UNIT V · LESSON 07

Information Theory & Coding

information · entropy · source coding (Huffman, Shannon–Fano) · channel capacity · Shannon limit · linear block codes · cyclic codes · convolutional codes · Viterbi

20
questions in Unit V
Signals · DSP · Analog & Digital Comm

MEASURE INFORMATION & ENTROPY

- Self-information **$I(x) = \log_2(1/p)$ bits** (rare events carry more information)
- Units: bit ($\log_2$), nat (ln), decit/Hartley ($\log_{10}$); 1 nat = 1.443 bits
- **Entropy $H = -\Sigma\, p_i \log_2 p_i$** bits/symbol (average information)
- H is **maximum = $\log_2 M$** when all M symbols are equiprobable; H = 0 if one symbol is certain
- Binary source: $H(p) = -p \log p - (1 - p) \log(1 - p)$; max **1 bit** at p = 0.5
- **Information rate R = r H** bits/s (r symbols/s)
- Example: p = {½, ¼, ⅛, ⅛} → H = 0.5 + 0.5 + 0.375 + 0.375 = **1.75 bits/symbol**
- Extension of source: $H(X^n) = n\, H(X)$

CHANNELS JOINT & CONDITIONAL ENTROPY

- H(X, Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)
- **Mutual information I(X; Y) = H(X) − H(X|Y) = H(Y) − H(Y|X)** ≥ 0
- H(X|Y) = **equivocation** (information lost in the channel)
- Independent X, Y: I(X; Y) = 0; noiseless channel: I = H(X)

COMPRESSION SOURCE CODING

- Average length **$\bar{L} = \Sigma\, p_i\, l_i$**; efficiency **$\eta = H / \bar{L}$**; redundancy = 1 − η
- **Shannon's source coding theorem**: $H \le \bar{L} < H + 1$ (block coding → $\bar{L}/n \to H$)
- **Kraft inequality**: $\Sigma\, 2^{-l_i} \le 1$ for a prefix (instantaneous) code
- **Prefix code**: no codeword is a prefix of another → uniquely & instantaneously decodable
- **Huffman** (optimal, bottom-up merging of two least probable) vs **Shannon–Fano** (top-down splitting into equal-probability halves)
- Others: arithmetic coding, **Lempel–Ziv** (ZIP, GIF), run-length coding

WORKED HUFFMAN EXAMPLE

Symbol	p	Code	Length
A	0.4	0	1
B	0.2	10	2
C	0.2	111	3
D	0.1	1100	4
E	0.1	1101	4

- $\bar{L}$ = 0.4 + 0.4 + 0.6 + 0.4 + 0.4 = **2.2 bits/symbol**
- H = 0.529 + 2(0.464) + 2(0.332) = **2.122 bits** → η = **96.4 %**
- Fixed-length code would need 3 bits

Huffman codes are not unique (tie-breaking), but $\bar{L}$ is always the same minimum.

DMC CHANNEL MODELS

- **Discrete memoryless channel**: transition matrix P(y|x)
- **Binary symmetric channel (BSC)** with crossover p: **C = 1 − H(p)** bits/use
- p = 0.1 → C = 1 − 0.469 = **0.531**; p = 0.5 → C = 0 (useless channel)
- **Binary erasure channel** (erasure ε): C = 1 − ε
- Noiseless channel: C = $\log_2$ M
- **Channel capacity C = max I(X; Y)** over input distributions

BASICS ERROR-CONTROL CODING

- **FEC** (forward error correction) vs **ARQ** (retransmission)
- Code rate **r = k/n**; redundancy n − k
- **Hamming distance** d between codewords; **minimum distance d_{min}**
- Detect up to **s = d_{min} − 1** errors; correct up to **t = ⌊(d_{min} − 1)/2⌋**

AWGN CAPACITY SHANNON–HARTLEY

- **C = B $\log_2$(1 + S/N)** bits/s
- **Channel coding theorem**: if R ≤ C, arbitrarily small error probability is possible with suitable coding; if R > C, it is not
- Example: B = 3.1 kHz, S/N = 30 dB (1000) → C ≈ **30.9 kbps**
- As B → ∞: **C_∞ = 1.44 S/N_0**
- **Shannon limit**: E_b/N_0 ≥ ln 2 = **−1.6 dB**
- Bandwidth–SNR trade-off: wider bandwidth allows lower SNR

BASICS ERROR-CONTROL CODING (CONTD.)

