香港中文大學
The Chinese University of Hong Kong

7.6 Feedback Capacity

Feedback

Feedback

- Feedback is common in practical communication systems for correcting possible errors which occur during transmission.

Feedback

- Feedback is common in practical communication systems for correcting possible errors which occur during transmission.
- Daily examples: phone conversation, classroom teaching.

Feedback

- Feedback is common in practical communication systems for correcting possible errors which occur during transmission.

- Daily examples: phone conversation, classroom teaching.

- Data communication: the receiver may request a packet to be retransmitted if the *parity check* bits received are incorrect (Automatic Repeat-reQuest).

Feedback

- Feedback is common in practical communication systems for correcting possible errors which occur during transmission.
- Daily examples: phone conversation, classroom teaching.
- Data communication: the receiver may request a packet to be retransmitted if the *parity check* bits received are incorrect (Automatic Repeat-reQuest).
- The transmitter can at any time decide what to transmit next based on the feedback so far.

Feedback

- Feedback is common in practical communication systems for correcting possible errors which occur during transmission.
- Daily examples: phone conversation, classroom teaching.
- Data communication: the receiver may request a packet to be retransmitted if the *parity check* bits received are incorrect (Automatic Repeat-reQuest).
- The transmitter can at any time decide what to transmit next based on the feedback so far.
- Can feedback increase the channel capacity?

Feedback

- Feedback is common in practical communication systems for correcting possible errors which occur during transmission.
- Daily examples: phone conversation, classroom teaching.
- Data communication: the receiver may request a packet to be retransmitted if the *parity check* bits received are incorrect (Automatic Repeat-reQuest).
- The transmitter can at any time decide what to transmit next based on the feedback so far.
- Can feedback increase the channel capacity?
- Not for DMC, even with complete feedback!

Definition 7.18 An (n, M) code with complete feedback for a discrete memoryless channel with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ is defined by encoding functions

$$f_i : \{1, 2, \cdots, M\} \times \mathcal{Y}^{i-1} \to \mathcal{X}$$

for $1 \leq i \leq n$ and a decoding function

$$g : \mathcal{Y}^n \to \{1, 2, \cdots, M\}.$$

Definition 7.18 An (n, M) code with complete feedback for a discrete memoryless channel with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ is defined by encoding functions

$$f_i : \{1, 2, \cdots, M\} \times \mathcal{Y}^{i-1} \to \mathcal{X}$$

for $1 \leq i \leq n$ and a decoding function

$$g : \mathcal{Y}^n \to \{1, 2, \cdots, M\}.$$

Notations: $\mathbf{Y}^i = (Y_1, Y_2, \cdots, Y_i)$, $X_i = f_i(W, \mathbf{Y}^{i-1})$

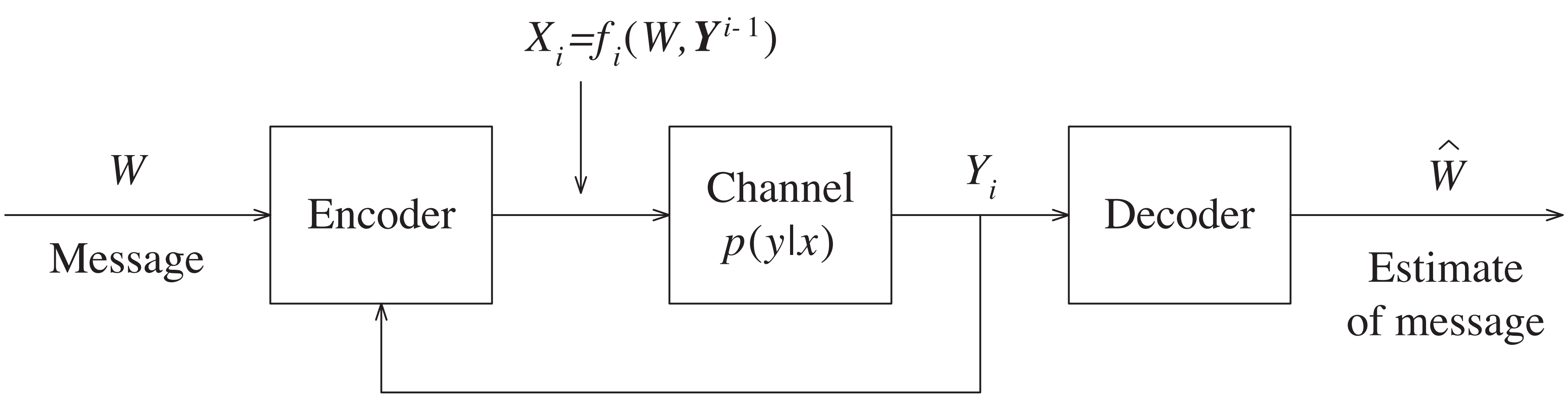
Definition 7.18 An (n, M) code with complete feedback for a discrete memoryless channel with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ is defined by encoding functions

