

Kompression mit Huffman-Codierung

Bei der herkömmlichen Text-Codierung mithilfe der ASCII-Tabelle besitzt jedes Zeichen eine Codelänge von genau 8 Bit. Die Idee hinter der Huffman-Codierung ist die, dass häufig vorkommende Zeichen mit einer kürzeren Codelänge auskommen als weniger häufig vorkommende Zeichen.

Ein Beispiel kennst du bereits: **Morsecode** verwendet als Code für das häufig vorkommende E lediglich ein • wohingegen ein Q in den den deutlich längeren Morsecode --•– übersetzt wird. Beim Morsen muss man jedoch neben den beiden Symbolen „kurz“ und „lang“ jedoch auch noch mit Pausen arbeiten um die unterschiedlichen Zeichen unterscheiden zu können. Das ist nötig, da der Morsecode nicht präfixfrei ist, das bedeutet, dass es Codes von Zeichen gibt, die gleichzeitig auch als Anfang eines anderen Codes interpretiert werden können. Beispiel: Das „E“ mit seinem Code „•“ ist gleichzeitig auch der Anfang des Zeichens „A“ mit dem Code „•-“. Damit ist der Morsecode ohne zusätzliche Pause (= drittes Zeichen) nicht mehr eindeutig entzifferbar:

- - - • • • - • - • = ABC
- - - • • • - • - • = EGFN

Der Huffman-Code bildet solch einen präfixfreien Code, der gleichzeitig den häufig vorkommenden Zeichen eine kürzere Codelänge generiert.

From:

<https://wiki.qg-moessingen.de/> - QG Wiki

Permanent link:

<https://wiki.qg-moessingen.de/faecher:informatik:oberstufe:codierung:huffmancodierung:start?rev=1634226990>

Last update: **14.10.2021 17:56**

