Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

A Markov Decision Approach to Feedback Channel Capacity

Sekhar Tatikonda

Abstract— **We introduce a general framework for treating channels with memory and feedback. First, we generalize Massey's concept of *directed information* [7] and use it to characterize the feedback capacity of general channels. Second, we present coding results for Markov channels. Third, a dynamic programming framework for computing the capacity of Markov channels with output feedback is presented.**

I. INTRODUCTION

This paper presents a general framework for proving coding theorems for channels with memory and feedback. The problem of optimal channel coding goes back to Shannon's original work [8]. The channel coding problem with feedback goes back to early work by Shannon, Dobrushin, and others [9], [3]. Because of increased demand for wireless communication and networked systems there is a renewed interest in this problem. Feedback can increase the capacity of a noisy channel, decrease the complexity of the encoder and decoder, and reduce latency.

Recently Verdú and Han presented a very general formulation of the channel coding problem without feedback [12]. Specifically they provided a coding theorem for finite alphabet channels with arbitrary memory. They worked directly with the information density and provided a Feinstein-like lemma for the converse result. Here we generalize that formulation to the case of channels with feedback. In this case we require the use of code-functions as opposed to codewords. A code-function maps a message and the channel feedback information into a channel input symbol.

We first convert the channel coding problem with feedback into a new channel coding problem without feedback. The channel inputs in this new channel are code-functions. Unfortunately the space of code-functions can be quite complicated to work with. We show that we can work directly with the original space of channel inputs by making explicit the relationship between code-function distributions and channel input distributions. This relationship allows us to convert a mutual information optimization problem over code-function distributions into a directed information optimization problem over channel input distributions. We then show that for Markov channels this latter optimization can be solved using a dynamic programming formulation. The concept of directed information was introduced by Massey [7].

Due to space limitations we do not present proofs here. Complete details can be found in [11].

S. Tatikonda is with the Department of Electrical Engineering, Yale University, New Haven, CT, 06520, USA sekhar.tatikonda@yale.edu

II. DIRECTED INFORMATION

Let A and $\{A_t\}$ be random elements in the finite set $\mathcal{A}$. Similarly let B and $\{B_t\}$ be random elements in the finite set $\mathcal{B}$. Let $\mathcal{A}^T$ and $\mathcal{B}^T$ represent the T-fold product spaces. We use "log" to represent logarithm base 2.

A *stochastic kernel* on $\mathcal{B}$ given $\mathcal{A}$ is a function $P_{B|A}(\cdot|\cdot)$ such that: (1) $P_{B|A}(\cdot|a)$ is a probability measure on $\mathcal{B}$ for each fixed $a \in \mathcal{A}$ and (2) $P_{B|A}(b|\cdot)$ is a measurable function on $\mathcal{A}$ for each fixed $b \in \mathcal{B}$. We will use the notation $P(dB \mid a)$ to represent the measure $P_{B|A}(\cdot \mid a)$. We use the notation $X - Y - Z$ to denote that the random elements X, Y, Z form a Markov chain.

Definition 2.1: We are given an ordered sequence of random variables $A_1, ..., A_N$ with joint measure P_{A^N}. Let $I = \{i_1, ..., i_K\} \subseteq \{1, ..., N\}$ where $1 \leq i_1 < i_2 < ... < i_K \leq N$. Let $I^c = \{1, ..., N\} \setminus I$. Let $A^I = (A_{i_1}, ..., A_{i_K})$. Define A^{I^c} similarly. Then the *directed stochastic kernel of* A^I *with respect to* A^{I^c} is

$$\vec{P}_{A^I|A^{I^c}}(dA^I \mid a^{I^c}) = \bigotimes_{k=1}^{K} P_{A_{i_k}|A^{i_k-1}}(dA_{i_k} \mid a^{i_k-1}).$$

For each a^{I^c} the directed stochastic kernel $\vec{P}_{A^I|A^{I^c}}(dA^I \mid a^{I^c})$ is a well defined measure. For $J \subset I$ it is generically true that $\sum_{a^{I \setminus J} \in \mathcal{A}^{I \setminus J}} \vec{P}_{A^I|A^{I^c}}(a^{I \setminus J},\ a^J \mid a^{I^c})$ $\neq \vec{P}_{A^J|A^{I^c}}(a^J \mid a^{I^c})$. For example, given $P(dA_1, dA_2, dA_3)$ with the obvious time ordering: $\sum_{a_1 \in \mathcal{A}} \vec{P}(a_1,\ a_3 \mid a_2)$ $= \sum_{a_1 \in \mathcal{A}} (P(a_3 \mid a_1, a_2) P(a_1))$. This does not equal $P(a_3 \mid a_2) = \vec{P}(a_3 \mid a_2)$ unless $A_1 - A_2 - A_3$ forms a Markov chain.

Definition 2.2: The *directed information* is defined as

$$I(A^I \to A^{I^c}) = D(P_{A^I, A^{I^c}} \mid \vec{P}_{A^I|A^{I^c}} P_{A^{I^c}}) \tag{1}$$

where $D(\cdot \mid \cdot)$ is the divergence and $P_{A^I, A^{I^c}}(dA^I, dA^{I^c}) = \vec{P}_{A^{I^c}|A^I}(dA^{I^c} \mid a^I) \otimes \vec{P}_{A^I|A^{I^c}}(dA^I \mid a^{I^c})$ and $\vec{P}_{A^I|A^{I^c}} P_{A^{I^c}}(dA^I, dA^{I^c}) = \vec{P}_{A^I|A^{I^c}}(dA^I \mid a^{I^c}) \otimes P_{A^{I^c}}(dA^{I^c})$.

