

Übungsblatt 6

Grundbegriffe der Informatik — Winter 2023/24

Tutor*in: Tutorium Nr.:

Nach-,Vorname 1: ,

Matr.nr. 1:

Nach-,Vorname 2: ,

Matr.nr. 2:

Ausgabe: 5. Dezember 2023, 14:30 Uhr

Abgabe: 15. Dezember 2023, 12:30 Uhr

Bitte beachten Sie die Hinweise auf der letzten Seite.

*Von Tutor*in auszufüllen:*

Blatt 6: / 21

Blätter 1 – 6, Stud. 1: / 122

Blätter 1 – 6, Stud. 2: / 122

Aufgabe 1 - Entschlüsselung (6 Punkte)

Gegeben seien die Alphabete $A = \{a, b, c, d\}$ und $B = \{0, 1\}$ sowie die Funktionen $f: A \rightarrow B^*$, $g: A \rightarrow B^*$ und $h: A \rightarrow B^*$, die gemäß der folgenden Tabelle definiert sind. Den Funktionswert $h(d)$ sollen Sie in Teilaufgabe (e) angeben.

x	a	b	c	d
$f(x)$	00	10	01	11
$g(x)$	00	10	01	011
$h(x)$	00	10	01	???

Dazu sei $f^{**}: A^* \rightarrow B^*$ der von f und $g^{**}: A^* \rightarrow B^*$ der von g induzierte Homomorphismus.

- Geben Sie $f^{**}(\text{cbd})$, $g^{**}(\text{cba})$ und $g^{**}(\text{da})$ an. (1 Punkt)
- Sind f^{**} und g^{**} ε -frei? Sind sie präfixfrei? Begründen Sie jeweils kurz Ihre Antwort. (1 Punkt)
- Geben Sie eine Decodierung $l_f: B^* \rightarrow (A \cup \{\perp\})^*$ von f^{**} an. (1 Punkt)
- Geben Sie eine Decodierung $l_g: B^* \rightarrow (A \cup \{\perp\})^*$ von g^{**} an. (2 Punkte)
- Ergänzen Sie die Definition von $h: A \rightarrow B^*$ in der Tabelle oben, so dass der von h induzierte Homomorphismus h^{**} keine Decodierung besitzt. Begründen Sie, warum es keine Decodierung gibt. (1 Punkt)

Aufgabe 2 - Blockschrift (5 Punkte)

In Kapitel 8 der Vorlesung wurde erwähnt, dass es eine bijektive Übersetzung zwischen Wörtern der Länge 4 in binärer Zahlenrepräsentation und Wörtern der Länge 1 in hexadezimaler Repräsentation gibt. In dieser Aufgabe wollen wir weitere solche Übersetzungen finden.

Sei dazu $k \in \mathbb{N}_+$ mit $k \geq 2$. Ein Wort der Länge $n \in \mathbb{N}_+$ über dem Alphabet Z_k wird in dieser Aufgabe als k -ärer n -Block bezeichnet.

- In dieser Teilaufgabe sei zunächst $k = 16, l = 8$. Geben Sie die folgenden Elemente von Z_{16}^3, Z_8^4 an: (1 Punkt)
 - $x_1 \in Z_{16}^3$ mit $\text{Num}_{16}(x_1) = 10$
 - $x_2 \in Z_8^3$ mit $\text{Num}_8(x_2) = 10$
 - $x_3 \in Z_{16}^4$ mit $\text{Num}_{16}(x_3) = 4087$

- $x_4 \in Z_8^4$ mit $\text{Num}_8(x_4) = 4087$
- b) Gibt es eine Funktion $s : Z_{16}^3 \rightarrow Z_8^4$, die (1) bijektiv und (2) eine Übersetzung ist? Begründen Sie Ihre Antwort. (1 Punkt)
- c) Seien $k, l, m, n \in \mathbb{N}_+$ mit $k \geq 2$ und $l \geq 2$. Geben Sie eine Bedingung an, die notwendig und hinreichend dafür ist, dass eine bijektive Übersetzung $t : Z_k^n \rightarrow Z_l^m$ existiert. (1 Punkt)
- d) Zeigen Sie, dass Ihre Bedingung aus der vorherigen Teilaufgabe hinreichend ist. (2 Punkte)

Aufgabe 3 - Auf heißer Spur (5 Punkte)

Detektivin Lovelace ist noch immer davon überzeugt, dass Agent Hoare kein Spion war und ist entschlossen, ihn von den üblen Anschuldigungen gegen ihn reinzuwaschen. Dass Agent Zuse nur schlampig gegen ihn ermittelt zu haben scheint, bestärkt sie nur in diesem Vorhaben.

