1 Codes

1.1 Ordnungen auf X^*

Es sei $X = \{a_1, \ldots, a_r\}$ ein gemäß $a_1 < a_2 < \ldots < a_r$ geordnetes Alphabet.

Definition 1.1 Lexikographische Ordnung

$$w <_{lex} v \quad :\Leftrightarrow \quad \begin{cases} w \sqsubset v & \text{, oder} \\ \exists u \exists i \exists j (i < j \wedge u a_i \sqsubseteq w \wedge u a_j \sqsubseteq v) & . \end{cases}$$

Definition 1.2 Quasilexikographische Ordnung

$$w <_{ql} v \quad :\Leftrightarrow \quad \begin{cases} |w| < |v| & \text{, oder} \\ |w| = |v| \text{ und } w <_{lex} v & . \end{cases}$$

Eigenschaften

Es sei $X = \{0, \ldots, r-1\}$ ein in der natürlichen Ordnung geordnetes Alphabet, und wir setzen $0.w := \sum_{i=1}^{|w|} x_i \cdot r^{-i} \in \mathbb{R}$ für $w = x_1 \cdots x_{|w|},\ x_i \in X$. Dann gelten

1. Aus $w \leq_{lex} v$ folgt $0.w \leq 0.v$.
2. $w <_{lex} v \iff 0.w < 0.v$ für $v \notin X^* \cdot 0$, und
3. $w <_{lex} v \iff (0.w < 0.v \vee v \in w \cdot 0 \cdot 0^*)$
   $\iff (0.w < 0.v \vee w \sqsubset v)$ für $w, v \in \{0, \ldots, r-1\}^*$.
4. Aus $w \sqsubseteq v$ folgt $0 \leq 0.v - 0.w \leq r^{-|w|} - r^{-|v|}$, und
5. $w \sqsubseteq v \iff 0 \leq 0.v - 0.w \leq r^{-|w|} - r^{-|v|}$ für $v \notin X^* \cdot 0$.

1.2 Codes

Definition 1.3 [Code]
Eine Sprache $C \subseteq X^*$ heißt *(eindeutig decodierbarer) Code* : $\Longleftrightarrow$ Für Wörter $w_1, \ldots, w_n, v_1, \ldots, v_m$ folgen aus $w_1 \cdots w_n = v_1 \cdots v_m$ stets $n = m$ und $w_i = v_i$.

Definition 1.4 [Codierung]
Eine Abbildung $\varphi : Y \to X^*$ heißt *(eindeutig decodierbare, verlustfreie) Codierung* : $\Longleftrightarrow$

1. φ ist eineindeutig, und
2. das Bild $\varphi(Y)$ ist ein eindeutig decodierbarer Code.

Satz 1.1 (Ungleichung von *Kraft*) *Es seien* $r = |X|$ *und* $(\ell_i)_{i \in I}$ $(I \subseteq \mathbb{N}_+)$ *mit* $\ell_1 \leq \ell_2 \leq \ldots$ *eine endliche oder unendliche Familie von natürlichen Zahlen mit der Eigenschaft* $\sum_{i \in I} r^{-\ell_i} \leq 1$. *Dann gibt es einen Präfixcode* $C = \{w_i : i \in I\}$ *mit* $|w_i| = \ell_i$.

Algorithmus KRAFT $(\ell_1, \ell_2, \ldots)$;; $\ell_1 \leq \ell_2 \leq \ldots$, $n = |I|$
$i := 0$, $\ell_0 := 0$, $M_0 := \{e\}$, $C_0 := \emptyset$
While $i \neq n$ do
  $i := i + 1$
  $M := M_{i-1} \cdot X^{\ell_i - \ell_{i-1}}$
  $w_i :=$ erstes Wort in der lexikographischen Ordnung von M
  $C_i := C_{i-1} \cup \{w_i\}$, $M_i := M \setminus \{w_i\}$
Endwhile
Output $C := C_n$

Definition 1.5 [Bernoulli-Maß]
Ein Bernoulli Maß auf X^* ist eine Abbildung $\mu : X^* \to [0,1]$ die die folgenden Bedingungen erfüllt:

1. $\sum_{a \in X} \mu(a) = 1$, und
2. $\forall w, v (w, v \in X^* \to \mu(w \cdot v) = \mu(w) \cdot \mu(v))$.