$$f_i : \{1, 2, \cdots, M\} \times \mathcal{Y}^{i-1} \rightarrow \mathcal{X}$$

for $1 \leq i \leq n$ and a decoding function

$$g : \mathcal{Y}^n \rightarrow \{1, 2, \cdots, M\}.$$

Notations: $\mathbf{Y}^i = (Y_1, Y_2, \cdots, Y_i)$, $X_i = f_i(W, \mathbf{Y}^{i-1})$



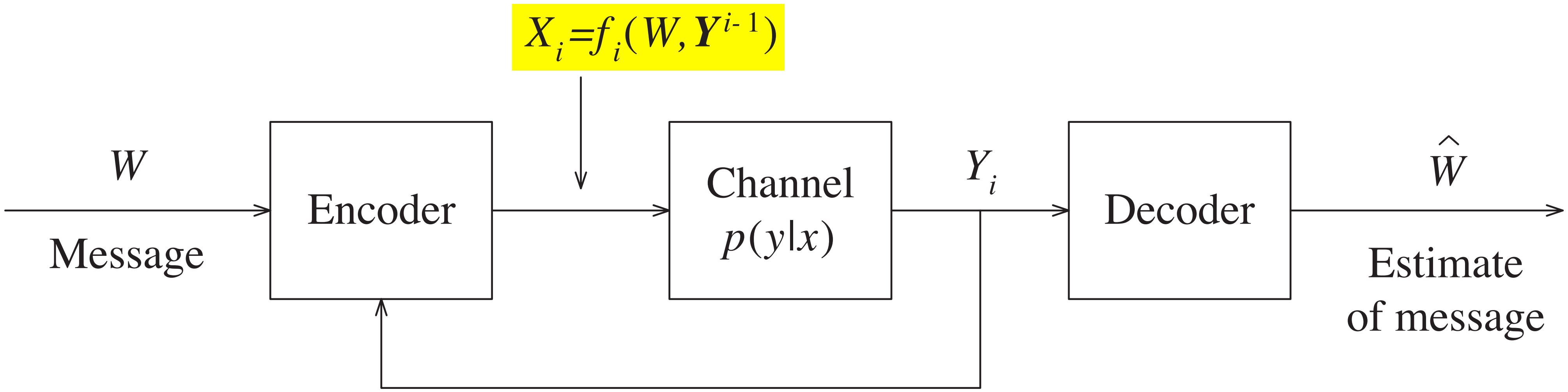
Definition 7.18 An (n, M) code with complete feedback for a discrete memoryless channel with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ is defined by encoding functions

