Informatik IV - Tutorium XVI & XVII

Tut Nr. 9 – Fehlerkorrigierende Codes

David Münch

Universität Karlsruhe (TH)
Fakultät für Informatik
IBDS Prautzsch

Dank an Yusuf für die Zusammenarbeit.

3. Juli 2009

Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

Inhaltsverzeichnis

Inhaltsverzeichnis

Inhaltsverzeichnis

Inhaltsverzeichnis

Organisatorisches

Email: muenchdavid@gmail.com

https://www.stud.uni-karlsruhe.de/~uhbro/

Tutorium 16: Freitags 8:00 Uhr - Raum -107
Tutorium 17: Freitags 9:45 Uhr - Raum -119

Schein / Übungsblätter

Bearbeitete Hausaufgaben können abgeben werden.
Es gibt vier Scheinklausuren mit je 50 Punkten an folgenden Terminen:

- 19.05.2009 $\surd$ $\varnothing 31$ Punkte
- 09.06.2009 $\surd$ $\varnothing 35$ Punkte
- 02.07.2009 $\surd$ $\varnothing 39$ Punkte
- 23.07.2009 8:00 Uhr $\ddot\smile$

120 Punkte aus den Scheinklausuren sind notwendig, um einen Schein zu erhalten.

Schein / Übungsblätter

Bearbeitete Hausaufgaben können abgeben werden.

Es gibt vier Scheinklausuren mit je 50 Punkten an folgenden Terminen:

- 19.05.2009 $\surd$ $\varnothing 31$ Punkte
- 09.06.2009 $\surd$ $\varnothing 35$ Punkte
- 02.07.2009 $\surd$ $\varnothing 39$ Punkte
- 23.07.2009 8:00 Uhr $\ddot\smile$

120 Punkte aus den Scheinklausuren sind notwendig, um einen Schein zu erhalten.

Ab 140 Punkten gibt es einen Notenbonus von ca. 1/3 Note.

Schein / Übungsblätter

Bearbeitete Hausaufgaben können abgeben werden.
Es gibt vier Scheinklausuren mit je 50 Punkten an folgenden Terminen:

- 19.05.2009 $\surd$ $\varnothing 31$ Punkte
- 09.06.2009 $\surd$ $\varnothing 35$ Punkte
- 02.07.2009 $\surd$ $\varnothing 39$ Punkte
- 23.07.2009 8:00 Uhr $\ddot\smile$

120 Punkte aus den Scheinklausuren sind notwendig, um einen Schein zu erhalten.

Ab 140 Punkten gibt es einen Notenbonus von ca. 1/3 Note.

Abgabe in Zweiergruppe erlaubt und ausdrücklich erwünscht!

Schein / Übungsblätter

Bearbeitete Hausaufgaben können abgeben werden.
Es gibt vier Scheinklausuren mit je 50 Punkten an folgenden Terminen:

- 19.05.2009 $\surd$ $\varnothing 31$ Punkte
- 09.06.2009 $\surd$ $\varnothing 35$ Punkte
- 02.07.2009 $\surd$ $\varnothing 39$ Punkte
- 23.07.2009 8:00 Uhr $\ddot\smile$

120 Punkte aus den Scheinklausuren sind notwendig, um einen Schein zu erhalten.

Ab 140 Punkten gibt es einen Notenbonus von ca. 1/3 Note.

Abgabe in Zweiergruppe erlaubt und ausdrücklich erwünscht!

Bitte folgendes Deckblatt verwenden:
`http://www.stud.uni-karlsruhe.de/~unbdh/deckblatt/index.php?course=8`

Literatur

Boehm, Prautzsch: Numerical Methods. AK Peters 1993. ISBN 3-528-06350-5
http://www.ubka.uni-karlsruhe.de/hylib-bin/suche.cgi?opacdb=UBKA_OPAC&fbt=7319953&nd=3204657

Ash: Information Theory. Dover 1990. ISBN 0-486-66521-6
http://www.ubka.uni-karlsruhe.de/hylib-bin/suche.cgi?opacdb=UBKA_OPAC&nd=9866904

Goos: Vorlesungen über Informatik. Bd. 4, Springer 1998. ISBN 3-540-60650-5
http://www.ubka.uni-karlsruhe.de/hylib-bin/suche.cgi?opacdb=UBKA_OPAC&fbt=9316367&nd=6568301

Was wollen wir heute erreichen?

Was wollen wir heute erreichen?

- 3. Scheinklausur besprechen

Was wollen wir heute erreichen?