Folgerung 1.2 *Ist μ Bernoulli-Maß auf X^*, so gelten*

1. $\mu(e) = 1$, *und*
2. $\mu(W \cdot V) \leq \mu(W) \cdot \mu(V)$, *wobei für $\prod_{a \in X} \mu(a) > 0$ und $\mu(W), \mu(V) < \infty$ die Gleichheit genau dann eintritt, wenn jedes $u \in W \cdot V$ genau eine Faktorisierung $u = w \cdot v$ mit $w \in W$ und $v \in V$ hat.*

Satz 1.3 (Satz von MacMillan) *Ist μ ein Bernoulli-Maß auf X^*, so gilt für jeden Code $C \subseteq X^*$ die Beziehung $\mu(C) \leq 1$.*

1.3 Präfix-Codes und Bäume

Definition 1.6 [r-verzweigter Wurzelbaum]
Ein *höchstens r-verzweigter Wurzelbaum* kann interpretiert werden als präfixabgeschlossene Teilmenge $W \subseteq \{0, 1, \ldots, r-1\}^*$ mit der Wurzel $e \in \{0, 1, \ldots, r-1\}^*$ und den Nachfolgern $w \cdot i$ ($i \in M_w \subseteq \{0, 1, \ldots, r-1\}$) des Knotens $w \in \{0, 1, \ldots, r-1\}^*$.

Wörter $w \in W$ ohne Nachfolger in W sind die Blätter des Baumes, die anderen Wörter $w \in W$ sind die inneren Knoten.

Folgerung 1.4 *Ist $W \subseteq \{0, 1, \ldots, r-1\}^*$ präfixabgeschlossen und gilt $|W| \neq 1$, so bildet die Menge der Blätter von W einen Präfixcode.*

1.4 Mittlere Codewortlänge

Definition 1.7 Es sei $Y = \{a_1, \ldots, a_k\}$, und es sei $P = (p_1, \ldots, p_k)$ mit $\sum_{i=1}^k p_i = 1$ ein k-Vektor nichtnegativer Zahlen. Dann nennen wir (Y, P) eine *Quelle*.

Weiter nennen wir $H_m(P) := \sum_{i=1}^k -p_i \cdot \log_m p_i$ die *Entropie* der Quelle (Y, P).

Satz 1.5 (Abschätzung für die minimale mittlere Codewortlänge) *Es seien (Y, P) eine Quelle und $\varphi : Y \to X^*$ eine eindeutig decodierbare Codierung. Dann gilt für die mittlere Codewortlänge, $\overline{l(\varphi)}$, von φ die Beziehung:*

$$\overline{l(\varphi)} := \sum_{i=1}^k p_i \cdot |\varphi(a_i)| \geq H_{|X|}(P)\,.$$

Satz 1.6 (Shannon) *Zu jeder Häufigkeitsverteilung $P = (p_1, \ldots, p_k)$ und zu jedem Alphabet X gibt es einen Präfix-Code $C \subseteq X^*, C = \{w_1, \ldots, w_k\}$, mit einer mittleren Codewortlänge $\overline{l(C)}$, die der Abschätzung $\overline{l(C)} \leq H_{|X|}(P) + 1$ genügt.*

Definition 1.8 Eine eindeutig decodierbare Codierung $\varphi : \{a_1, \ldots, a_k\} \to \{x_1, \ldots, x_m\}^*$ heißt *alphabetisch*, wenn $\varphi(a_i) <_{\text{lex}} \varphi(a_{i+1})$.

Satz 1.7 *Zu jeder Quelle (Y, P) und zu jedem Alphabet X gibt es eine alphabetische Codierung φ, mit einer mittleren Codewortlänge $\overline{l(\varphi)}$, die der Abschätzung $\overline{l(\varphi)} \leq H_{|X|}(P) + 2$ genügt.*

Ternärer Shannon – Code

Häufigkeit	r_{j-1}	ternär	$\ell_j = \lceil \log_3 \frac{1}{p_j} \rceil$	Codewort
$p_1 = 0.4$	0.0	0.0000000	1	0
$p_2 = 0.2$	0.4	0.1012101	2	10
$p_3 = 0.1$	0.6	0.1210121	3	121
$p_4 = 0.1$	0.7	0.2002200	3	200
$p_5 = 0.1$	0.8	0.2101210	3	210
$p_6 = 0.1$	0.9	0.2200220	3	220