We can recover Massey's definition of directed information [7] by applying definition 2.2 to $A^I = A^T$ and $A^{I^c} = B^T$ with time-ordering: $A_1, B_1, A_2, ...$. Then $I(A^T \to B^T)$ $= \sum_{t=1}^{T} I(A^t;\ B_t \mid B^{t-1})$. Unlike the chain rule for mutual information the superscript on A in the summation is "t" and not "T". From definition 2.2 one can easily show: $I(A^T \to B^T) =$

$$E\left[\log \frac{P_{A^T|B^T}(A^T \mid B^T)}{\vec{P}_{A^T|B^T}(A^T \mid B^T)}\right] = E\left[\log \frac{\vec{P}_{B^T|A^T}(B^T \mid A^T)}{P_{B^T}(B^T)}\right]$$

The first equality shows that the directed information is the ratio between the posterior distribution and a "causal" prior distribution.

Note that $I(A^T;B^T) = E\left[\log \frac{\vec{P}(B^T \mid A^T)\vec{P}(A^T \mid B^T)}{P(B^T)P(A^T)}\right] = I(A^T \to B^T) + I(B^T \to A^T)$. If there is no feedback then $A_t - A^{t-1} - B^{t-1}$ forms a Markov chain. Hence $I(B^T \to A^T) = \sum_{t=1}^T I(A_t; B^{t-1} \mid A^{t-1}) = 0$. There is no "information" flowing from the receiver to the transmitter. We can conclude that $I(A^T;B^T) \geq I(A^T \to B^T)$ with equality if and only if there is no feedback [7].

III. Channels with Feedback

Here we formulate the feedback channel coding problem. A *channel* is a family of stochastic kernels $\{P(dB_t \mid a^t, b^{t-1})\}_{t=1}^T$ where T may be infinity. These channels are nonanticipative because the conditioning includes only a^t, b^{t-1}. A *message set* is a set $\mathcal{W} = \{1, ..., M\}$. A *channel code-function* is a sequence of T deterministic measurable maps $\{f_t\}_{t=1}^T$ such that $f_t : \mathcal{B}^{t-1} \to \mathcal{A}$ which takes $b^{t-1} \mapsto a_t$. Let $f^T = \{f_t\}_{t=1}^T$. Denote the set of all code-functions by $\mathcal{F}_T = \{f^T : f^T \text{ is a code-function}\}$. A *channel code* or *encoder*, is a set of M channel code-functions denoted by $f^T[w],\ w \in \mathcal{W}$. For message w at time t with channel feedback b^{t-1} the channel encoder outputs $f_t[w](b^{t-1})$. A *channel code without feedback*, is a set of M channel codewords denoted by $a^T[w],\ w \in \mathcal{W}$. For message w at time t the channel encoder outputs $a_t[w]$ independent of the past channel outputs b^{t-1}. A *channel decoder* is a map $g : \mathcal{B}^T \to \mathcal{W}$ taking $b^T \mapsto w$. The decoder waits till it observes all the channel outputs before reconstructing the input message.

To compute the different "information" measures we need to determine the joint measure: $P_{A^T,B^T}(dA^T, dB^T) = \bigotimes_{t=1}^T P_{A_t|A^{t-1},B^{t-1}}(dA_t|a^{t-1}, b^{t-1}) \otimes P_{B_t|A^t,B^{t-1}}(dB_t|a^t, b^{t-1})$. To compute the joint measure we need to specify the kernels $\{P(dA_t \mid a^{t-1}, b^{t-1})\}_{t=1}^T$. These kernels are determined by specifying an encoder.

A. Interconnection of Code-Functions to the Channel

Now we are ready to interconnect the pieces: channel, code-functions, encoder, and decoder. We follow Dobrushin's program and define a joint measure over the variables of interest that is consistent with the different components [4]. We will define a new channel without feedback that connects the code-functions to the channel outputs.

Let P_{F^T} be a distribution on $\mathcal{F}_T$. For example P_{F^T} may be a distribution that places mass $1/M$ on each of M different code-functions. Given a distribution on code-functions P_{F^T}, a channel $\{P(dB_t \mid a^t, b^{t-1})\}_{t=1}^T$, and the deterministic relations $a_t = f_t(b^{t-1})$ we need to construct a new channel that interconnects the random variables F^T to the random variables B^T. Call this channel $\{Q(dB_t \mid f^t, b^{t-1})\}_{t=1}^T$. We use "$Q$" to denote the new joint measure $Q(dF^T, dA^T, dB^T)$ that we will construct. The following three reasonable properties should hold for our new channel.

Definition 3.1: A measure $Q(dF^T, dA^T, dB^T)$ is *consistent* with a code-function distribution P_{F^T}, the relations $a_t = f_t(b^{t-1})$, and channel $\{P(dB_t \mid a^t, b^{t-1})\}_{t=1}^T$ if

(1) *There is no feedback to the code-functions in the new channel*: The measure on $\mathcal{F}_T$ is chosen at time 0. Thus it cannot causally depend on the B_t's. Specifically we require that $F_t - F^{t-1} - B^{t-1}$ be a Markov chain under Q. Thus $Q(dF_t \mid F^{t-1} = f^{t-1},\ B^{t-1} = b^{t-1}) = P(dF_t \mid f^{t-1}) \quad Q - a.s.$

(2) *The channel input is a function of the past outputs:* For each t, $A_t = F_t(B^{t-1})\ \ Q - a.s.$

(3) *The new channel preserves the properties of the underlying channel:* $Q(dB_t \mid F^t = f^t,\ A^t = a^t,\ B^{t-1} = b^{t-1}) = P(dB_t \mid a^t, b^{t-1}) \quad Q - a.s.$

Next we show there exists a unique *consistent* measure Q.