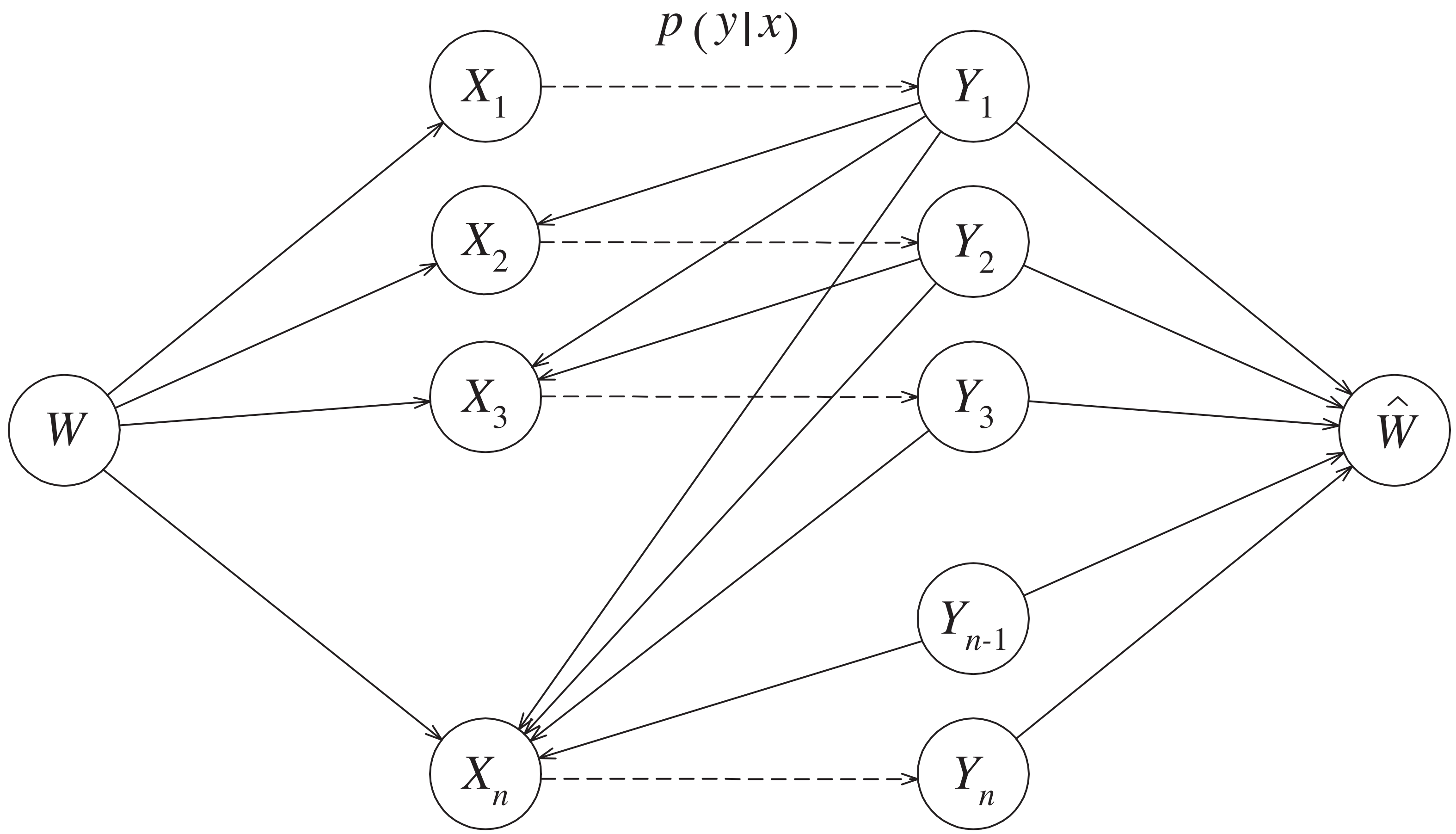
$$f_i : \{1, 2, \cdots, M\} \times \mathcal{Y}^{i-1} \rightarrow \mathcal{X}$$

for $1 \leq i \leq n$ and a decoding function

$$g : \mathcal{Y}^n \rightarrow \{1, 2, \cdots, M\}.$$

Notations: $\mathbf{Y}^i = (Y_1, Y_2, \cdots, Y_i)$, $X_i = f_i(W, \mathbf{Y}^{i-1})$





- The above is the dependency graph for a channel code with feedback, which is equivalent to

$$q(w, \mathbf{x}, \mathbf{y}, \hat{w}) = q(w) \left(\prod_{i=1}^{n} q(x_i | w, \mathbf{y}^{i-1}) \right) \left(\prod_{i=1}^{n} p(y_i | x_i) \right) q(\hat{w} | \mathbf{y})$$

  for all $(w, \mathbf{x}, \mathbf{y}, \hat{w}) \in \mathcal{W} \times \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{W}$ such that $q(w, \mathbf{y}^{i-1}), q(x_i) > 0$ for $1 \leq i \leq n$ and $q(\mathbf{y}) > 0$, where $\mathbf{y}^i = (y_1, y_2, \cdots, y_i)$.

Definition 7.19 A rate R is achievable with complete feedback for a discrete memoryless channel $p(y|x)$ if for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code with complete feedback such that

$$\frac{1}{n} \log M > R - \epsilon$$

and

$$\lambda_{max} < \epsilon.$$

Definition 7.19 A rate R is achievable with complete feedback for a discrete memoryless channel $p(y|x)$ if for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code with complete feedback such that

$$\frac{1}{n} \log M > R - \epsilon$$

and

$$\lambda_{max} < \epsilon.$$

Definition 7.20 The feedback capacity, C_{FB}, of a discrete memoryless channel is the supremum of all the rates achievable by codes with complete feedback.