$\bar{l} = 2.0\,, \qquad H_{\mathcal{P}} = 1.465$

Binärer Shannon – Code

Häufigkeit	r_{j-1}	binär	$\ell_j = \lceil \log_2 \frac{1}{p_j} \rceil$	Codewort
$p_1 = 0.4$	0.0	0.000000000	2	00
$p_2 = 0.2$	0.4	0.011001101	3	011
$p_3 = 0.1$	0.6	0.100110011	4	1001
$p_4 = 0.1$	0.7	0.101100110	4	1011
$p_5 = 0.1$	0.8	0.110011010	4	1100
$p_6 = 0.1$	0.9	0.111001101	4	1110

$\bar{l} = 3.0\,, \qquad H_{\mathcal{P}} = 2.322$

Binärer Alphabetischer Code

Häufigkeit	r_{j-1}	binär	$\ell_j = \lceil \log_2 \frac{1}{p_j} \rceil + 1$	Codewort
$p_1 = 0.1$	0.05	0.000011010	5	00001
$p_2 = 0.1$	0.15	0.001001101	5	00100
$p_3 = 0.4$	0.4	0.011001101	3	011
$p_4 = 0.1$	0.65	0.101001101	5	10100
$p_5 = 0.2$	0.8	0.110011010	4	1100
$p_6 = 0.1$	0.95	0.111100110	5	11110

$\bar{l} = 4.0\,, \qquad H_{\mathcal{P}} = 2.322$

2 Sprachentheoretische Eigenschaften von Codes

2.1 Maximalität und Vollständigkeit von Codes

Definition 2.1 [Maximale Codes] Eine Sprache $C \subseteq X^*$ heißt *maximaler Code* : $\Longleftrightarrow$
Für jeden Code $C' \subseteq X^*$ mit $C \subseteq C'$ gilt $C' = C$.

Folgerung 2.1 *Es sei μ ein nicht-entartetes Bernoulli-Maß auf X^*. Ist $C \subseteq X^*$ ein Code mit $\mu(C) = 1$, so ist C maximal.*

Satz 2.2 (Über die Existenz von maximalen Codes)
Zu jedem Code $C \subseteq X^*$ gibt es einen maximalen Code, der C enthält.

Shannons Kanalkapazität

$$C(\Phi) := \lim_{t\to\infty} \frac{\log_r |\{x_1 \cdots x_t : x_1 \cdots x_t \text{ ist Ausgabe von } \Phi\}|}{t}$$

Vollständigkeit von Codes

$$\mathbf{T}(W) := \{v : \exists w, v_1, v_2 (w \in W \wedge v_1, v_2 \in X^* \wedge v_1 \cdot v \cdot v_2 = w)\}$$

Definition 2.2 1. Eine Sprache $W \subseteq X^*$ heißt *dicht*, falls $\mathbf{T}(W) = X^*$ erfüllt ist.

2. Anderenfalls nennen wir eine Sprache $W \subseteq X^*$ *mager*.

Folgerung 2.3 *Ist $W \subseteq X^*$ mager, so gilt $\mu(W) < \infty$ für jedes Bernoulli-Maß $\mu : X^* \to (0,1]$ auf X^*.*

Lemma 2.4 *Es sei μ ein Bernoulli-Maß auf X^*. Dann gelten:*

1. *Ist $\mu(W^*) = \infty$, so gilt $\mu(W) \geq 1$.*
2. *Ist $C \subseteq X^*$ ein Code, so gilt genau dann $\mu(C^*) = \infty$, wenn $\mu(C) = 1$ ist.*

Definition 2.3 Eine Sprache $W \subseteq X^*$ heißt *vollständig*, falls W^* dicht ist.

Satz 2.5 *Ist $W \subseteq X^*$ eine magere und vollständige Sprache, so gilt $\mu(W^*) = \infty$ für jedes Bernoulli-Maß $\mu : X^* \to (0,1]$ auf X^*.*

Lemma 2.6 *Es sei $V \subseteq X^*$ und es gebe ein festes $k \in \mathbb{N}$ derart, dass $\mathbf{T}(V) \subseteq \{v : \exists w_1 \exists w_2 (|w_1|, |w_2| \leq k \wedge w_1 \cdot v \cdot w_2 \in V\}$.*

Dann gilt für jedes Bernoulli-Maß μ mit $\mu_{\min} := \min\{\mu(x) : x \in X\} > 0$ die Ungleichung

$$\mu(\mathbf{T}(V)) \geq \mu(V) \geq \frac{\mu_{\min}^{2k}}{(k+1)^2} \cdot \mu(\mathbf{T}(V)) .$$