- 3. Scheinklausur besprechen
- Hamming-Codes

Aufgabe 1

1. Sei X eine Zufallsvariable über $\mathcal{A} = \{a_1, \dots, a_6\}$ mit den Wahrscheinlichkeiten $p_i = p(a_i)$ wobei

$$p_1 = \frac{1}{3}\ , \qquad p_2 = p_3 = p_4 = \frac{1}{9}\ , \qquad p_5 = p_6 = \frac{1}{6}\ .$$

Konstruieren Sie dazu einen ternären Huffman-Code $l : \mathcal{A} \to \{0, 1, 2\}$.

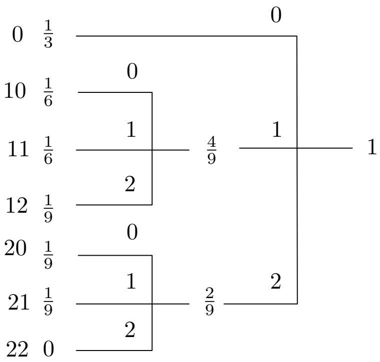
Aufgabe 1

1. Sei X eine Zufallsvariable über $\mathcal{A} = \{a_1, \ldots, a_6\}$ mit den Wahrscheinlichkeiten $p_i = p(a_i)$ wobei

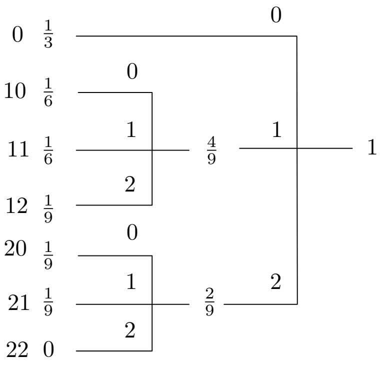
$$p_1 = \frac{1}{3}\ , \qquad p_2 = p_3 = p_4 = \frac{1}{9}\ , \qquad p_5 = p_6 = \frac{1}{6}\ .$$

   Konstruieren Sie dazu einen ternären Huffman-Code $l : \mathcal{A} \to \{0, 1, 2\}$.
2. Ändern Sie zwei der p_i so ab, dass der Huffman-Code redundanzfrei ist. (ohne Beweis)

Lösung


0 $\frac{1}{3}$
10 $\frac{1}{6}$
11 $\frac{1}{6}$
12 $\frac{1}{9}$
20 $\frac{1}{9}$
21 $\frac{1}{9}$
22 0
0
1
2
0
1
2
$\frac{4}{9}$
$\frac{2}{9}$
0
1
2
1

Lösung



Für $p_5 = \frac{1}{3}$ und $p_6 = 0$ ist der Huffman-Code redundanzfrei.

Aufgabe 2

1. Zeigen Sie, dass ein binärer symmetrischer Kanal mit Fehlerwahrscheinlichkeit ρ die Kanalkapazität $\log 2 - H(\rho, 1-\rho)$ hat. Sie können benutzen, dass bei einem binären symmetrischen Kanal und gleichverteilter Eingabe auch die Ausgabe gleichverteilt ist.

Aufgabe 2

1. Zeigen Sie, dass ein binärer symmetrischer Kanal mit Fehlerwahrscheinlichkeit ρ die Kanalkapazität $\log 2 - H(\rho, 1-\rho)$ hat. Sie können benutzen, dass bei einem binären symmetrischen Kanal und gleichverteilter Eingabe auch die Ausgabe gleichverteilt ist.
2. Zeigen Sie für zwei stochastisch unabhängige Zufallsvariablen X und Y, dass

$$H(X,Y) = H(X) + H(Y)$$

gilt.

Lösung

$$
\begin{aligned}
C &= \max_{\mathbf{p}} I(X,Y) \\
&= \max_{\mathbf{p}} \left(H(Y) - H(Y|X)\right) \\
&= \max_{\mathbf{p}} \left(H(Y) - \sum p(x) H(Y|x)\right) \\
&= \max_{\mathbf{p}} \left(H(Y) - \sum p(x) H(\rho, 1-\rho)\right) \\
&= \max_{\mathbf{p}} H(Y) - H(\rho, 1-\rho) \\
&= \log 2 - H(\rho, 1-\rho)
\end{aligned}
$$

Lösung

$$
\begin{aligned}
H(X,Y) &= -\sum_{x,y} p(x,y)\log p(x,y) \\
&= -\sum_{x,y} p(x)p(y)\log(p(x)p(y)) \\
&= -\sum_{x,y} p(x)p(y)\log(p(x)) - \sum_{x,y} p(x)p(y)\log(p(y)) \\
&= -\sum_{x} p(x)\log(p(x)) - \sum_{x} p(y)\log(p(y)) \\
&= H(X) + H(Y)
\end{aligned}
$$

Aufgabe 3

Betrachten Sie den binären Übertragungskanal $X \rightarrow Y$ mit den bedingten Wahrscheinlichkeiten $p(y|x)$:

$$p(0|0) = \frac{2}{3} \qquad p(1|0) = \frac{1}{3}$$

$$p(0|1) = 0 \qquad p(1|1) = 1$$

Eine Tabelle für die numerischen Werte von $-p \cdot \log p$ finden Sie im Anhang.

