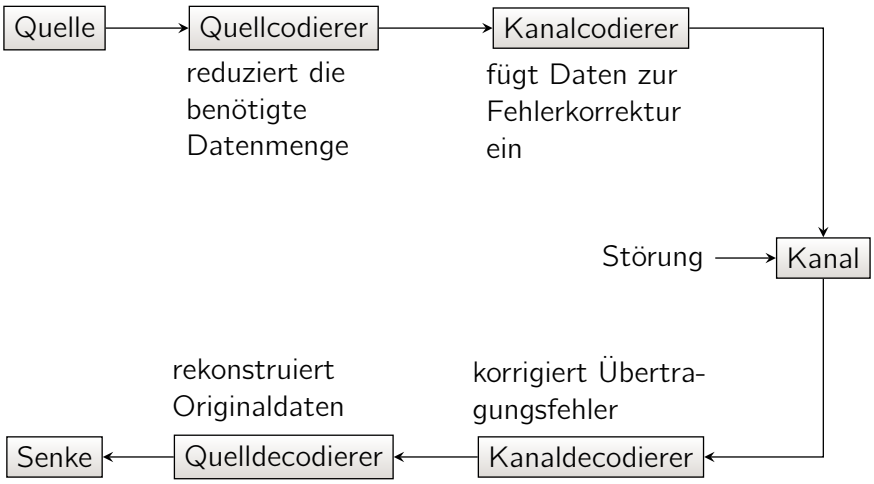
13 Rück- und Ausblick

13.1 Entropie diskreter Quellen
13.2 Kanäle
13.3 Quellcodierung
13.4 Kanalcodierung

13 Rück- und Ausblick

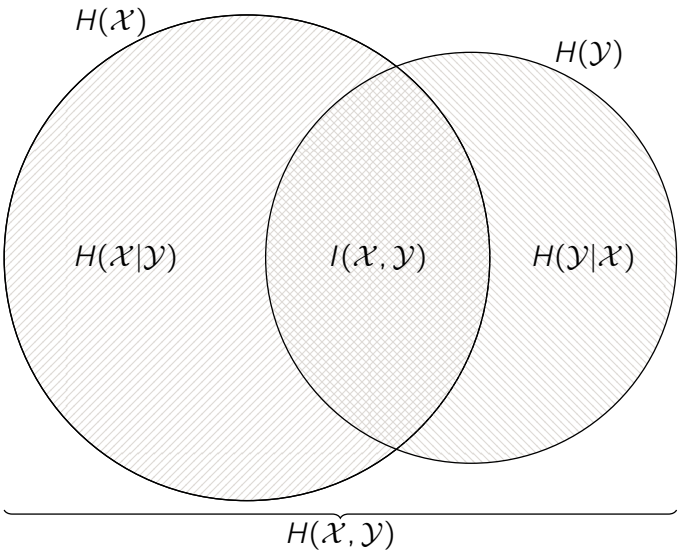


13.1 Entropie diskreter Quellen

13.1.1 Was wir gemacht haben

- Eine Informationsquelle $\mathcal{X}$ liefert Symbole x aus einem Alphabet X; die Auftrittswahrscheinlichkeit der Symbole ist $p_X(x)$.
- Für unabhängige Symbole gilt für den durchschnittlichen Informationsgehalt, die Entropie
$$H(\mathcal{X}) = -\sum_{x\in X} p_X(x)\,\mathrm{ld}\,p_X(x).$$
- Werden Symbole aus zwei Quellen $\mathcal{X}$ und $\mathcal{Y}$ paarweise zusammengefasst, so gilt gilt für die Verbundentropie
$$H(\mathcal{X},\mathcal{Y}) \le H(\mathcal{X}) + H(\mathcal{Y}).$$
Gleichheit gilt dabei nur für zwei unabhängige Quellen.
- Die zusätzliche Information, die $\mathcal{Y}$ liefert, wenn $\mathcal{X}$ bekannt ist, ist die bedingte Entropie
$$H(\mathcal{Y}|\mathcal{X}) = -\sum_{x\in X}\sum_{y\in Y} p_X(x)p_{Y|X}(y|x)\,\mathrm{ld}\,p_{Y|X}(y|x).$$
Es gilt $H(\mathcal{X},\mathcal{Y}) = H(\mathcal{X}) + H(\mathcal{Y}|\mathcal{X})$ und $H(\mathcal{Y}|\mathcal{X}) \le H(\mathcal{Y})$.

