

Vorlesung

Kommunikationstechnik

Laborübung

Wörterbuch-Codierung

Prof. Dr. Dirk Staehle

Die Abgabe erfolgt durch Hochladen der bearbeiteten Word-Datei in Moodle.

Bearbeitung in Zweier-Teams

Team-Mitglied 1:

Team-Mitglied 2:

1 Aufgabenstellung

In dieser Laboraufgabe wollen wir die beiden Verfahren Lempel-Ziv (LZ77) und Lempel-Ziv-Welch (LZW) implementieren und vergleichen. Finden Sie sich daher wieder wie bei der Entropie-Codierung zu Team-Paaren zusammen, von denen ein Team die Codierung und Decodierung nach Lempel-Ziv und nach Lempel-Ziv-Welch implementiert.

2 Lempel-Ziv

Implementieren Sie eine Klasse `LempelZiv`, die die Codierung nach dem Lempel-Ziv-Verfahren durchführt. Die Klasse erhält als Attribute

- die Anzahl Bits für die Codierung der Rückwärtsreferenz
- die Anzahl Bits für die Länge der Zeichenkette

Die Klasse soll über die folgenden Methoden verfügen:

- `bitstring, charD=encode(message)`
- `message=decode(bitstring,charD)`

Das Dictionary `charD` enthält die verwendete Codierung der Buchstaben zu Bits. Codieren Sie alle Buchstaben mit der gleichen Anzahl von Bits, die Sie anhand der in der Nachricht vorkommenden Buchstaben festlegen können.

Tipps:

- `set(message)` liefert Ihnen die Buchstaben
- `f'{i:0{n}b}'` stellt die Zahl `i` als Bitstring mit `n` Stellen dar
- `int(bin_str,2)` wandelt einen binären String in eine Dezimalzahl um

3 Lempel-Ziv-Welch

Implementieren Sie eine Klasse `LempelZivWelch`, die die Codierung nach dem Lempel-Ziv-Welch-Verfahren durchführt.

Die Klasse soll über die folgenden Methoden verfügen:

- `bitstring=encode(message,charL)`
- `message=decode(bitstring, charL)`

Die Liste `charL` enthält die Zeichen, mit denen das Wörterbuch initialisiert werden soll. Verwenden Sie in der Aufgabe nur die in der Nachricht vorkommenden Zeichen.

4 Test

Überprüfen Sie ihre Implementierung anhand der Beispiele aus der Vorlesung.

5 Performance

5.1 Einfluss der Codierung bei Lempel-Ziv

Bestimmen Sie für unsere Test-Nachricht, den RFC 2324, die Kombination aus Anzahl Bits für Rückwärtsreferenz und Zeichenkettenlänge, die das optimale Ergebnis liefert.

5.2 Vergleich

Wenden Sie die beiden Verfahren auf unsere Test-Nachricht an, den RFC 2324. Welche ist besser?

Tipp: Hier kommt der „Spezialfall“ bei der Decodierung vor.