Definition 7.19 A rate R is achievable with complete feedback for a discrete memoryless channel $p(y|x)$ if for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code with complete feedback such that

$$\frac{1}{n} \log M > R - \epsilon$$

and

$$\lambda_{max} < \epsilon.$$

Definition 7.20 The feedback capacity, C_{FB}, of a discrete memoryless channel is the supremum of all the rates achievable by codes with complete feedback.

Proposition 7.21 The supremum in the definition of C_{FB} in Definition 7.20 is the maximum.

Proof Follows from Definition 7.19. See textbook for details.

Definition 7.19 A rate R is achievable with complete feedback for a discrete memoryless channel $p(y|x)$ if for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code with complete feedback such that

$$\frac{1}{n} \log M > R - \epsilon$$

and

$$\lambda_{max} < \epsilon.$$

Definition 7.20 The feedback capacity, C_{FB}, of a discrete memoryless channel is the supremum of all the rates achievable by codes with complete feedback.

Proposition 7.21 The supremum in the definition of C_{FB} in Definition 7.20 is the maximum.

Proof Follows from Definition 7.19. See textbook for details.

Remark Since a channel code without feedback is a special case of a channel code with complete feedback, $C_{FB} \geq C$.

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof

1. The Markov chain

$$(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

   holds because the channel is memoryless.

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof

1. The Markov chain

$$(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

holds because the channel is memoryless.

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \rightarrow X_i \rightarrow Y_i$$

forms a Markov chain.

Proof

1. The Markov chain

$$(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}) \rightarrow X_i \rightarrow Y_i$$

holds because the channel is memoryless.

Lemma 7.22 For all $1 \le i \le n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof

1. The Markov chain

$$(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

   holds because the channel is memoryless.

Lemma 7.22 For all $1 \le i \le n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof

1. The Markov chain

$$(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

holds because the channel is memoryless.

2. Then

$$0 = I(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}; Y_i | X_i)$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof

1. The Markov chain

$$(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

   holds because the channel is memoryless.

2. Then

$$0 = I(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}; Y_i | X_i)$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof

1. The Markov chain

$$(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

   holds because the channel is memoryless.

2. Then

$$0 = I(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}; Y_i | X_i)$$

Lemma 7.22 For all $1 \le i \le n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof

1. The Markov chain

$$(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

   holds because the channel is memoryless.

2. Then

$$\begin{aligned} 0 &= I(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}; Y_i | X_i) \\ &= I(W, \mathbf{Y}^{i-1}; Y_i | X_i) \end{aligned}$$

Lemma 7.22 For all $1 \le i \le n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof

1. The Markov chain

$$(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

holds because the channel is memoryless.

2. Then

$$\begin{aligned} 0 &= I(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}; Y_i | X_i) \\ &= I(W, \mathbf{Y}^{i-1}; Y_i | X_i) \end{aligned}$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \rightarrow X_i \rightarrow Y_i$$

forms a Markov chain.

Proof

1. The Markov chain

   $$(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}) \rightarrow X_i \rightarrow Y_i$$

   holds because the channel is memoryless.

2. Then

$$\begin{aligned} 0 &= I(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}; Y_i | X_i) \\ &= I(W, \mathbf{Y}^{i-1}; Y_i | X_i) + I(\mathbf{X}^{i-1}; Y_i | W, X_i, \mathbf{Y}^{i-1}) \end{aligned}$$

Lemma 7.22 For all $1 \le i \le n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof

1. The Markov chain

$$(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

holds because the channel is memoryless.

2. Then

$$\begin{aligned} 0 &= I(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}; Y_i | X_i) \\ &= I(W, \mathbf{Y}^{i-1}; Y_i | X_i) + I(\mathbf{X}^{i-1}; Y_i | W, X_i, \mathbf{Y}^{i-1}) \end{aligned}$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof

1. The Markov chain

$$(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

holds because the channel is memoryless.

2. Then

$$\begin{aligned} 0 &= I(\underline{W}, \underline{\mathbf{X}^{i-1}}, \underline{\mathbf{Y}^{i-1}}; Y_i | X_i) \\ &= I(\underline{W, \mathbf{Y}^{i-1}}; Y_i | X_i) + I(\underline{\mathbf{X}^{i-1}}; Y_i | \underline{W}, X_i, \underline{\mathbf{Y}^{i-1}}) \end{aligned}$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof

1. The Markov chain

$$(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

holds because the channel is memoryless.

