

5. Woche

Perfekte und Optimale Codes, Schranken

Packradius eines Codes (Wiederholung)

Definition Packradius eines Codes

Sei C ein (n, M, d) -Code. Der Packradius $pr(C) \in \mathbb{N}$ von C ist die größte Zahl, so dass die Hammingkugeln $B^n(\mathbf{c}, pr(C))$ für alle $\mathbf{c} \in C$ disjunkt sind.

Korollar

Sei C ein (n, M, d) -Code.

- 1 Der Packradius von C ist $pr(C) = \lfloor \frac{d-1}{2} \rfloor$.
- 2 C ist genau v -fehlerkorrigierend gdw $pr(C) = v$.

0^n als „standard“ Codewort

Lemma Äquivalenter Code mit 0^n

Sei C ein (n, M, d) -Code. Dann gibt es einen (n, M, d) -Code C' mit $0^n \in C'$.

Beweis: Übung. Wichtige Idee, man kann ein beliebiges Wort x nehmen, und einen Code C zusammen mit einer Abbildung $C \rightarrow C'$ konstruieren, so dass $x \mapsto 0^n$ und alle Abstände respektiert werden.

Beweisidee vom Korollar.1: wenn wir Kugeln größeren Radius nehmen, können wir explizit nichtleere Schnitte konstruieren. Übung.

Perfekte Codes

Die an den Codewörtern zentrierten Kugeln und mit Radius $\frac{d-1}{2}$ überschneiden einander nicht. Also $M \cdot V^n \left(\left\lfloor \frac{d-1}{2} \right\rfloor \right) \leq 2^n$ und somit $M \leq \frac{2^n}{V^n \left(\left\lfloor \frac{d-1}{2} \right\rfloor \right)}$ d.h. $M \leq \frac{2^n}{\sum_{i=0}^{\lfloor (d-1)/2 \rfloor} \binom{n}{i}}$. (Kugelpackungsschranke).
Max in der Tat in einigen Fällen erreichbar.

Definition Perfekter Code

Sei $C \subseteq \{0, 1\}^n$ ein (n, M, d) -Code. C heißt *perfekt*, falls

$$M \cdot V^n \left(\left\lfloor \frac{d-1}{2} \right\rfloor \right) = 2^n.$$

D.h. die maximalen disjunkten Hammingkugeln um die Codeworte partitionieren $\{0, 1\}^n$.

Nicht für alle (n, M, d) , die obige Bedingung erfüllen, gibt es auch einen Code.

Perfekte Codes

- $\{0, 1\}^n$ ist ein $(n, 2^n, 1)$ -Code
 - ▶ Packradius ist 0, Hammingkugeln bestehen nur aus Codewort selbst.
 - ▶ Perfekter Code, aber nutzlos für Fehlerkorrektur.
- $R(n)$ ist für ungerade n ein perfekter $(n, 2, n)$ -Code.
 - ▶ $2 \cdot \sum_{i=0}^{\frac{n-1}{2}} \binom{n}{i} = 2 \cdot \frac{2^n}{2} = 2^n$
 - ▶ Code ist nutzlos, da er nur zwei Codeworte enthält.
- Der Golay Code $(23, 2^{12}, 7)$ ist perfekt.
 - ▶ $2^{12} \cdot \sum_{i=0}^3 \binom{23}{i} = 2^{11} \cdot 2^{12} = 2^{23}$
- Der Hamming Code $\mathcal{H}(h) = (n, M, d) = (2^h - 1, 2^{n-h}, 3)$ ist perfekt.
 - ▶ $2^{n-h} (1 + 2^h - 1) = 2^n$
- Die einzigen perfekten, binären v -fehlerkorrigierenden Codes mit $v \geq 2$ sind Repetitions codes und der obige Golay Code.

Die Größe $A(n, d)$ und optimale Codes

Definition Optimaler Code

Wir definieren

$$A(n, d) = \max\{M \mid \exists \text{ binärer } (n, M, d) - \text{Code}\}$$

Ein (n, M, d) -Code heißt optimal, falls $M = A(n, d)$.

