Variable blocklength communication with feedback

Gauri Joshi
Graduate Seminar in Area 1

EECS MIT

9th Nov 2010

Motivation

- Shannon - feedback does not improve capacity
- Can help in the non-asymptotic regime

Length required to achieve 90% capacity on a $C = 1/2$ BSC

- Fixed length with feedback, $l > 3100$ bits
- Variable length with feedback, $l < 200$ bits

Even a simple termination signal sent by the source indicating end of transmission helps a lot

System Model

- Discrete Memoryless Channel

$$P_{Y_i|X_1^i, Y_1^{i-1}} = P_{Y_i|X_i} = P_{Y_1|X_1}$$

- Input and Output Alphabet $\mathcal{A}$ and $\mathcal{B}$
- Transition matrix $\mathcal{P}$ - $p_{i,j}$ is the probability of transmitting i^{th} input symbol and receiving j^{th} output symbol
- $W \in 1, 2, ..M$ equiprobable message to be transmitted - mapped to input alphabet $\mathcal{A}$

Part I

Feedback in non-asymptotic regime

Fixed Blocklength without feedback

An (l, M, ϵ) code,

- Encoder $X_n = f(W)$, Decoder $\hat{W} = g(Y^l)$
- The fundamental limit of coding is,

$$M^*(l, \epsilon) = \max\{M : \exists (l, M, \epsilon) \text{ code}\}$$

- Maximum information we can send is,

$$\log M^*(l, \epsilon) = lC - \sqrt{lV} Q^{-1}(\epsilon) + O(\log l)$$

- V - Channel dispersion - measures the stochastic variability of the channel as compared to a deterministic channel of the same capacity.
- In presence of variable-length coding with feedback the $\frac{1}{\sqrt{l}}$ penalty term is eliminated

Fixed Blocklength with feedback

An (l, M, ϵ) code

- Noiseless Feedback of Y's to the encoder
- Encoder $X_n = f(W, Y^{n-1})$
- Feedback does not help remove the penalty term

$$\log M_b^*(l, \epsilon) = nC - \sqrt{nV}Q^{-1}(\epsilon) + O(\log n)$$

- $\log M_b^*$ increases hardly by 2-3 bits as compared to $\log M^*$.

Variable Blocklength without feedback

- Allow a non-vanishing probability of error ϵ
- The capacity increases to give ϵ-capacity

Theorem

For any non-anticipatory channel with capacity C that satisfies the strong converse for fixed-blocklength codes (without feedback), the ϵ-capacity under variable-length coding without feedback, is

$$C_\epsilon = \frac{C}{1 - \epsilon}$$

Variable blocklength with feedback

- Encoder $X_n = f(W, Y^{n-1})$
- Decoder $\hat{W} = g_\tau(Y^\tau)$
- Stopping time τ on $\sigma\{Y_1, ..Y_n\}$ such that $\mathbb{E}(\tau) \leq l$
- $Pr(\hat{W} \neq W) \leq \epsilon$

The fundamental limit of VLF coding is,

$$M_f^*(l, \epsilon) = \max M : \exists (l, M, \epsilon) - \text{VLF} \quad \text{code}$$

VLF codes with termination

- In VLF codes, the decoder decides the stopping time τ as a function of outputs $Y^{\tau-1}$. It is conveyed to source through feedback
- Source sends a termination signal to receiver on a separate reliable channel - VLFT code
- Stopping time τ on $\sigma\{W, Y_1, ..Y_n\}$ such that $\mathbb{E}(\tau) \leq l$

The fundamental limit of VLFT coding is,

$$M_t^*(l, \epsilon) = \max M : \exists (l, M, \epsilon) - \text{VLFT code}$$

Variable Blocklength with feedback

Theorem

For an arbitrary DMC with capacity C we have for any $0 < \epsilon < 1$

$$\log M_f^*(l,\epsilon) = \frac{lC}{1-\epsilon} + O(\log l) \quad (1)$$

$$\log M_t^*(l,\epsilon) = \frac{lC}{1-\epsilon} + O(\log l) \quad (2)$$

More precisely, we have,

$$\frac{lC}{1-\epsilon} - \log l + O(1) \leq \log M_f^*(l,\epsilon) \leq \frac{lC}{1-\epsilon} + O(1) \quad (3)$$

$$\log M_f^*(l,\epsilon) \leq \log M_t^*(l,\epsilon) \leq \frac{lC + \log l}{1-\epsilon} + O(1) \quad (4)$$