- Die Transinformation $I(\mathcal{X},\mathcal{Y}) = H(\mathcal{X}) - H(\mathcal{X}|\mathcal{Y}) = H(\mathcal{Y}) - H(\mathcal{Y}|\mathcal{X})$ gibt an, wieviel Information $\mathcal{X}$ und $\mathcal{Y}$ übereinander tragen.
- Die Zusammenhänge zwischen der verschiedenen Größen lassen sich mit Hilfe des Venn-Diagramms darstellen:



13.1.2 Was wir nicht gemacht haben

- Quellen mit abhängigen Ereignissen, insbesondere Markov-Quellen.
- Quellen mit kontinuierlichem Wertebereich (z.B. reelle Zahlen). Kernidee: Übergang von Summen zu Integralen.

13.2 Kanäle

13.2.1 Was wir gemacht haben

- Ein Kanal wird mit zwei Quellen modelliert; eine für die Symbole am Kanaleingang, eine für die Symbole am Kanalausgang.
- Die Transinformation beschreibt, wie viel Information vom Kanaleingang zum Kanalausgang gelangt; sie hängt von der Übergangswahrscheinlichkeit des Kanal und der Verteilung der Eingangsdaten ab.
- Die Kanalkapazität ist das Maximum der Transinformation über alle Eingangsverteilungen (und damit nur noch von der Übergangswahrscheinlichkeit des Kanals abhängig.)
- Für einen symmetrisch gestörten Binärkanal mit der Fehlerwahrscheinlichkeit p beträgt die Kanalkapazität

$$K = 1 + (1-p) \cdot \mathrm{ld}(1-p) + p \cdot \mathrm{ld}(p).$$

- Für einen analogen AWGN-Kanal mit der Singalleistung σ_x und der Rauschleistung σ_s beträgt die Kanalkapzität

$$K = \frac{1}{2} \mathrm{ld}\left(1 + \frac{\sigma_x^2}{\sigma_s^2}\right).$$

13.2.2 Was wir nicht gemacht haben

- Kaskadierte Kanäle.
- Kanäle mit mehreren Sendern und Empfängern (Broadcast).
- Relay-Kanäle.

13.3 Quellcodierung

13.3.1 Was wir gemacht haben

- Aufgabe: Darstellung der Symbole einer Quelle mit möglichst wenigen Bits.
- Die Entropie ist eine untere Schranke für die mittlere Codewortlänge jedes eindeutig decodierbaren Codes.
- Präfix-Codes können die gleiche mittlere Codewortlänge erreichen wie allgemeine eindeutig decodierbare Codes.
- Shannon: Werden Blöcke von Symbolen auf einmal codiert, so erreicht die mittlere Codewortlänge für große Blocklängen asymptotisch die Entropie.
- Die Huffman-Codierung, bei der sukzessive die beiden Symbole mit der geringsten Auftrittswahrscheinlichkeit zusammengefasst werden, ist optimal, d.h. keinen andere Codierung erreicht eine geringere mittlere Codewortlänge.
- Die arithmetische Codierung (bzw. Bereichscodierung) ist zwar nicht optimal, hat aber den Vorteil, dass nicht explizit ein Codebuch erzeugt werden muss.
- Allerdings nutzen weder Huffman-Codierung noch arithmetische Codierung die Abhängigkeiten von Symbolen unteinander nicht aus; dies geschieht am besten durch andere, mehr oder weniger applikationsspezifische Algorithmen.
- Für Texte wird häufig die Familie der Lempel-Ziv-Algorithmen verwendet, die aus den Daten ein Wörterbuch generieren, um häufig auftretende Symbolfolgen kompakt darzustellen.