- Bestimmung von $A(n, d)$ ist offenes Problem.
- Zeigen hier obere und untere Schranken für $A(n, d)$.
- Für kleine Werte von n, d bestimmen wir $A(n, d)$ wie folgt:
 - ▶ Zeigen $A(n, d) \leq M$.
 - ▶ Konstruieren (n, M, d) -Code.
- $A(n, d) \leq 2^n$ für $d \in [n]$: höchstens 2^n Codeworte der Länge n .
- $A(n, 1) = 2^n$: $C = \{0, 1\}^n$.
- $A(n, n) = 2$: $R(n)$.
- $A(n, d) \leq A(n, d')$ für $1 \leq d' \leq d \leq n$ (Übung).

Erstes nicht-triviales Resultat

Satz

$$A(4, 3) = 2.$$

- Sei C ein optimaler $(4, M, 3)$ -Code. OBdA $0000 \in C$.
- Worte mit Abstand mindestens 3 von 0000:

0111, 1011, 1101, 1110, 1111.

- Je zwei Worte besitzen Abstand höchstens 2, d.h. $A(4, 3) \leq 2$.
- Für $C = \{0000, 0111\}$ gilt $d(C) = 3$ und damit $A(4, 3) = 2$.

Verkürzen eines Codes

Definition Verkürzter Code

Sei C ein (n, M, d) -Code und $j \in [n]$, $b \in \{0, 1\}$. Der *bezüglich b -Bit an j -ter Position verkürzte Code* C' entsteht aus C durch

- 1 Einschränkung auf Codeworte aus C , deren j -tes Bit b ist.
- 2 Herausstreichen der j -ten Stelle.

- **Bsp:** Verkürzen von $C = \{001, 010, 101\}$ bezüglich 0-Bit an 1. Position liefert $C' = \{01, 10\}$.
- Beachte C besitzt Abstand 1, aber $d(C') = 2$.

Satz Verkürzter Code

Sei C ein (n, M, d) -Code und C' ein verkürzter Code. Dann gilt $d(C') \geq d$.

- Betrachten nur die bezüglich einer Stelle j konstanten Codeworte.
- Stelle j kann nicht zum Abstand beitragen.

Rekursive Schranke für $A(n, d)$

Lemma Rekursive Schranke

Für $n \geq 2$ gilt: $A(n, d) \leq 2 \cdot A(n-1, d)$.

- Sei C ein optimaler (n, M, d) -Code.
- Sei C_b der bezügl. b -Bit, $b \in \{0, 1\}$ an 1. Position verkürzte Code.
- Aus $d(C_b) \geq d$ folgt

$$\begin{aligned} A(n, d) = M &= |C_0| + |C_1| \leq A(n-1, d(C_0)) + A(n-1, d(C_1)) \\ &\leq A(n-1, d) + A(n-1, d). \end{aligned}$$

Korollar

$$A(5, 3) = 4.$$

- $A(5, 3) \leq 2 \cdot A(4, 3) = 4$.
- $C = \{00000, 11100, 00111, 11011\}$ besitzt $d(C) = 3$.

Schnitt von Strings

Definition Hamminggewicht, Schnitt

Seien $\mathbf{x}, \mathbf{y} \subseteq \{0, 1\}^n$. Das Hamminggewicht $w(\mathbf{x})$ von $\mathbf{x}$ ist definiert als die Anzahl von Einsen in $\mathbf{x}$.