Lemma 3.1: Given P_{F^T}, the channel $\{P(dB_t \mid a^t, b^{t-1})\}_{t=1}^T$, and the relations $a_t = f_t(b^{t-1})$ there exists a unique consistent measure $Q(dF^T, dA^T, dB^T)$ on $\mathcal{F}_T \times \mathcal{A}^T \times \mathcal{B}^T$. Furthermore the channel from $\mathcal{F}_T$ to $\mathcal{B}^T$ for each $t = 1, ..., T$ is determined Q-a.s.:

$$Q(dB_t|F^t = f^t, B^{t-1} = b^{t-1}) = P(dB_t|f^t(b^{t-1}), b^{t-1})$$

A distribution P_W on $\mathcal{W}$ induces a measure P_{F^T} on $\mathcal{F}_T$:

Corollary 3.1: A distribution P_W on $\mathcal{W}$, a channel code $\{f^T[w]\}_{w=1}^M$, and channel $\{P(dB_t|a^t, b^{t-1})\}_{t=1}^T$ uniquely define a measure $Q(dW, dA^T, dB^T)$ on $\mathcal{W} \times \mathcal{A}^T \times \mathcal{B}^T$.

B. Channel Codes and Channel Capacity

Let the distribution P_W on the message set $\mathcal{W}$ be the uniform distribution.

Definition 3.2: A (T, M, ϵ) *channel code* over time horizon T consists of M code-functions, a channel decoder g, and an error probability satisfying: $\frac{1}{M}\sum_{w=1}^M \Pr(w \neq g(b^T) \mid w) \leq \epsilon$. A (T, M, ϵ) *channel code without feedback* over time horizon T consists of M codewords, a channel decoder g, and an error probability satisfying: $\frac{1}{M}\sum_{w=1}^M \Pr(w \neq g(b^T) \mid w) \leq \epsilon$.

Definition 3.3: R is an *ϵ-achievable rate* if, for every $\delta > 0$ there exists, for sufficiently large T, a (T, M, ϵ) channel code with rate $\frac{\log M}{T} > R - \delta$. The maximum $\epsilon - achievable$ rate is the called the *$\epsilon-$capacity* and denoted C_ϵ^{o}. The *operational channel capacity* is defined as the maximal rate that is achievable for all $0 < \epsilon < 1$ and is denoted C^{o}. Analogous definitions for $C_\epsilon^{\mathrm{o,nfb}}$ and $C^{\mathrm{o,nfb}}$ hold for the case without feedback.

The superscript "o" and "nfb" represent the words "operational" and "no feedback."

Definition 3.4: A *channel input distribution* is a sequence of kernels $\{P(dA_t \mid a^{t-1}, b^{t-1})\}_{t=1}^T$. A *channel input distribution without feedback* is a channel input distribution with the further condition that for each t the kernel $P(dA_t \mid a^{t-1}, b^{t-1})$ is independent of b^{t-1}. (Specifically $P(dA_t \mid a^{t-1}, b^{t-1}) = P(dA_t \mid a^{t-1}, \tilde{b}^{t-1})\ \ \forall b^{t-1}, \tilde{b}^{t-1}$.)

When computing the capacity of a channel it will turn out that we will need to know the convergence properties of the random variables $\frac{1}{T}\log\frac{P_{A^T,B^T}(A^T,B^T)}{\vec{P}_{A^T|B^T}P_{B^T}(A^T,B^T)}$. This is the *normalized information density* discussed in [12]. If there are reasonable regularity properties, like information stability, then these random variables will converge in probability to a deterministic limit. In the absence of any such structure we are forced to follow Verdú and Han's lead and define the following "floor" and "ceiling" limits [12].

The *limsup in probability* of a sequence of random variables $\{X_t\}$ is defined as the smallest extended real number α such that $\forall \epsilon > 0 \lim_{t\to\infty} \Pr[X_t \geq \alpha + \epsilon] = 0$. The *liminf in probability* of a sequence of random variables $\{X_t\}$ is defined as the largest extended real number α such that $\forall \epsilon > 0 \lim_{t\to\infty} \Pr[X_t \leq \alpha - \epsilon] = 0$.

Let $\vec{i}(a^T; b^T) = \log \frac{P_{A^T,B^T}(a^T,b^T)}{\vec{P}_{A^T|B^T} P_{B^T}(a^T,b^T)}$. For a sequence of joint measures $\{ P_{A^T,B^T} \}_{T=1}^{\infty}$ let $\underline{I}(A \to B) = \liminf_{\text{in prob}} \frac{1}{T}\vec{i}(A^T; B^T)$ and $\overline{I}(A \to B) = \limsup_{\text{in prob}} \frac{1}{T}\vec{i}(A^T; B^T)$.

Lemma 3.2: For a sequence of joint measures $\{P_{A^T,B^T}\}_{T=1}^{\infty}$ we have $\underline{I}(A \to B) \leq \liminf_{T\to\infty} \frac{1}{T} I(A^T \to B^T) \leq \limsup_{T\to\infty} \frac{1}{T} I(A^T \to B^T) \leq \overline{I}(A \to B)$.

Let $\mathcal{D}_T = \{\{P(dA_t \mid a^{t-1}, b^{t-1})\}_{t=1}^T\}$ be the set of all channel input distributions. Let $\mathcal{D}_T^{\text{nfb}} = \{P(dA_t \mid a^{t-1}, b^{t-1})\}_{t=1}^T\}$ be the set of all channel input distributions without feedback. We now define the mutual information optimization problems. For finite T let $C_T = \sup_{\mathcal{D}_T} \frac{1}{T} I(A^T \to B^T)$ and $C_T^{\text{nfb}} = \sup_{\mathcal{D}_T^{\text{nfb}}} \frac{1}{T} I(A^T \to B^T) = \sup_{\mathcal{D}_T^{\text{nfb}}} \frac{1}{T} I(A^T; B^T)$. For the infinite horizon case let $C = \sup_{\{\mathcal{D}_t\}_{t=1}^{\infty}} \underline{I}(A \to B)$ and $C^{\text{nfb}} = \sup_{\{\mathcal{D}_t^{\text{nfb}}\}_{t=1}^{\infty}} \underline{I}(A \to B) = \sup_{\{\mathcal{D}_t^{\text{nfb}}\}_{t=1}^{\infty}} \underline{I}(A; B)$.