2. Then

$$\begin{aligned} 0 &= I(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}; Y_i | X_i) \\ &= I(W, \mathbf{Y}^{i-1}; Y_i | X_i) + I(\mathbf{X}^{i-1}; Y_i | W, X_i, \mathbf{Y}^{i-1}) \end{aligned}$$

which implies $I(W, \mathbf{Y}^{i-1}; Y_i | X_i) = 0$, or $(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$.

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \rightarrow X_i \rightarrow Y_i$$

forms a Markov chain.

Proof

1. The Markov chain

$$(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}) \rightarrow X_i \rightarrow Y_i$$

   holds because the channel is memoryless.

2. Then

$$\begin{aligned} 0 &= I(W, \mathbf{X}^{i-1}, \mathbf{Y}^{i-1}; Y_i | X_i) \\ &= I(W, \mathbf{Y}^{i-1}; Y_i | X_i) + I(\mathbf{X}^{i-1}; Y_i | W, X_i, \mathbf{Y}^{i-1}) \end{aligned}$$

   which implies $I(W, \mathbf{Y}^{i-1}; Y_i | X_i) = 0$, or $(W, \mathbf{Y}^{i-1}) \rightarrow X_i \rightarrow Y_i$.

Proof for $C_{FB} \leq C$

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W; \mathbf{Y}) + H(W|\mathbf{Y}).$$

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W; \mathbf{Y}) + H(W|\mathbf{Y}).$$

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W; \mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$I(W; \mathbf{Y}) = H(\mathbf{Y}) - H(\mathbf{Y}|W)$$

Proof for $C_{FB} \le C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned} I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \end{aligned}$$

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned}
I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, \underline{X_i})
\end{aligned}$$

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned} I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\underline{\mathbf{Y}^{i-1}, W}, \underline{X_i}) \end{aligned}$$

Proof for $C_{FB} \le C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W; \mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned} I(W; \mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \end{aligned}$$

Lemma 7.22 For all $1 \le i \le n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned} I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\cancel{\mathbf{Y}^{i-1}}, W, X_i) \end{aligned}$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned}
I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i)
\end{aligned}$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \le C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned}
I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\le \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i)
\end{aligned}$$

Lemma 7.22 For all $1 \le i \le n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned}
I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\leq \sum_{i=1}^{n} \underline{H(Y_i)} - \sum_{i=1}^{n} H(Y_i|X_i)
\end{aligned}$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned}
I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\leq \sum_{i=1}^{n} \underline{H(Y_i)} - \sum_{i=1}^{n} \underline{H(Y_i|X_i)}
\end{aligned}$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \le C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned}
I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\le \sum_{i=1}^{n} \underline{H(Y_i)} - \sum_{i=1}^{n} \underline{H(Y_i|X_i)} \\
&= \sum_{i=1}^{n} \underline{I(X_i;Y_i)}
\end{aligned}$$

Lemma 7.22 For all $1 \le i \le n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned}
I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&= \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq nC.
\end{aligned}$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \le C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned}
I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\le \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&= \sum_{i=1}^{n} I(X_i;Y_i) \\
&\le nC.
\end{aligned}$$

Lemma 7.22 For all $1 \le i \le n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned}
I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&= \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq nC.
\end{aligned}$$

4. Second,

$$H(W|\mathbf{Y}) = H(W|\mathbf{Y}, \hat{W}) \leq H(W|\hat{W}) \approx 0.$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \rightarrow X_i \rightarrow Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned}
I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&= \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq nC.
\end{aligned}$$

4. Second,

$$H(W|\mathbf{Y}) = H(W|\mathbf{Y}, \underline{\hat{W}}) \leq H(W|\hat{W}) \approx 0.$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned}
I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&= \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq nC.
\end{aligned}$$

4. Second,

$$H(W|\mathbf{Y}) = H(W|\mathbf{Y},\hat{W}) \leq H(W|\hat{W}) \approx 0.$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned} I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\ &\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\ &= \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq nC. \end{aligned}$$

4. Second,

$$H(W|\mathbf{Y}) = H(W|\cancel{\mathbf{Y}}, \hat{W}) \leq H(W|\hat{W}) \approx 0.$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned}
I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&= \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq nC.
\end{aligned}$$

4. Second,

$$H(W|\mathbf{Y}) = H(W|\mathbf{Y}, \hat{W}) \leq H(W|\hat{W}) \approx 0.$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned} I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\ &\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\ &= \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq nC. \end{aligned}$$