Seien $\mathbf{x} = x_1 \dots x_n$, $\mathbf{y} = y_1 \dots y_n$. Dann ist der *Schnitt* definiert als

$$\mathbf{x} \cap \mathbf{y} = x_1 \cdot y_1 \dots x_n \cdot y_n$$

Lemma Abstand via Gewicht

Seien $\mathbf{x}, \mathbf{y} \subseteq \{0, 1\}^n$. Dann gilt

$$d(\mathbf{x}, \mathbf{y}) = w(\mathbf{x}) + w(\mathbf{y}) - 2w(\mathbf{x} \cap \mathbf{y})$$

- Sei $\#_{bc}$: Anzahl Positionen mit b in $\mathbf{x}$ und c in $\mathbf{y}$, $b, c \in \{0, 1\}$.
- Es gilt

$$\begin{aligned} d(\mathbf{x}, \mathbf{y}) &= \#_{10} + \#_{01} = (\#_{10} + \#_{11}) + (\#_{01} + \#_{11}) - 2\#_{11} \\ &= w(\mathbf{x}) + w(\mathbf{y}) - 2w(\mathbf{x} \cap \mathbf{y}). \end{aligned}$$

Ungerade d genügen

Satz

Sei $d \geq 1$ ungerade. Dann existiert ein (n, M, d) -Code C gdw ein $(n+1, M, d+1)$ -Code C' existiert.

“ $\Leftarrow$ ”: Sei C' ein $(n+1, M, d+1)$ -Code.

- Seien $\mathbf{c}, \mathbf{c}' \in C'$ mit $d(\mathbf{c}, \mathbf{c}') = d+1$ und i eine Position mit $\mathbf{c}_i \neq \mathbf{c}'_i$.
- Lösche i -te Position aus C' . Resultierender Code C besitzt $d(C) = d$ und Länge n .

“ $\Rightarrow$ ”: Sei C ein (n, M, d) Code.

- C : Erweitere C um Paritätsbit, so dass $w(\mathbf{c})$ gerade für alle $\mathbf{c} \in C$.
- Mittels Lemma

$$d(\mathbf{x}, \mathbf{y}) = w(\mathbf{x}) + w(\mathbf{y}) - 2w(\mathbf{x} \cap \mathbf{y}) \equiv 0 \pmod{2} \quad \text{für alle } \mathbf{x}, \mathbf{y} \in C'.$$

- Wegen $d = d(C)$ ungerade und $d \leq d(C') \leq d+1$ folgt $d(C') = d+1$.

Einige Werte von $A(n, d)$ (Quelle: Sloane 1982)

Korollar

$A(n, d) = A(n + 1, d + 1)$ für $d \geq 1$ ungerade.

n	5	6	7	8	9	10	11	16
$d = 3$	4	8	16	20	40	72-79	144-158	2560-3276
$d = 5$	2	2	2	4	6	12	24	256-340
$d = 7$	-	-	2	2	2	2	4	36-37

Gilbert-Varshamov-Schranke (Sphere-Covering) und Kugelpackungsschranke (Sphere-Packing)

Satz Kugelschranken

$$\frac{2^n}{V^n(d-1)} \leq A(n, d) \leq \frac{2^n}{V^n(\lfloor \frac{d-1}{2} \rfloor)}.$$

Untere Schranke, d.h. Gilbert-Varshamov-Schranke (Sphere-Covering)

- Sei C ein optimaler (n, M, d) -Code, d.h. $M = A(n, d)$.
- Für alle $\mathbf{x} \in \{0, 1\}^n \exists \mathbf{c} \in C: d(\mathbf{x}, \mathbf{c}) < d$. (Warum?)

$$\{0, 1\}^n \subseteq \bigcup_{i=1}^M B^n(\mathbf{c}_i, d-1) \Rightarrow 2^n \leq V^n(d-1) \cdot M.$$

Obere Schranke, d.h. Kugelpackungsschranke (Sphere-Packing)

- Bei „Perfekten Codes“ haben wir gezeigt, dass für einen (n, M, d) -Code C gilt $M \leq \frac{2^n}{V^n(\lfloor \frac{d-1}{2} \rfloor)}$. Insbesondere gilt dies auch für C optimal, d.h. mit $M = A(n, d)$.

