CC

Noiseless coding - entropy

1. DMC
2. BSC, BEC
3. information rate $\rho(C) = \frac{1}{n}log_2 m$; error rate = $e(\hat{C}) = \max_{x \in M} P(error|xsent)$
4. **transmit reliably** at rate R if there exists $(C_n)_{n=1}^{\infty}$ with each C_n a code of length n such that $lim_{n\to\infty}\rho(C_n) = R$ & $lim_{n\to\infty}\hat{e}(C_n) = 0$.
5. A code is a function $c : A \to B^*$, $c(a)$ are codewords; $c^* : A^* \to B^*$
6. decipherable: induced map c^* is injective
7. block code: all words same length; comma code;
8. prefix-free code: is a code where no codeword is a prefix of any other distinct word
9. **Kraft's inequality**: $|A| = m, |B| = a$, $c : A \to B^*$ has word lengths $l_1, \dots, l_m$. Then $\sum_{i=1}^{m} a^{-l_i} \leq 1$
10. A prefix-free code exists if and only if Kraft's inequality holds
11. (McMillan). Any decipherable code satisfies Kraft's inequality
12. Cor: A decipherable code with prescribed word lengths exists if and only if a prefix-free code with the same word lengths exists

1. $H(X) = -\sum_{i=1}^{b} p_i \log p_i$
2. note: $H(p)' = \log \frac{1-p}{p}$, $p = \frac{1}{2}$ giving entropy 1
3. **Gibb's inequality**: $-\sum_{i=1}^{n} p_i \log p_i \leq -\sum_{i=1}^{n} p_i \log q_i$. [hint: $\ln q_i/p_i \leq q_i/p_i - 1$]
4. Cor: $H(p_1, p_2, \dots, p_n) \leq \log n$
5. **Shannon's Noiseless Coding Theorem**: $H(X)/\log a \leq E[S] < H(X)/\log a + 1$ [left: Gibb's, $q_i = a^{-l_i}/D$; right: $l_i = lower[-\log_a p_i] + 1$]
6. Shannon-Fano Coding
7. Huffman Coding is optimal [lemma: $p_i p_j, l_i l_j$, ; maximal length differ only one last]
8. $H(X, Y)$
9. $H(X,Y) \leq H(X) + H(Y)$ [Gibb's, p_{ij} replace by $p_i q_j$]

Error correcting codes - noisy channels

1. binary $[m, n]$-code, Hamming distance
2. ideal observer, maximal likelihood(maximising $P(x$ received | c sent)), munimum distance [later two equivalent if $p < 1/2$]
3. d-error detecting: changing up to d digits in each codeword cannot produce another; e-error correcting if knowing that $x \in 0, 1^n$ differs fom a codeword in at most e places we can deduce the codeword.
4. Repetition Code: $[n, 2]$-code, info rate $1/n$
5. Simple parity check: $[n, 2^{n-1}]$, info rate $\frac{n-1}{n}$
6. Hamming code; [7,16,3]-code, 1-error-correcting
7. $[n, m, d]$-code. Minimum distance d, $(d-1)$-error-detecting, $[\frac{d-1}{2}]$-error-correcting

1. $V(n,r) = |B(x,r)| = \sum_{i=0}^{r} \binom{n}{i}$
2. **Hamming's bound:** e-error correcting code C of length n has $|C| \leq \frac{2^n}{V(n,e)}$ (as pairwise disjoint balls)
3. Perfect [example: Hamming, binary repetition]
4. $A(n, d+1) \leq A(n,d)$
5. $2^n/(V(n,d-1)) \leq |A(n,d)| \leq 2^n/(V(n,[(d-1)/2]))$
6. C^+, C^-, C'

Information theory - Shannon's theorems

1. Definition. We model n uses of a channel by the nth extension, with input alphabet A^n and output alphabet B^n. A code C of length n is a function $M \to A^n$ where M is the set of possible messages. Implicitly we also have a decoding rule $B^n \to M$. The size of C is $m = |M|$. The information rate is $\rho(C) = 1/n \log_2 m$. The error rate is $\hat{e(C)} = \max_{x \in M} P(error|xsent)$.
2. transmit reliably at rate R; capacity is the supremum of all reliable transmission rates.
3. $H(X|Y) = H(X,Y) - H(Y)$
4. $H(X|Y) \leq H(X)$
5. Let X, Y, Z be random variables. Then $H(X|Y) \leq H(X|Y,Z) + H(Z)$
6. Fano's inequality: Let X, Y be random variables taking values in A, |A| = m. Let p = P(X != Y). Then $H(X|Y) \leq H(p) + plog(m-1)$ [hint: define $Z = 0$ if $X = Y$]
7. $I(X;Y) := H(X) - H(X|Y)$
8. (information) capacity is $\max_X I(X;Y)$
9. Shannon's Second Coding Theorem: For a DMC, the operational capacity is equal to the information capacity
10. The nth extension of a DMC with information capacity C has information capacity nC

Linear Codes

1. Def of linear codes
2. rank of linear codes: dimension as a F_2-vector spac
3. (n,k,d)-code: length n, rank k, mim distance d – is $[n, 2^k]$ code.
4. weight of $x \in F_2^n$ is $w(x) = d(x,0)$
5. The minimum distance of a linear code is the minimum weight of a non-zero code word
6. parity check code of P is$C = x \in F_2^n : p \cdot x = 0, \forall p \in P$
7. $P = (1, \ldots, 1)$ gives the simple parity check code; $P = (1,0,1,0,1,0,1), (0,1,1,0,0,1,1), (0,0,0,1,1,1,1)$ gives Hamming's $[7, 16, 3]$-code
8. Every parity check code is linear
9. Dual code: $C^\perp = \{x \in F_2^n : x \cdot y = 0, \forall y \in C\}$
10. $dim(C) + dim(C^\perp) = n$. So $(C^\perp)^\perp = C$
11. Let C be a (n,k)-code. A generator matrix G for C is a $k \times n$ matrix with rows a basis of C; A parity check matrix H for C is a generator matrix for $C^\perp$. It is a $(n-k) \times n$ matrix. The codewords of C can be viewed either as: • Linear combinations of rows of G; • Linear dependence relations between the columns of H