Wie groß sind $H(X)$, $H(Y)$, $H(X|Y)$, $H(Y|X)$, $I(X,Y)$ und $H(X,Y)$, wenn $H(Y)$ maximal ist?

Lösung

$H(Y)$ ist maximal, wenn $P(Y{=}1) = P(Y{=}0) = \frac{1}{2}$. Aus $P(Y{=}0) = \frac{2}{3} \cdot P(X{=}0)$ folgt für die Quelle $P(X{=}0) = \frac{1}{2} : \frac{2}{3} = \frac{3}{4}$ und $P(X{=}1) = 1 - P(X{=}0) = \frac{1}{4}$.
Daher ist die Quellen-Entropie $H(X) = H(\frac{1}{4}, \frac{3}{4}) = 0{,}500$ bit $+$ $0{,}311$ bit $= 0{,}811$ bit.

Die Verbund-Entropie berechnet sich aus den Wahrscheinlichkeiten $P(X{=}x, Y{=}y) = P(X{=}x) \cdot p(y|x)$ aller kombinierten Fälle:

$$H(X,Y) \;=\; H\left(\frac{1}{2}, 0, \frac{1}{4}, \frac{1}{4}\right) \;=\; 3 \cdot 0{,}500 \text{ bit} \;=\; \frac{3}{2} \text{ bit}$$

Lösung

Sodann lassen sich die bedingten Informationsgehalte ermitteln:

$$\begin{aligned} H(X|Y) &= H(X,Y) - H(Y) \\ &= \frac{3}{2}\text{ bit} - 1\text{ bit} = \frac{1}{2}\text{ bit} \\ H(Y|X) &= H(X,Y) - H(X) \\ &= 1{,}5\text{ bit} - 0{,}811\text{ bit} = 0{,}689\text{ bit} \end{aligned}$$

Nun läßt sich auch die (Trans-)Information des Kanals in diesem Betriebsmodus angeben:

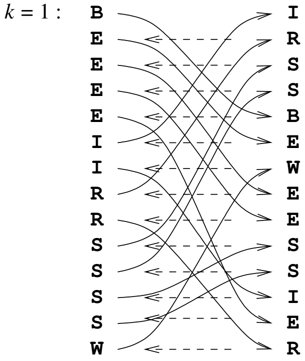
$$\begin{aligned} I(X,Y) &= H(X;Y) \\ &= H(X) - H(X|Y) = H(Y) - H(Y|X) \\ &= 0{,}311\text{ bit} \end{aligned}$$

Aufgabe 4

Unterziehen Sie die Zeichenkette `IRSSBEWEESSIER` mit Startindex $k = 1$ einer inversen Burrows-Wheeler-Transformation und notieren Sie das Ergebnis.

Lösung

Die ursprüngliche Zeichenkette, deren Transformierte gegeben ist, läßt sich wie folgt rekonstruieren:



Das Ergebnis ist also: “BESSERWISSEREI”

Aufgabe 5

Vervollständigen Sie die Dekodierungstabelle des LZW-Verfahrens:

Eingabe	Ausgabe	Wörterbuch
0	0	$2 = \_\_\_$
1	1	
2		
3		
5		
6		
4		
8		
9		

Lösung

Eingabe	Ausgabe	Wörterbuch
0	0	$2 = 01$
1	1	$3 = 10$
2	01	$4 = 011$
3	10	$5 = 101$
5	101	$6 = 1011$
6	1011	$7 = 10110$
4	011	$8 = 0110$
8	0110	$9 = 01100$
9	01100	

Das Ergebnis ist also: 010110101101101011001100

Hamming-Codes

Ein Hamming-Code ist ein linearer Code, der eine Kontrollmatrix A besitzt, die wie folgt aufgebaut ist:

- A hat n Spalten und m Zeilen, wobei gilt: $n \leq 2^m$.

Hamming-Codes

Ein Hamming-Code ist ein linearer Code, der eine Kontrollmatrix A besitzt, die wie folgt aufgebaut ist:

- A hat n Spalten und m Zeilen, wobei gilt: $n \leq 2^m$.
- Die Zeilen von A sind linear unabhängig.

Hamming-Codes

Ein Hamming-Code ist ein linearer Code, der eine Kontrollmatrix A besitzt, die wie folgt aufgebaut ist:

- A hat n Spalten und m Zeilen, wobei gilt: $n \leq 2^m$.
- Die Zeilen von A sind linear unabhängig.
- Die Spalten von A sind verschieden.