4. Second,

$$H(W|\mathbf{Y}) = H(W|\mathbf{Y}, \hat{W}) \leq H(W|\hat{W}) \approx 0.$$

Then we have

$$\log M = I(W;\mathbf{Y}) + H(W|\mathbf{Y}) \leq nC.$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \rightarrow X_i \rightarrow Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned}
I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&= \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq nC.
\end{aligned}$$

4. Second,

$$H(W|\mathbf{Y}) = H(W|\mathbf{Y}, \hat{W}) \leq H(W|\hat{W}) \approx 0.$$

Then we have

$$\log M = I(W;\mathbf{Y}) + H(W|\mathbf{Y}) \leq nC.$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned} I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\ &\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\ &= \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq nC. \end{aligned}$$

4. Second,

$$H(W|\mathbf{Y}) = H(W|\mathbf{Y}, \hat{W}) \leq H(W|\hat{W}) \approx 0.$$

Then we have

$$\log M = I(W;\mathbf{Y}) + H(W|\mathbf{Y}) \leq nC.$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned}
I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&= \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq nC.
\end{aligned}$$

4. Second,

$$H(W|\mathbf{Y}) = H(W|\mathbf{Y}, \hat{W}) \leq H(W|\hat{W}) \approx 0.$$

Then we have

$$\log M = I(W;\mathbf{Y}) + H(W|\mathbf{Y}) \leq nC.$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned} I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\ &\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\ &= \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq nC. \end{aligned}$$

4. Second,

$$H(W|\mathbf{Y}) = H(W|\mathbf{Y}, \hat{W}) \leq H(W|\hat{W}) \approx 0.$$

Then we have

$$\log M = I(W;\mathbf{Y}) + H(W|\mathbf{Y}) \leq nC.$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \rightarrow X_i \rightarrow Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned}
I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&= \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq nC.
\end{aligned}$$

4. Second,

$$H(W|\mathbf{Y}) = H(W|\mathbf{Y}, \hat{W}) \leq H(W|\hat{W}) \approx 0.$$

Then we have

$$\log M = I(W;\mathbf{Y}) + H(W|\mathbf{Y}) \leq nC.$$

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \leq C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned} I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\ &= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\ &\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\ &= \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq nC. \end{aligned}$$

4. Second,

$$H(W|\mathbf{Y}) = H(W|\mathbf{Y}, \hat{W}) \leq H(W|\hat{W}) \approx 0.$$

Then we have

$$\log M = I(W;\mathbf{Y}) + H(W|\mathbf{Y}) \leq nC.$$

5. Formally, apply Fano's inequality to upper bound $H(W|\hat{W})$. Upon filling in the ϵ's and δ's as in the proof of the converse of the channel coding theorem, we conclude that

$$R \leq C$$

for any rate R achievable with complete feedback.

Lemma 7.22 For all $1 \leq i \leq n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Proof for $C_{FB} \le C$

1. Consider any code with complete feedback.

2. Consider

$$\log M = H(W) = I(W;\mathbf{Y}) + H(W|\mathbf{Y}).$$

3. First,

$$\begin{aligned}
I(W;\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|\mathbf{Y}^{i-1}, W, X_i) \\
&= H(\mathbf{Y}) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\le \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&= \sum_{i=1}^{n} I(X_i;Y_i) \\
&\le nC.
\end{aligned}$$

4. Second,

$$H(W|\mathbf{Y}) = H(W|\mathbf{Y},\hat{W}) \le H(W|\hat{W}) \approx 0.$$

Then we have

$$\log M = I(W;\mathbf{Y}) + H(W|\mathbf{Y}) \le nC.$$

5. Formally, apply Fano's inequality to upper bound $H(W|\hat{W})$. Upon filling in the ϵ's and δ's as in the proof of the converse of the channel coding theorem, we conclude that

$$R \le C$$

for any rate R achievable with complete feedback.

Lemma 7.22 For all $1 \le i \le n$,

$$(W, \mathbf{Y}^{i-1}) \to X_i \to Y_i$$

forms a Markov chain.

Remarks

Remarks

1. Although feedback does not increase the capacity of a DMC, the availability of feedback often makes coding much simpler. See Example 7.23.

Remarks

1. Although feedback does not increase the capacity of a DMC, the availability of feedback often makes coding much simpler. See Example 7.23.

2. In general, if the channel has memory, feedback can increase the capacity.