Folgerung 2.7 *Ist $C \subseteq X^*$ ein magerer und vollständiger Code, so gilt $\mu(C) = 1$ für jedes Bernoulli-Maß $\mu : X^* \to (0,1]$ auf X^*.*

$\mathcal{M} := \{\mu \mid \mu : X^* \to (0,1]\}$ sei die Menge aller nicht-entarteten Bernoulli-Maße auf X^*. Für $W \subseteq X^*$ gilt:

$$\begin{array}{lcl} W \text{ ist dicht} & \Longleftarrow & \exists\mu(\mu \in \mathcal{M} \wedge \mu(W) = \infty) \\ & & \Uparrow \\ & & \forall\mu(\mu \in \mathcal{M} \to \mu(W) = \infty) \end{array}$$

Für **mageres** (d.h. nicht dichtes) $L \subseteq X^*$ gilt darüberhinaus:

$$\begin{array}{lcl} & & \exists\mu(\mu \in \mathcal{M} \wedge \mu(L) \geq 1) \\ & & \Uparrow \\ L \text{ ist vollständig} & \Longrightarrow & \forall\mu(\mu \in \mathcal{M} \to \mu(L) \geq 1) \end{array}$$

Der Vervollständigungssatz von *Ehrenfeucht* und *Rozenberg*

Definition 2.4 Wir nennen ein Wort $y \in X^*$ *unberandet*, falls $y \neq e$ und kein echtes Präfix $e \sqsubset w \sqsubset y$ zugleich auch Suffix von y ist.

Satz 2.8 *Es seien $C \subseteq X^*$ ein Code und $y \notin T(C^*)$ ein unberandetes Wort. Weiter sei $U := (X^* \setminus X^* \cdot y \cdot X^*) \setminus C^*$, dann ist $\hat{C} := C \cup y \cdot (U \cdot y)^*$ ein vollständiger Code.*

Lemma 2.9 *Es seien $v \in X^*$ ein unberandetes Wort, $V := X^* \setminus X^* \cdot y \cdot X^*$ die Menge aller Wörter, die y nicht als Infix enthalten. Dann ist $C := V \cdot y$ ein Präfix-Code.*

Satz 2.10 *Ist C ein maximaler Code, so ist C ein vollständiger Code.*

$\mathcal{M} := \{\mu \mid \mu : X^* \to (0,1]\}$ sei die Menge aller nicht-entarteten Bernoulli-Maße auf X^*. Dann gilt für **Codes** $C \subseteq X^*$:

$$\begin{array}{ccc} C \text{ ist maximal} & \Longleftarrow & \exists \mu(\mu \in \mathcal{M} \wedge \mu(C) = 1) \\ \Downarrow & & \Uparrow \\ C \text{ ist vollständig} & \overset{C \text{ mager}}{- - - \to} & \forall \mu(\mu \in \mathcal{M} \to \mu(C) = 1) \end{array}$$

Folgerung 2.11 *Ein Code $C \subseteq X^*$ ist genau dann vollständig, wenn C maximal oder dicht ist.*

Folgerung 2.12 *Ist $C \subseteq X^*$ ein Code, so gibt es einen vollständigen Code $\hat{C}$, der C enthält.*
Ist C darüber hinaus reguläre Sprache, so kann auch $\hat{C}$ regulär gewählt werden.

Lemma 2.13 *Es sei $L \subseteq X^*$ eine reguläre Sprache. Dann sind die folgenden Aussagen äquivalent:*

1. *L ist dicht, d.h. $\boldsymbol{T}(L) = X^*$.*
2. *Es existiert ein Bernoulli-Maß $\mu : X^* \to (0,1]$ mit $\mu(L) = \infty$.*
3. *Für alle Bernoulli-Maße $\mu : X^* \to (0,1]$ gilt $\mu(L) = \infty$.*

Folgerung 2.14 *Jeder reguläre Code $C \subseteq X^*$ ist mager.*

Definition 2.5 Ein Code $C \subseteq X^*$ heißt *präfix-maximal*: $\Longleftrightarrow$

1. C ist Präfix-Code, und
2. für jeden Präfix-Code $C' \supseteq C$ gilt $C' = C$.

Lemma 2.15 *Ein Präfix-Code $C \subseteq X^*$ ist genau dann präfix-maximal, wenn für es jedes $w \in X^*$ ein $v \in C$ mit $w \sqsubseteq v$ oder $v \sqsubseteq w$ gibt.*

Der Suffix-Code $\bigcup_{n\in\mathbb{N}} X^n \cdot b \cdot a^n$ enthält einen echten präfix-maximalen Teilcode $\hat{C}$, d.h. $\hat{C}$ ist nicht maximal als Code.