Hamming-Codes

Ein Hamming-Code ist ein linearer Code, der eine Kontrollmatrix A besitzt, die wie folgt aufgebaut ist:

- A hat n Spalten und m Zeilen, wobei gilt: $n \leq 2^m$.
- Die Zeilen von A sind linear unabhängig.
- Die Spalten von A sind verschieden.

Ist $A \cdot y = \vec{0}$ dann ist kein Fehler beim Übertragen von y aufgetreten.
Ist dagegen $A \cdot y = \vec{s} \neq \vec{0}$, dann ist die Position der Spalte von $\vec{s}$ in A das gekippte Bit.

Aufgabe

Betrachte den Hamming-Code, der die Gleichung $A \cdot x = 0$ löst, wenn gilt:

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & a_1 \\ 1 & 1 & 0 & 0 & a_2 \\ 1 & 0 & 1 & 0 & a_3 \\ 0 & 1 & 1 & 1 & a_4 \end{pmatrix}$$

a) Gib alle Codewörter an für den Fall $a_1 = a_2 = a_3 = a_4 = 1$.

Aufgabe

Betrachte den Hamming-Code, der die Gleichung $A \cdot x = 0$ löst, wenn gilt:

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & a_1 \\ 1 & 1 & 0 & 0 & a_2 \\ 1 & 0 & 1 & 0 & a_3 \\ 0 & 1 & 1 & 1 & a_4 \end{pmatrix}$$

a) Gib alle Codewörter an für den Fall $a_1 = a_2 = a_3 = a_4 = 1$.

b) Gib alle $\vec{a} = (a_1 ... a_4)^t$ an, sodass alle 1-Fehler korrigiert und alle 2-Fehler entdeckt werden können.

Aufgabe

Betrachte den Hamming-Code, der die Gleichung $A \cdot x = 0$ löst, wenn gilt:

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & a_1 \\ 1 & 1 & 0 & 0 & a_2 \\ 1 & 0 & 1 & 0 & a_3 \\ 0 & 1 & 1 & 1 & a_4 \end{pmatrix}$$

a) Gib alle Codewörter an für den Fall $a_1 = a_2 = a_3 = a_4 = 1$.

b) Gib alle $\vec{a} = (a_1 ... a_4)^t$ an, sodass alle 1-Fehler korrigiert und alle 2-Fehler entdeckt werden können.

c) Für welches $\vec{a} = (a_1 ... a_4)^t$ sind auch alle 2-Fehler korrigierbar?

Quellen

Pajor - Informatik 4 Tutorium SS2007

Prautzsch - Skript Informatik 4 SS2008

Wavelets for computer graphics: A primer
http://www.cis.udel.edu/~amer/CISC651/wavelets_for_computer_graphics_Stollnitz.pdf

Wikipedia

Reflexion

Was haben wir heute gelernt?

Reflexion

Was haben wir heute gelernt?

- 3. Scheinklausur besprochen

Reflexion

Was haben wir heute gelernt?

- 3. Scheinklausur besprochen
- Hamming-Codes

Noch Fragen?

Vorschau

Vorschau

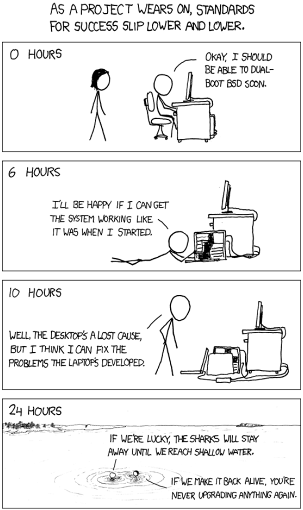
- JPEG-Komprimierung

Vorschau

- JPEG-Komprimierung
- Diskrete Cosinus Transformation

Vorschau

- JPEG-Komprimierung
- Diskrete Cosinus Transformation
- Haar-Wavelets


AS A PROJECT WEARS ON, STANDARDS FOR SUCCESS SLIP LOWER AND LOWER.
0 HOURS
OKAY, I SHOULD BE ABLE TO DUAL-BOOT BSD SOON.
6 HOURS
I'LL BE HAPPY IF I CAN GET THE SYSTEM WORKING LIKE IT WAS WHEN I STARTED.
10 HOURS
WELL, THE DESKTOP'S A LOST CAUSE, BUT I THINK I CAN FIX THE PROBLEMS THE LAPTOP'S DEVELOPED.
24 HOURS
IF WE'RE LUCKY, THE SHARKS WILL STAY AWAY UNTIL WE REACH SHALLOW WATER.
IF WE MAKE IT BACK ALIVE, YOU'RE NEVER UPGRADING ANYTHING AGAIN.