesamt nur MEM. Folglich ist das Codewort für E die Ziffernfolge 21. Nun ist klar, dass 1 21 1 die Kodierung für MEM ist.

Il resto della sequenza di cifre, cioè 2233321, codifica le lettere ORY. Il 2 da solo non può essere la parola in codice per O, altrimenti avremmo 00 all'inizio. La parola in codice per O deve quindi contenere almeno il 22. Alla fine, 1 e 21 sono già assegnati come parole in codice rispettivamente per M e E. La parola in codice per Y deve quindi essere almeno la sequenza 321. Tra 22 e 321 c'è 33, che deve essere la parola in codice per R: L'unica altra parola in codice possibile per R sarebbe 3. Quindi 3321 dovrebbe essere la parola in codice per Y - e inizierebbe con la parola in codice per R; questo è contrario alla regola. La divisione della parte posteriore è quindi 22 33 321.

## It's Informatics

Der Code, der in dieser Biberaufgabe beschrieben wird, ist ein Beispiel für einen *Prefixcode*. Ein Präfix ist eine Zeichenfolge zu Beginn einer anderen Zeichenfolge. Bei einem Prefixcode darf kein Codewort Präfix eines anderen Codeworts sein. Das heisst: kein Codewort darf mit einem anderen Codewort beginnen.

Bei Prefixcodes haben die Codewörter unterschiedliche Länge. Der Vorteil der Präfix-Regel ist, dass man keine Trennsymbole zwischen Codewörtern benötigt. Man kann immer erkennen, an welcher Stelle das nächste Codewort beginnt. Wenn man kurze Codewörter für häufig vorkommende Buchstaben wählt, kann man Texte sehr effizient kodieren und grosse Textmengen platzsparend speichern.

Die Huffman-Kodierung ist eine Methode, einen optimalen Prefixcode zu finden. Sie ist weit verbreitet und steckt z.B. hinter bekannten Datenformaten wie JPEG und MP3.

## This is Computational Thinking

Optional - not to be filled 2023

## Keywords and Websites

- Prefixcode: <https://de.wikipedia.org/wiki/Prefixcode>
- Huffman-Kodierung: <https://de.wikipedia.org/wiki/Huffman-Kodierung>
- Kryptographie: <https://de.wikipedia.org/wiki/Kryptographie>
- Kryptoanalyse: <https://de.wikipedia.org/wiki/Kryptoanalyse>