Verdú and Han proved the following theorem for the case without feedback [12].

Theorem 3.1: For channels without feedback $C^{\text{o,nfb}} = C^{\text{nfb}}$.

In a certain sense we already have the solution to the coding problem for channels with feedback. Specifically lemma 3.1 tells us that the feedback channel problem is equivalent to a new channel coding problem without feedback. This new channel is from $\mathcal{F}_T$ to $\mathcal{B}^T$ and has channel kernels defined in lemma 3.1. Thus we can directly apply theorem 3.1 to this new channel.

This can be a very complicated problem to solve. We would have to optimize the mutual information over distributions on code functions. The directed information optimization problem can often be simpler. One reason is that we can work directly on the original $\mathcal{A}^T \times \mathcal{B}^T$ space and not on the $\mathcal{F}_T \times \mathcal{B}^T$ space. The second half of this paper describes a stochastic control approach to solving this optimization. In the next section, though, we present the feedback coding theorem.

IV. Coding Theorem for Channels with Feedback

Theorem 4.1: For channels with feedback $C^{\text{o}} = C$.

We first give a high-level summary of the issues involved. The converse part is straightforward. For any channel code and channel we know by lemma 3.1 that there exists a unique consistent measure $Q(dF^T, dA^T, dB^T)$. From this measure we can compute the induced channel input distribution $\{Q(dA_t \mid a^{t-1}, b^{t-1})\}_{t=1}^T$. Now $\{Q(dA_t \mid a^{t-1}, b^{t-1})\}_{t=1}^T \in \mathcal{D}_T$ but it need not be the supremizing channel input distribution. Thus the directed information under the induced channel input distribution may be less than the directed information under the supremizing channel input distribution. This is how we will show $C^{\text{o}} \leq C$.

For the direct part we take the optimizing channel input distribution $\{P(dA_t \mid a^{t-1}, b^{t-1})\}_{t=1}^T$ and construct a distribution on code-functions P_{F^T}. We then prove the direct part of the coding theorem for the channel from $\mathcal{F}_T$ to $\mathcal{B}^T$ by the usual techniques for channels without feedback. The next lemma shows that the directed information measures are the same for both the "$\mathcal{F}_T - \mathcal{B}^T$" channel and the "$\mathcal{A}^T - \mathcal{B}^T$" channel.

Lemma 4.1: For every consistent joint measure $Q(dF^T, dA^T, dB^T)$ the following holds

$$\frac{Q_{F^T,B^T}(F^T,B^T)}{Q_{F^T} Q_{B^T}(F^T,B^T)} = \frac{Q_{A^T,B^T}(A^T,B^T)}{\vec{Q}_{A^T|B^T} Q_{B^T}(A^T,B^T)} \quad Q-a.s.$$

hence $I(F^T; B^T) = I(A^T \to B^T)$ for each finite T. Furthermore, if given a sequence of consistent measures $\{Q(dF^T, dA^T, dB^T)\}_{T=1}^{\infty}$ we have $\underline{I}(F; B) = \underline{I}(A \to B)$.

Now we can prove the feedback channel coding theorem 4.1. We first prove the converse part. Then we prove the direct part.

a) Converse Theorem: Choose a (T, M, ϵ) channel code $\{f^T[w]\}_{w=1}^M$. Place a prior probability $\frac{1}{M}$ on each code-function $f^T[w]$. By corollary 3.1 this defines a consistent measure $Q(dW, dA^T, dB^T)$. The following is a generalization of the Verdú-Han converse [12].

Lemma 4.2: Every (T, M, ϵ) channel code satisfies $\epsilon \geq$

$$Q_{A^T,B^T}\left(\frac{1}{T}\log\frac{Q_{A^T,B^T}(A^T,B^T)}{\vec{Q}_{A^T|B^T} Q_{B^T}(A^T,B^T)} \leq \frac{1}{T}\log M - \gamma\right) - \ 2^{-\gamma T} \qquad \forall \gamma > 0$$

Theorem 4.2: The channel capacity $C^o \leq C$.

b) Direct Theorem: We will prove the direct theorem via a random coding argument. The following is a generalization of Feinstein's lemma [5].

Lemma 4.3: Fix a time T and $0 < \epsilon < 1$. Fix a channel $\{P(dB_t \mid b^{t-1}, a^t)\}_{t=1}^T$. Then for all $\gamma > 0$ and channel input distributions $\{P(dA_t \mid a^{t-1}, b^{t-1})\}_{t=1}^T$, there exists an (T, M, ϵ) channel code for the channel that satisfies $\epsilon \leq$

$$P_{A^T,B^T}\left(\frac{1}{T}\log\frac{P_{A^T,B^T}(A^T,B^T)}{\vec{P}_{A^T|B^T} P_{B^T}(A^T,B^T)} \leq \frac{1}{T}\log M\gamma\right) + \ 2^{-\gamma T}.$$

Theorem 4.3: All rates less than C are achievable.

By combining theorems 4.2 and 4.3 we can conclude theorem 4.1. Specifically C is the feedback channel capacity. It should be clear that if we restrict ourselves to channels without feedback then we recover the original coding theorem by Verdú and Han [12].