Bsp $A(n, 3)$ für Sphere-Covering und Sphere-Packing

n	5	6	7	8	9	10	11	16
untere	2	3	5	7	12	19	31	479
$A(n, 3)$	4	8	16	20	40	72-79	144-158	2560-3276
obere	5	9	16	28	51	93	170	3855

Singleton-Schranke

Satz Singleton-Schranke

$$A(n, d) \leq 2^{n-d+1}$$

- Sei C ein optimaler (n, M, d) -Code. Entferne letzte $d - 1$ Stellen.
- Resultierende Codeworte sind alle verschieden, da sich $\mathbf{c} \in C$ in mindestens d Stellen unterscheiden.
- Es gibt M viele verkürzte unterschiedliche Codeworte der Länge $n - (d - 1)$:

$$M \leq 2^{n-d+1}.$$

Vereinfachte Plotkin-Schranke

Satz Vereinfachte Plotkin-Schranke

Sei $n < 2d$, dann gilt

$$A(n, d) \leq \frac{2d}{2d - n}.$$

- Sei C ein optimaler (n, M, d) – Code und $S = \sum_{i < j} d(\mathbf{c}_i, \mathbf{c}_j)$.
- Je zwei Codeworte besitzen Abstand mindestens d , d.h.
 $S \geq d \binom{M}{2}$.
- Betrachten erste Stelle in allen Codeworten:
 - ▶ Sei k die Anzahl der Nullen und $(M - k)$ die Anzahl der Einsen.
 - ▶ Erste Stelle liefert Beitrag von $k(M - k)$ zu S .
 - ▶ $k(M - k)$ ist maximal für $k = \frac{M}{2}$, d.h. $k(M - k) \leq \frac{M^2}{4}$.
 - ▶ Analog für jede der n Stellen, d.h. $S \leq \frac{nM^2}{4}$.
- Kombination beider Schranken und Auflösen nach M liefert

$$M \leq \frac{2d}{2d - n}.$$

Vergleich der oberen Schranken

n	7	8	9	10	11	12	13
$A(n, 7)$	2	2	2	2	4	4	8
Singleton	2	4	8	16	32	64	128
Plotkin	2	2	2	3	4	7	14

Die Rate eines Codes

Definition Rate eines Codes

Sei C ein (n, M, d) -Code.

- ➊ Die *Übertragungsrate* ist definiert als $\mathcal{R}(C) = \frac{\log_2(M)}{n}$.
- ➋ Die *Fehlerrate* ist definiert als $\delta(C) = \frac{\lfloor \frac{d-1}{2} \rfloor}{n}$.

Beispiele:

- $C = \{0^n\}$ hat Übertragungsrate 0, aber perfekte Fehlerkorrektur.
- $C = \{0, 1\}^n$ hat Übertragungsrate 1, aber keine Fehlerkorrektur.
- $\mathcal{R}(R(n)) = \frac{1}{n}$ und $\delta(R(n)) = \frac{\lfloor \frac{n-1}{2} \rfloor}{n}$.
 - ▶ Übertragungsrate konvergiert gegen 0, Fehlerrate gegen $\frac{1}{2}$.
- $\mathcal{R}(\mathcal{H}(h)) = \frac{n-h}{n} = 1 - \frac{h}{n}$ und $\delta(\mathcal{H}(h)) = \frac{1}{n}$.
 - ▶ Übertragungsrate konvergiert gegen 1, Fehlerrate gegen 0.

Shannons Theorem für fehlerbehaftete Kanäle

Codierungstheorem von Shannon für fehlerbehaftete Kanäle

Gegeben sei ein binärer symmetrischer Kanal Q mit Fehlerws p . Für alle $R < 1 + p \log_2 p + (1 - p) \log_2(1 - p) = 1 - H(Q)$ und alle $\epsilon > 0$ gibt es für hinreichend große n einen (n, M) -Code C mit Übertragungsrate $\mathcal{R}(C) \geq R$ und $\mathcal{W}_s(\text{Decodierfehler}) \leq \epsilon$.

- Beweis komplex, nicht-konstruktiv.
- Resultat gilt nur asymptotisch für genügend große Blocklänge.