Lemma 2.16 *Jeder (endliche) Präfix-Code ist in einem maximalen (endlichen) Präfix-Code enthalten.*
Jeder endliche maximale Präfix-Code ist auch maximal als Code.

Es sei $\varphi : X \to Y^*$ eine Codierung, deren Erweiterung auf X^* ebenfalls mit φ bezeichnet werde.

Lemma 2.17 *1. Ist $C \subseteq X^*$ ein Code, so ist auch $\varphi(C) \subseteq Y^*$ ein Code.*

2. Ist $C' \subseteq Y^*$ ein Code, so ist auch $\varphi^{-1}(C') \subseteq X^*$ ein Code.

2.2 Entscheidungsprobleme und Decodierverzögerung

Definition 2.6 Wir setzen für $C \subseteq X^*$:

$$\begin{aligned} w <_1 v &:\iff \exists u(u \in C \wedge u = w \cdot v) \\ \{e\}\ w <_2 v &:\iff \exists u(u \in C \wedge w = u \cdot v)\text{ , sowie} \\ w < v &:\iff w <_1 v \vee w <_2 v \end{aligned}$$

Lemma 2.18 (Levenštejn) *Es gibt genau dann eine C-Kette $v_1 < \ldots < v_n$, wenn es Wörter $w_1, \ldots, w_i \in C$ und $u_1, \ldots, u_j \in C$ mit $w_1 \neq u_1$ und $i + j = n$ derart gibt, daß*

1. $v_1 \in \{w_1, u_1\}$,

2. $u_1 \cdots u_j \cdot v_n = w_1 \cdots w_i$ und

3. $|v_n| \leq |w_i|$.

Folgerung 2.19 *Eine Sprache $C \subseteq X^* \setminus \{e\}$ ist genau dann ein Code, wenn es keine endliche C-Kette $v_1 < \ldots < v_n$ $(n \geq 2)$, die auf ein Wort $v_n \in C$ endet, gibt.*

Folgerung 2.20 *Es gibt ein Verfahren, welches zu endlichem $L \subseteq X^* \setminus \{e\}$ entscheidet, ob L ein Code ist.*

Satz 2.21 *Es gibt ein Verfahren, welches zu regulärem $L \subseteq X^* \setminus \{e\}$ entscheidet, ob L ein Code ist.*

Es gibt kein Verfahren, welches zu kontextfreiem $L \subseteq X^* \setminus \{e\}$ entscheidet, ob L ein Code ist.

Definition 2.7 [Decodierverzögerung] Es seien $C \subseteq X^*$ ein Code und $w \in C$.

- Das Codewort w hat die Decodierverzögerung $m(w) \in \mathbb{N}$, falls aus $wv_1 \ldots v_{m(w)} \sqsubseteq w'u$ mit $v_i, w' \in C$ und $u \in C^*$ stets $w = w'$ folgt.
- C hat endliche Decodierverzögerung, falls jedes $w \in C$ eine endliche Decodierverzögerung hat.
- C hat beschränkte Decodierverzögerung, falls es ein $n \in \mathbb{N}$ mit $\forall w(w \in C \to m(w) \leq n)$ gibt.

Beispiele für Decodierverzögerung

Verzögerung ist	aber nicht	Gegenbeispiel
–	endlich	$\{a, ab, bb\}$, Suffix-Code
endlich	beschränkt	$ab^*a \cup \{ab^{n+1}ab^na : n \in \mathbb{N}\}$ $\cup \{b^{n+1}ab^na : n \in \mathbb{N}\}$
n $(n \geq 1)$	$n-1$	$\{a, a^nb\}$
0 (Präfix-Code)	Bifix-Code	$\{a, ba\}$

Lemma 2.22 *Wenn für jedes $w \in C$ jede mit w beginnende C-Kette eine Länge $\leq l_w$ hat, so hat C endliche Decodierverzögerung.*

Lemma 2.23 *Es sei C ein Code.*

1. *Ist jede C-Kette nicht länger als ℓ, so hat C eine Decodierverzögerung von höchstens $\ell - 1$.*
2. *Hat C die Decodierverzögerung $m \in \mathbb{N}$, so ist keine C-Kette länger als $2m + 2$.*