V. Markov Channels and the Canonical Markov Channel

In this section we formulate the Markov channel feedback capacity problem. We then provide a channel coding theorem. Finally we introduce the *canonical Markov channel.*

As before let $\mathcal{A}, \mathcal{B}$ be spaces with a finite number of elements representing the channel input and channel output, respectively. Furthermore let $\mathcal{S}$ be a finite state space. Let S_t, A_t, B_t be measurable random elements taking values in $\mathcal{S}, \mathcal{A}, \mathcal{B}$ respectively.

There is a natural time-ordering on the random variables of interest:

$$W, S_1, A_1, B_1, S_2, ..., S_T, A_T, B_T, \ \hat{W} \qquad (2)$$

First, at time 0 a message W is produced and the initial state S_1 drawn. The order of events in each of the T epochs is described in (2). At beginning of t-th epoch the channel input symbol A_t is placed on the channel by the transmitter, then B_t is observed by the receiver, then the state of the system evolves to S_{t+1}, and then finally the receiver feeds back information to the transmitter. At the beginning of the $t+1$ epoch the transmitter uses the feedback information to produce the next channel input symbol A_{t+1}. Finally at time T, after observing B_T, the decoder outputs the reconstructed message $\hat{W}$.

Definition 5.1: A *Markov channel* consists of an initial state distribution $P(dS_1)$, the state transition stochastic kernels $\{P(dS_{t+1} \mid s_t, a_t)\}_{t=1}^{T-1}$, and the channel output stochastic kernels $\{P(dB_t \mid s_t, a_t)\}_{t=1}^{T}$. If the stochastic kernel $P(dS_{t+1} \mid s_t, a_t)$ is independent of a_t for each $t = 1, ..., T-1$ then we say the channel is a *Markov channel without ISI* (intersymbol interference.)
Note that we are assuming the kernels $\{P(dS_{t+1} \mid s_t, a_t)\}$ and $\{P(dB_t \mid s_t, a_t)\}$ are stationary (independent of time.)

In order to compute any "information" measure we will need the joint measure:

$$P(dS^T, dA^T, dB^T) = \bigotimes_{t=1}^{T} (P(dB_t|s_t, a_t)$$
$$\otimes P(dA_t|s^t, a^{t-1}, b^{t-1}) \otimes P(dS_t|s_{t-1}, a_{t-1})).$$

To complete the description of the joint measure we will need to *interconnect* a channel input distribution, $\{P(dA_t \mid s^t, a^{t-1}, b^{t-1})\}$, to the channel.

We will restrict the channel input distribution for the Markov channel to be a sequence of stochastic kernels of the form: $\{P(dA_t \mid a^{t-1}, b^{t-1})\}$. Specifically, we only allow the encoder access to channel output feedback. This will allow us to use the coding results for the general channel discussed in sections 3 and 4.

A. Coding Theorem for Markov Channels

One can convert any Markov channel into a general channel.

Lemma 5.1: Given a Markov channel, $\{P(dB_t \mid s_t, a_t), P(dS_{t+1} \mid s_t, a_t)\}$ we have $\vec{P}(b^T \mid a^T)$ $= \sum_{s^T} \vec{P}(s^T, b^T \mid a^T) = \sum_{s^T} \vec{P}(s^T \mid a^T) \ \vec{P}(b^T \mid s^T, a^T)$. Here the directed probabilities are with respect to the time-ordering given in (2).

If we have a Markov channel without ISI then $\vec{P}(b^T \mid a^T) = \sum_{s^T} P(s^T) \ \vec{P}(b^T \mid s^T, a^T)$.

Because we can convert any Markov channel with output feedback into a general channel of the form studied earlier in this paper we can define the operational channel capacity, C^{O}, for the Markov channel with feedback in exactly the same way we did in definition 3.3. We can also use the same definitions of capacity, C, as before. Thus we can directly apply theorem 4.1 to prove:

Theorem 5.1: For Markov channels we have $C^{\mathrm{O}} = C$.

B. The Sufficient Statistic Π_t and the Canonical Markov Channel

Here we introduce the *canonical* Markov channel associated with a given Markov channel. Computing $I(A^t; B_t \mid B^{t-1})$ requires knowing the measure $P(dA^t, dB^t)$. The state S_t does not show up explicitly in $I(A^t; B_t \mid B^{t-1})$. We can estimate the state S_t from the information in A^{t-1}, B^{t-1}. Define the sufficient statistics: $\Pi_t(\cdot) = P_{S_t|(A^{t-1}, B^{t-1})}(\cdot \mid A^{t-1}, B^{t-1}) \in \mathcal{P}(\mathcal{S})$. Here $\mathcal{P}(\mathcal{S})$ represents the space of all measures on $\mathcal{S}$. Note that statistic Π_t depends on information from *both* the transmitter and the receiver. Furthermore under any channel input distribution $S_t - \Pi_t - (A^{t-1}, B^{t-1})$ forms a Markov chain for each t.

Lemma 5.2: There exists a function Φ_Π such that $\Pi_{t+1} = \Phi_\Pi(\Pi_t, A_t, B_t)$.