Proof of converse of the theorem

Theorem

Consider an arbitrary DMC with capacity C. Then any (l, M, ϵ) VLF code with $0 \leq \epsilon \leq 1$ satisfies

$$\log M \leq \frac{Cl + h(\epsilon)}{1 - \epsilon},$$

whereas each (l, M, ϵ) VLFT code with $0 \leq \epsilon \leq 1$ satisfies

$$\log M \leq \frac{Cl + h(\epsilon) + (l+1)h(\frac{1}{l+1})}{1 - \epsilon} \leq \frac{Cl + \log(l+1) + h(\epsilon) + log(\epsilon)}{1 - \epsilon},$$

where $h(x) = -x \log x - (1-x)\log(1-x)$ is the binary entropy function.

Proof of converse of the theorem

By Fano's inequality we have,

$$\begin{aligned}(1-\epsilon)\log M &\leq I(W;Y^{\tau},\tau)+h(\epsilon)\\ &= I(W;Y^{\tau})+I(W;\tau|Y^{\tau})+h(\epsilon)\\ &\leq I(W;Y^{\tau})+H(\tau)+h(\epsilon)\\ &\leq I(W;Y^{\tau})+(l+1)h\left(\frac{1}{l+1}\right)+h(\epsilon),\end{aligned}$$

where, we upper bound $H(\tau)$ by solving the optimization problem:

$$\max_{\tau:\mathbb{E}[\tau]\leq l} H(\tau)=(l+1)h\left(\frac{1}{l+1}\right)$$

τ cannot convey more than $O(\log l)$ bits of information about the message

Proof of converse of the theorem

$$
\begin{aligned}
(1-\epsilon)\log M &\leq I(W;Y^{\tau}) + (l+1)h\left(\frac{1}{l+1}\right) + h(\epsilon) \\
&\leq Cl + (l+1)h\left(\frac{1}{l+1}\right) + h(\epsilon),
\end{aligned}
$$

We use the result from Burnashev which says that,

$$
I(W;Y^{\tau}) \leq C\mathbb{E}[\tau] \leq Cl.
$$

Concluding Remarks

- Variable length coding with feedback drastically reduces the average blocklength required to achieve a given probability of error by removing the $\frac{1}{\sqrt{l}}$ penalty term.
- Even simple decision-feedback codes with just the termination signal have performance very close to the VLFT codes

Part II

Optimal Error Exponents

Decision making method

- Find the posterior probability $p_j(n)$ of the j^{th} input symbol after n observations
- Calculate the likelihood functions $\log \frac{p_j(x_n)}{1-p_j(x_n)}$.
- Make a decision in favor of symbol X_j if the likelihood crosses $\log(1/\epsilon)$
- Probability of error,

$$P_e = \frac{1}{M} \sum_{j=1}^{M} \big(1 - p_j(x_n)\big) \leq \frac{\epsilon}{1+\epsilon} \leq \epsilon$$

Entropy of the posterior distribution

- Entropy of the posterior distribution $p(n)$ is defined as H_n

$$\mathbb{E}(H_n - H_{n+1}) \leq C$$
$$\mathbb{E}(\log H_n - \log H_{n+1}) \leq C_1$$

where $C_1 > C$ is the maximal relative entropy between output distributions.

$$C_1 = \max_{i,k} \sum_{l=1}^{K} p_{i,l} \log \frac{p_{i,l}}{p_{k,l}} = \max_{i,k} D(\mathbf{p}_i || \mathbf{p}_k)$$

Burnashev's Error exponent

- We know that without feedback, the error exponent is $E(R) = (C - R)$. i.e the probability of error with blocklength l is, $P_e \leq e^{-El}$
- With variable length and feedback we get the error exponent,

$$E(R) = C_1 \left(1 - \frac{R}{C}\right)$$

Yamamoto Itoh scheme

- Simple two-phase coding scheme that achieves this error exponent
- Phase 1 - Transmit message for γN symbols
- Phase 2 - Transmit correct/error signal for $n = (1 - \gamma)N$ symbols
- If in error, retransmit the message in the next block
- Probability of error $= P_E = P_{1e} P_{ce}$