Verborgен in einer geheimen Schublade seines Schreibtisches findet Lovelace einen an sie adressierten Brief. Hoare scheint ihr eine letzte Nachricht hinterlassen zu haben! Doch ihre Enttäuschung ist groß, als der Umschlag nur einen Zettel mit wenigen Zeilen enthält, alle davon offenbar Kauderwelsch. Zuoberst steht das Wort

verschluesslemich

- a) Stellen Sie einen Huffman-Baum zu diesem Wort auf und geben Sie die daraus resultierende Huffman-Codierung jedes auftretenden Zeichens an. (2 Punkte)

Die Detektivin ist enttäuscht, dass diese Spur nicht zu einem Ziel zu führen scheint. Auf dem Zettel steht aber auch noch das Folgende:

000111110100100000010110
11101110100001001100
100010110001011

Was wollte der verstorbene Agent ihr wohl damit mitteilen? Helfen Sie der Detektivin, das herauszufinden!

Auf der Innenseite des Briefumschlags findet Lovelace etwas, das wie die unvollständige Definition einer Codierung aussieht:

x	a	e	h	n	r	s	t	u	v	z
N_x	2	3	1	2	2	2			1	
$h_m(x)$			00000	011	100	101	110	111		0001

„Aha! Hoare wollte uns einen Hinweis geben, wie wir seinen Code entschlüsseln können!“, meint sie.

- b) Sei $c \in Z_2^*$ die oben stehende Code-Nachricht. Die partielle Tabelle gibt die absoluten Häufigkeiten N_x der Buchstaben der Ursprungsnachricht $m \in \{a, e, h, n, r, s, t, u, v, z\}^*$ an. Die Funktion h_m ist eine Huffman-Codierung für m .

Rekonstruieren Sie die fehlenden Einträge in der Tabelle. Sie sind unter der Annahme, dass h_m eine Huffman-Codierung ist, eindeutig.

Geben Sie das Urbild m von c unter h_m mit $h_m(m) = c$ an. (3 Punkte)

Zufrieden lehnt sich Lovelace in ihrem Schreibtischstuhl zurück: „Na also, geht doch.“

Aufgabe 4 - Der Weisheit letzter Schluss (5 Punkte)

In der Vorlesung wurden bisher Huffman-Codierungen betrachtet, die auf ein bestimmtes Wort zugeschnitten sind. Doch was passiert, wenn man eine Codierung festlegen möchte, um damit bisher unbekannte Texte zu codieren?

In dieser Aufgabe sei das deutsche Alphabet $A = \{a, b, \dots, z\}$ der Buchstaben ohne Umlaute und ß gegeben. Es gilt also $|A| = 26$.

- a) Durch einen Homomorphismus $h_{fix} : A^* \rightarrow Z_2^*$ soll jeder Buchstabe auf ein Wort derselben Länge $n \in \mathbb{N}_+$ abgebildet werden, so dass $|h_{fix}(x)| = n$ für alle $x \in A$. Welchen Wert muss n mindestens annehmen, damit die Übersetzung h_{fix} ein Code ist? Begründen Sie Ihre Antwort kurz. (1 Punkt)
- b) Informieren Sie sich über die relative Häufigkeiten der Zeichen aus A , zum Beispiel unter [Wikipedia](#)¹. Erstellen Sie darauf basierend eine Huffman-Codierung $h_{huffman} : A^* \rightarrow Z_2^*$ für deutsche Texte. Vervollständigen Sie

¹Dort ist auch ß aufgelistet. Aber die Häufigkeit der Buchstaben ohne ß addiert sich schon auf 100%

x		a	b	c	d	e	f	g	h	i	j	k	l	m
N_x		17	1	13	5	27	4	7	26	20	0	2	5	3
x		n	o	p	q	r	s	t	u	v	w	x	y	z
N_x		21	0	0	0	17	12	23	7	0	8	0	0	2

Tabelle 1: Die Anzahlen der Vorkommen der Zeichen aus A im gegebenen Beispieltext

dazu den Huffman-Baum in Abbildung 1, indem Sie die Teilbäume $*_1$ und $*_2$ angeben, und behandeln Sie die relativen Zeichenhäufigkeiten so wie die Anzahlen der Zeichenvorkommen in einem konkreten Wort. (2 Punkte)

- c) Wie viele Zeichen aus Z_2 hat $h_{\text{huffman}}(\mathfrak{T})$ für einen deutschen Text \mathfrak{T} mit $|\mathfrak{T}| = n$, der der angegebenen Buchstabenverteilung folgt, im Durchschnitt? (1 Punkt)
- d) Betrachten Sie den folgenden Text \mathfrak{F} aus *Faust. Der Tragödie zweiter Teil* von Johann Wolfgang von Goethe:

*Daran erkenn' ich den gelehrten Herrn!
 Was ihr nicht tastet, steht euch meilenfern,
 Was ihr nicht fasst, das fehlt euch ganz und gar,
 Was ihr nicht rechnet, glaubt ihr, sei nicht wahr,
 Was ihr nicht waegt, hat fuer euch kein Gewicht,
 Was ihr nicht muenzt, das, meint ihr, gelte nicht.*

Die Anzahlen der Zeichenvorkommen für diesen Text finden Sie in Tabelle 1.