We now define the *canonical* Markov channel, $\{P(d\Pi_{t+1} \mid \pi_t, a_t), P(dB_t \mid \pi_t, a_t)\}$, associated with the Markov channel $\{P(dS_{t+1} \mid s_t, a_t), P(dB_t \mid s_t, a_t)\}$. By lemma 5.1 we can convert this Markov channel into an equivalent general channel $\{P(dB_t \mid a^t, b^{t-1})\}$. For this channel $P(b_t \mid a^t, b^{t-1}) = \sum_{s_t} P(b_t \mid s_t, a_t) \ P(s_t \mid a^t, b^{t-1}) = \sum_{s_t} P(b_t \mid s_t, a_t) \ \pi_t(s_t) = P(b_t \mid \pi_t, a_t)$. Thus $B_t - (A_t, \Pi_t) - (A^{t-1}, B^{t-1})$ forms a Markov chain. The new state Π_t evolves as:

$$P(\pi_{t+1} \in \Omega \mid \pi_t, a_t) = \sum_{\{b_t \text{ such that } \Phi_\Pi(\pi_t, a_t, b_t) \in \Omega\}} P(b_t \mid \pi_t, a_t)$$

for any Borel measurable set $\Omega \subset \mathcal{P}(\mathcal{S})$.

The next lemma shows that the use of $\{\Pi_t\}$ can simplify the form of the directed information.

Lemma 5.3: $I(F^T \to B^T) = I(A^T \to B^T) = \sum_{t=1}^{T} I(A_t, \Pi_t; B_t \mid B^{t-1})$
Proof: The first equality follows from lemma 4.1. The second equality follows from noting that $I(A^T \to B^T) = \sum_{t=1}^{T} I(A^t; B_t \mid B^{t-1})$ and

$$\begin{aligned} I(A^t; B_t \mid B^{t-1}) &= I(A_t, \Pi_t; B_t \mid B^{t-1}) \\ &\quad + I(A^{t-1}; B_t \mid \Pi_t, A_t, B^{t-1}) \\ &\quad - I(\Pi_t; B_t \mid A^t, B^{t-1}) \\ &= I(A_t, \Pi_t; B_t \mid B^{t-1}) \end{aligned}$$

Because Π_t is a function of A^{t-1}, B^{t-1} we see that $I(\Pi_t; B_t \mid A^t, B^{t-1}) = 0$. Because $A^{t-1} - (\Pi_t, A_t, B^{t-1}) - B_t$ is a Markov chain we see that $I(A^{t-1}; B_t \mid \Pi_t, A_t, B^{t-1}) = 0$. $\square$

In summary, every Markov channel with feedback that depends only on the channel outputs can be converted into canonical Markov channel with state Π_t. The canonical Markov channel has the property that Π_t is a function of the channel inputs and channel outputs only. Any residual "randomness" in S_t is captured by Π_t.

Most existing results in the literature examine non-ISI Markov channels. Note, though, that even if the original Markov channel $\{P(dS_{t+1} \mid s_t, a_t)\}$ does not have ISI it is generically the case that the corresponding canonical Markov channel $\{P(d\Pi_{t+1} \mid \pi_t, a_t)\}$ does have ISI. The channel input distribution has two roles: transmitting information about the message and probing the state of the channel. This phenomena is well known in the theory of partially observed stochastic control and is called the *dual effect* [2]. The ISI facilitates probing of the state.

VI. Markov Channels with Output Feedback

Here we formulate the following optimization problem for Markov channels as an infinite horizon average cost problem: $\sup_{\mathcal{D}_\infty} \liminf_{T\to\infty} \frac{1}{T} I(A^T \to B^T) =$

$$\sup_{\mathcal{D}_\infty} \liminf_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} I(A_t, \Pi_t; B_t \mid B^{t-1}). \tag{3}$$

To compute $I(A_t, \Pi_t; B_t \mid B^{t-1})$ we need to know the measure: $P(d\Pi_t, dA_t, dB_t \mid b^{t-1}) =$

$$P(dB_t \mid \pi_t, a_t) \otimes P(dA_t \mid \pi_t, b^{t-1}) \otimes P(d\Pi_t \mid b^{t-1}). \tag{4}$$

By lemma 5.2 we know that the pair Π_t, B^{t-1} is a function of the encoder's information pattern (A^{t-1}, B^{t-1}). Thus the stochastic kernels $\{P(dA_t \mid \pi_t, b^{t-1})\}$ constitute a valid channel input distribution.

To formulate the optimization in (3) as a stochastic control problem we need to specify the state space, the control actions, and the running cost. On first glance it may appear that the encoder should choose control actions of the form $P(dA_t)$ based on its information pattern A^{t-1}, B^{t-1}. From (4) we can see that this decision would only need to based on Π_t, B^{t-1}. Unfortunately we cannot write the running cost in terms of actions of the form $P(dA_t)$. To see this observe that the argument under the expectation in $I(A_t, \Pi_t; B_t \mid B^{t-1}) = E\left[\log \frac{P(B_t \mid A_t, \Pi_t)}{P(B_t \mid B^{t-1})}\right]$ can be written as

$$\log \frac{P(b_t \mid a_t, \pi_t)}{P(b_t \mid b^{t-1})}$$

$$= \log \frac{P(b_t \mid a_t, \pi_t)}{\int P(b_t \mid \tilde{a}_t, \tilde{\pi}_t) P(d\tilde{a}_t \mid \tilde{\pi}_t, b^{t-1}) P(d\tilde{\pi}_t \mid b^{t-1})}. \tag{5}$$

This depends on $P_{A_t \mid \Pi_t, B^{t-1}}(\cdot \mid \cdot, b^{t-1})$ and not $P_{A_t \mid \Pi_t, B^{t-1}}(\cdot \mid \pi_t, b^{t-1})$.

To treat this issue we take the view that the decoder chooses a control $u(dA, d\Pi)$ in the control space $\mathcal{U} = \mathcal{P}(\mathcal{A} \times \mathcal{P}(\mathcal{S}))$. The space $\mathcal{U}$ is a Polish space (i.e. a complete, separable metric space.)