- **Hamming bound**: $2^{n-k} \geq \Sigma_{i=0}^{t} C(n, i)$; equality → **perfect code**
- Coding gain: reduction in required E_b/N_0 for the same BER

(N, K) LINEAR BLOCK CODES

- Codeword **c = m G**; generator **G = [I_k | P]** (systematic)
- Parity-check matrix **H = [P^T | I_{n-k}]**, c H^T = 0
- **Syndrome s = r H^T** depends only on the error pattern → look-up table decoding
- d_{min} = minimum weight of non-zero codewords (linear codes)
- **Hamming codes**: n = 2^m − 1, k = n − m, **d_{min} = 3** (single-error correcting, perfect): (7,4), (15,11)
- Repetition code (n, 1): d_{min} = n
- Single parity check (k + 1, k): d_{min} = 2 (detects 1 error)

MEMORY CONVOLUTIONAL CODES

- Encoder with shift registers: rate **k/n**, **constraint length K** (memory + 1)
- Represented by generator sequences, **state diagram**, **trellis**, code tree
- Number of states = **2^{K-1}** (k = 1)
- **Free distance d_{free}** determines error correction
- Rate ½, K = 3, generators (111, 101) → d_{free} = **5**
- Decoding: **Viterbi** (maximum-likelihood, trellis, survivor paths) — hard / soft decision; sequential decoding (Fano)
- **Turbo codes** (parallel concatenated + interleaver, iterative) and **LDPC** codes approach the Shannon limit (3G/4G/5G, Wi-Fi)
- **Interleaving** spreads burst errors into random errors

POLYNOMIALS CYCLIC CODES

- Any cyclic shift of a codeword is a codeword
- **Generator g(x)** of degree n − k divides **x^n + 1**
- Systematic encoding: remainder of x^{n-k} m(x) ÷ g(x) = parity bits
- (7,4) Hamming cyclic code: g(x) = 1 + x + x^3
- Implemented with **shift registers** (LFSR); syndrome by division
- **CRC**, **BCH** (multiple random errors), **Reed–Solomon** (non-binary, **burst errors**: CDs, DVDs, QR codes; corrects (n − k)/2 symbols)
- Golay (23,12) perfect, t = 3

Q → A RAPID RECALL

Information of event with p = 1/8	3 bits
Entropy of 8 equiprobable symbols	3 bits
Max entropy of binary source	1 bit (p = ½)
Coding efficiency formula	H / $\bar{L}$
Optimal prefix code	Huffman
Capacity of BSC with p = 0	1 bit/use
Shannon capacity formula	B $\log_2(1 + S/N)$
Shannon limit for E_b/N_0	−1.6 dB
Errors corrected with d_{min} = 7	3

Q → A RAPID RECALL (CONTD.)

d_{min} of Hamming code	3
Codes for burst errors on CDs	Reed–Solomon
Decoder for convolutional codes	Viterbi

TRAP EXAM TRAPS

- More probable symbols get **shorter** codewords
- Capacity rises only **logarithmically** with SNR but linearly with bandwidth (until noise grows)
- Detection capability $d_{min} - 1 \neq$ correction capability $\lfloor (d_{min} - 1)/2 \rfloor$
- Kraft inequality is necessary for prefix codes; it does not make every code with those lengths prefix-free
- BSC with $p = 1$ has capacity 1 (just invert the bits)