Geben Sie jeweils die Längen der Buchstaben-Codierung $h_{\text{fix}}(\mathfrak{F})$ aus Teilaufgabe (a) und der Huffman-Codierung $h_{\text{huffman}}(\mathfrak{F})$ aus Teilaufgabe (b) an. (1 Punkt)

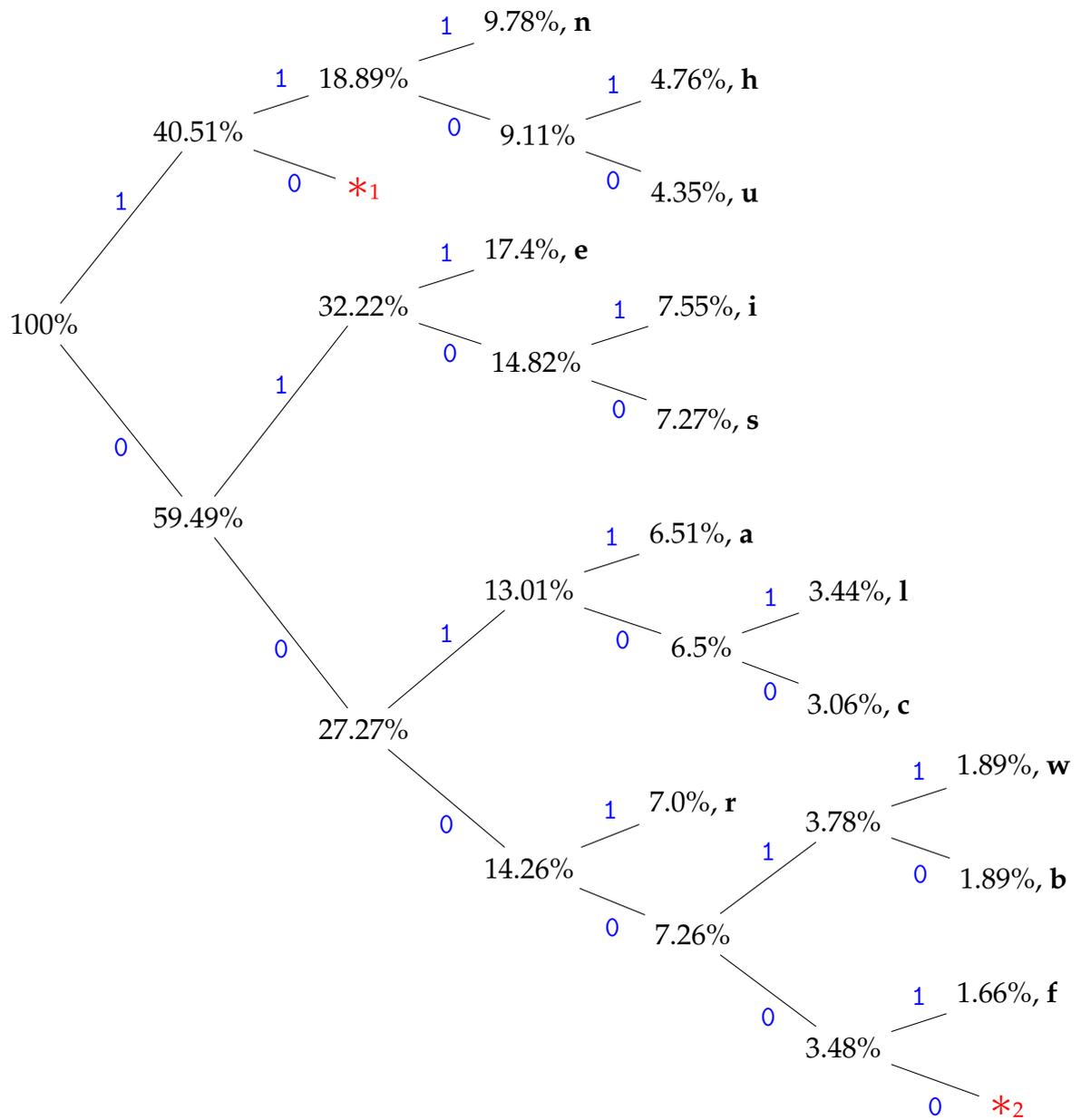


Abbildung 1: Unvollständiger Huffman-Baum zur relativen Zeichenhäufigkeit in deutschen Texten

Bitte beachten Sie die folgenden Hinweise:

- Lösungen **müssen** handschriftlich erstellt werden
- Ihre Abgabe sollte die erste Seite dieser Datei als Deckblatt haben
- Ihre Abgabe muss **rechtzeitig** erfolgen

Außerdem, wenn Sie Ihre Ausarbeitung über die Abgabekästen im Keller des Informatik-Gebäudes abgeben:

- Ihre Abgabe muss in der oberen **linken** Ecke zusammengeheftet werden
- Tablet-Ausdrucke sind zulässig

Wenn Sie Ihre Ausarbeitung online über ILIAS abgeben, dann achten Sie darauf:

- Ihre Abgabe muss **genau eine** PDF-Datei sein
- Scans und lesbare Fotos sind zulässig
- Abgabe erfolgt unter „Tutorien“ im Ordner **Ihres** Tutoriums