At time t the decoder's state is given by its information pattern: (B^{t-1}, U^{t-1}). The policy at time t is given by $\mu_t : (b^{t-1}, u^{t-1}) \mapsto u_t(dA_t, d\Pi_t) \in \mathcal{U}$. Note that because the encoder has access to b^{t-1} it too can compute the control u_t. The dynamics are given by

$$\begin{aligned}
& P\left((b^t, u^t) \mid (b^{t-1}, u^{t-1}), u_t\right) \\
= \; & P(b_t \mid b^{t-1}, u^t) \\
= \; & \int P(b_t \mid a_t, \pi_t, b^{t-1}, u^t) P(da_t, d\pi_t \mid b^{t-1}, u^t) \\
= \; & \int P(b_t \mid a_t, \pi_t) u_t(da_t, d\pi_t)
\end{aligned}$$

The cost at time t is given by

$$c_t\left((b^{t-1}, u^{t-1}), u_t\right) =$$

$$\int P(db_t|\pi_t, a_t) u_t(da_t, d\pi_t) \log \frac{P(b_t|a_t, \pi_t)}{\int P(b_t|\tilde{a}_t, \tilde{\pi}_t) u_t(d\tilde{a}_t, d\tilde{\pi}_t)} \tag{6}$$

In order for this cost to be consistent with (4) and (5) the decoder must choose a control action u_t such that the marginal $u_t(d\Pi_t) = P^\mu(d\Pi_t \mid b^{t-1}, u^{t-1})$. Here P^μ represents the measure under the control policy $\{\mu_t\}$. Note that the only information in (b^{t-1}, u^{t-1}) relevant to choosing the control u_t is $P^\mu(d\Pi_t \mid b^{t-1}, u^{t-1})$. Thus we can simplify the state of the stochastic control problem.

Denote the decoder's estimate of the state of the canonical Markov channel by: $\Gamma_t^d(d\Pi_t) = P^\mu(d\Pi_t \mid B^{t-1}, U^{t-1}) \in \mathcal{P}(\mathcal{P}(\mathcal{S}))$. At time t the estimate Γ_t^d will be the simplified state of our stochastic control problem. Thus the control law can be rewritten as

$$\mu_t : \mathcal{P}(\mathcal{P}(\mathcal{S})) \to \mathcal{U} \quad \text{taking} \quad \gamma_t^d \mapsto u_t(dA_t, d\Pi_t) \in \mathcal{U}(\gamma_t^d)$$

where the control constraint is

$$\mathcal{U}(\gamma^d) = \left\{u(dA, d\Pi) \in \mathcal{U} \; : \; u(d\Pi) = \gamma^d(d\Pi)\right\}.$$

Note that for each γ^d the control constraint space $\mathcal{U}(\gamma^d)$ is Polish.

The following lemma ensures that $\Gamma_t^d(\cdot)$ is well defined and can be determined independently of the policy $\{\mu_t\}$ in place. That is $P^\mu(d\Pi_t \mid b^{t-1}, u^{t-1}) = P(d\Pi_t \mid b^{t-1}, u^{t-1})$.

Lemma 6.1: There exists a function Φ_{Γ^d} such that $\Gamma_{t+1}^d = \Phi_{\Gamma^d}(\Gamma_t^d, B_t, U_t)$ for $U_t \in \mathcal{U}(\Gamma_t^d)$. Thus $P(d\Pi_t \mid b^{t-1}, u^{t-1})$ can be determined independently of the policy in place.

For any Borel measurable set $\Omega \subset \mathcal{P}(\mathcal{P}(\mathcal{S}))$ and $u_t \in \mathcal{U}(\gamma_t^d)$ the dynamics are given by

$$\begin{aligned}
& P(\gamma_{t+1}^d \in \Omega \mid \gamma_t^d, u_t) \\
= & \int P(\gamma_{t+1}^d \in \Omega | \gamma_t^d, u_t, b_t, a_t, \pi_t) P(db_t, da_t, d\pi_t | \gamma_t^d, u_t) \\
\overset{(a)}{=} & \int P(\gamma_{t+1}^d \in \Omega | \gamma_t^d, u_t, b_t) P(db_t | \pi_t, a_t) u_t(da_t | \pi_t) \gamma_t^d(d\pi_t)
\end{aligned}$$

where (a) follows from lemma 6.1.

Finally we redefine the running cost (6) for our simplified state: $c(\gamma_t^d, u_t) =$

$$\int P(db_t \mid \pi_t, a_t) u_t(da_t, d\pi_t) \log \frac{P(b_t \mid a_t, \pi_t)}{\int P(b_t \mid \tilde{a}_t, \tilde{\pi}_t) u_t(d\tilde{a}_t, \tilde{\pi}_t)}.$$

For any policy $\{\mu_t\}$ satisfying the control constraints we see that equation (4) factors $P^\mu(d\Pi_t, dA_t, dB_t \mid b^{t-1}) = P(dB_t \mid \pi_t, a_t) \otimes P^\mu(dA_t \mid \pi_t, b^{t-1}) \otimes P^\mu(d\Pi_t \mid b^{t-1})$. Let $\mu^t(b^{t-1}) = (\mu_1, \mu_2(b_1), ..., \mu_t(b^{t-1}))$. Then $P^\mu(d\Pi_t \mid b^{t-1}) = P(d\Pi_t \mid b^{t-1}, \mu^{t-1}(b^{t-2}))$ and $P^\mu(dA_t \mid \pi_t, b^{t-1}) =$

$$\mu_t(P^\mu(d\Pi_t \mid b^{t-1}))(dA_t \mid \pi_t) = P^\mu(dA_t \mid \pi_t, \gamma_t^d) \quad (7)$$

is a valid channel input distribution.