13.3.2 Was wir nicht gemacht haben

- Applikationsspezifische Codierverfahren für nicht textähnliche Daten: Video, Audio, ...
- Burrows-Wheeler-Transformation: Sortiert textähnliche Daten so um, dass sich Ketten gleicher Symbole bilden ($\rightarrow$ anschließende Lauflängencodierung), sich die Originalreihenfolge aber mit minimaler Nebeninformation wiederherstellen lässt.
- Verlustbehaftet Codierung: Minimale Störung bei gegebener maximaler Datenrate oder minimale Datenrate bei gegebener maximaler Störung.

13.4 Kanalcodierung

13.4.1 Was wir gemacht haben

- Aufgabe: Einfügen von Redundanz, die es erlaubt, bei einer Übertragung aufgetretene Fehler zu erkennen oder sogar zu korrigieren.
- Wichtige Kenngröße eines Codes: Der minimale Abstand d zwischen zwei Codewörtern. Es können t Fehler erkannt werden, wenn $t < d$, und t Fehler korrigiert werden, wenn $2t < d$.
- Shannon: Die Kanalkapazität ist eine obere Schranke für die Coderate, wenn mit verschwindend geringer Restfehlerwahrscheinlichkeit übertragen werden soll. Für große Codewortlängen erreichen zufällige Codes diese Schranke.
- Praktisch sind zufällige Codes mit großen Codewortlängen nicht handhabbar; die Codes müssen eine Struktur bekommen.
- Lineare Codes ($\mathcal{C}(\boldsymbol{a}_1 + \boldsymbol{a}_2) = \mathcal{C}(\boldsymbol{a}_1) + \mathcal{C}(\boldsymbol{a}_2)$): Codierung und Fehlererkennung durch Matrix-Multiplikation, Decodierung mit Syndromtabelle (speicheraufwändig) oder mit Belief Propagation (für dünn besetzte Prüfmatrizen).
- Zyklische Codes (jede zyklische Verschiebung eines Codeworts wieder ein Codewort): Codierung und Fehlererkennung durch Polynommultiplikation (oder Division), Decodierung im Allgemeinen mit Syndromtabelle.

- Faltungscodes: Codierung durch (bitweise) Faltung, Decodierung mit Viterbi-Algorithmus (sehr effizient für niedrigen Grad der Generatorpolynome. Nachteil: empfindlich gegen Bündelfehler.
- Turbo-Codes: Paralleler Einsatz zweier Faltungscodes mit Interleaver; Empfindlichkeit gegen Bündelfehler reduziert durch gegenseitiges Helfen der Komponentendecoder.

Code	Länge n	Rate	Decodierung
Wiederholung	beliebig	$\frac{1}{n} \to 0$	Mehrheitsentscheidung, korrigiert $\lfloor (n-1)/2 \rfloor$ Fehler
Parity-Check	beliebig	$\frac{n-1}{n} \to 1$	erkennt einen Fehler, keine Korrektur
Hamming	$2^k - 1$	$\frac{2^k-1-k}{2^k-1} \to 1$	korrigiert einen Fehler mittels Syndrom
CRC	beliebig	$\to 1$	nur zur Fehlererkennung
BCH/RS	typ. < 1000	beliebig	Lösung eines Gleichungssystems, um Fehler zu korrigieren
LDPC	beliebig	typ. $\frac{1}{2}$	iterativ; erlaubt Annäherung an Shannon-Grenze
Turbo	beliebig	typ. $\frac{1}{3}$–$\frac{1}{2}$	iterativ; erlaubt Annäherung an Shannon-Grenze

13.4.2 Was wir nicht gemacht haben

- Weitere spezielle Codes (Golay, Reed-Muller, ...)
- Galois-Felder (theoretische Grundlage z.B. der BCH-Codes)
- Kombination von Codes
- Detailliertere Analyse von Faltungscodes
- Entwurf und Analyse von LDPC-Codes
- Entwurf und Analyse von Turbo-Codes
- Dirty-Paper-Coding: Codierung bei im Sender bekannter Störung