Thus under $P^\mu(A^T, B^T, U^T)$ we have $E^\mu\left[c\left(\Gamma_t^d, \mu_t(\Gamma_t^d)\right)\right]$

$$\begin{aligned} &= E^\mu\left[\int P(db_t \mid a_t, \pi_t)\mu_t(\Gamma_t^d)(da_t, d\pi_t)\right. \\ &\qquad \left. \times \log \frac{P(b_t \mid a_t, \pi_t)}{\int P(b_t \mid \tilde{a}_t, \tilde{\pi}_t)\mu_t(\Gamma_t^d)(d\tilde{a}_t, d\tilde{\pi}_t)}\right] \\ &= E^\mu\left[\int P(db_t \mid a_t, \pi_t) P^\mu(da_t, d\pi_t \mid \Gamma_t^d)\right. \\ &\qquad \left. \times \log \frac{P(b_t \mid a_t, \pi_t)}{\int P(b_t \mid \tilde{a}_t, \tilde{\pi}_t) P^\mu(d\tilde{a}_t, d\tilde{\pi}_t \mid \Gamma_t^d)}\right] \\ &= I(A_t, \Pi_t; B_t \mid \Gamma_t^d) \end{aligned}$$

Finally by (7) we see $I(A_t, \Pi_t; B_t \mid \Gamma_t^d) = I(A_t, \Pi_t; B_t \mid B^{t-1})$.

In summary, at the end of the $t-1$-th epoch the receiver observes B_{t-1} and feeds this back to the transmitter. At the beginning of the t-th epoch the encoder computes Π_t and Γ_t^d and the decoder computes Γ_t^d. The pair (Π_t, Γ_t^d) can be viewed as the state of the encoder and Γ_t^d can be viewed as the state of the decoder. The evolution of Γ_t^d and Π_t are described in lemmas 6.1 and 5.2 respectively. Both the encoder and decoder know the policy $\{\mu_t\}$ and hence both can determine the control action $U_t(\cdot, \cdot) = \mu_t(\Gamma_t^d)$. The encoder, which has computed π_t, then randomly draws an A_t from the distribution $U_t(dA_t \mid \pi_t)$ and places it on the channel.

We now present the infinite horizon average cost verification theorem. First note that the cost $0 \le c\left(\gamma_t^d, u_t\right) \le \log|\mathcal{B}|$ is bounded. The state space $\mathcal{P}(\mathcal{P}(\mathcal{S}))$ and the constrained actions spaces $\mathcal{U}(\gamma^d)$ are all Polish.

Theorem 6.1: If there exists a bounded number V^*, a bounded function $w : \gamma^d \mapsto w(\gamma^d) \in I\!R$, and a policy μ^* achieving the supremum for each γ^d in the following average cost optimality equation (ACOE):

$$V^* + w(\gamma^d) = \sup_{u \in \mathcal{U}(\gamma^d)} \left(c(\gamma^d, u) + \int w(\tilde{\gamma}^d) P(d\tilde{\gamma}^d | \gamma^d, u) \right) \quad (8)$$

then the supremizing μ^* corresponds to the optimal channel input distribution for the optimization given in (3) and V^* is the feedback capacity.

Proof: See theorems 6.2 and 6.3 of [1].

Presenting conditions for the existence of a solution to the ACOE will take us too far afield. A good survey can be found in [1]. Note that the ACOE, equation (8), can be viewed as an implicit "single-letter characterization" of the capacity of the Markov channel.

VII. Conclusion

In this paper we presented a general coding theorem for channels with feedback. To prove this theorem we used Massey's concept of directed information. We showed the equivalence of the channel with feedback to another channel without feedback. We then examined the class of Markov channels with output feedback. We showed that the problem of feedback coding for Markov channels can be cast as a partially observed optimal stochastic control problem. Consequently one can now use the tools of exact and approximate dynamic programming to compute the capacity of a large class of Markov channels with feedback.

References

[1] A. Arapostathis, V. Borkar, E. Fernández-Gaucherand, M. Ghosh, and S. Marcus, "Discrete-Time Controlled Markov Processes with Average Cost Criterion: A Survey." Siam Journal on Control and Optimization, Vol. 31, No. 2, pp. 282-344, March 1993.

[2] D. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific, 1995.

[3] R. Dobrushin, "Transmission of Information in Channels with Feedback." Theor. Prob. Appl., 3, N4, pp. 395-412, 1958.

[4] R. Dobrushin, "A General Formulation of the Basic Shannon Theorem in Information Theory," Uspekhi Math. Nauk, 14, No. 6, pp. 3-103, 1959.

[5] A. Feinstein, "A New Basic Theorem of Information Theory." IRE Trans. on Information Theory, Vol 4, pp. 2-22, 1954

[6] G. Kramer, "Capacity Results for the Discrete Memoryless Network." IEEE Trans. Info. Theory, vol. IT-49, Issue 1, pp. 4-21, January 2003.

[7] J. Massey, "Causality, Feedback, and Directed Information." Proceedings of the 1990 International Symposium on Information Theory and its Applications (ISITA-90), pp. 303-305.

[8] C. Shannon, "A Mathematical Theory of Communication," Bell System Technical Journal, Vol. 27 (July and October 1948), pp. 379-423 and 623-656.

[9] C. Shannon, "The Zero Error Capacity of a Noisy Channel," Institute of Radio Engineers, Transactions on Information Theory, Vol. IT-2 (September, 1956), pp. S8-S19.

[10] S. Tatikonda, "Control Under Communication Constraints." MIT Ph.D. thesis, August 2000.

[11] S. Tatikonda and S. Mitter, "The Capacity of Channels with Feedback," manuscript in progress, 2005.

[12] S. Verdú, T. Han, "A general formula for channel capacity." IEEE Trans. Info. Theory, vol. IT-40, pp. 1147 - 1157, July 1994.