Information Theory of Intercellular Signal Transduction

Andrew W. Eckford
Dept. of EECS
York University, Toronto, Canada
Email: aecford@yorku.ca

Peter J. Thomas
Dept. of Mathematics
Case Western Reserve University, USA
Email: pjthomas@case.edu

Abstract—The cells in the human body use intercellular signal transduction to organize themselves and regulate their work; multicellular life could not exist without it. In this paper, we discuss why information theory provides a useful toolbox to approach this problem. We also discuss how to calculate the capacity of intercellular signal transduction in some specific examples. Finally, we describe the challenges and future prospects of this avenue of research.

I. INTRODUCTION

Molecular communications, a rapidly growing subfield of information theory, addresses the fundamental limits of communication at micron and submicron length scales, and strategies for achieving those limits [1]. Driven in part by the anticipated need for nanoscale agents to communicate and coordinate their actions, for instance for nanomedicine applications [2], communications engineering can find inspiration in naturally evolved approaches for signaling systems in environments hostile to traditional wired and wireless solutions – such as the interior of the human body.

Living cells monitor their environments and signal to one another using chemical signals communicated via diffusion, and transduced by receptor proteins [3], [4]. Early information theoretic analyses of biochemical signaling focused on chemotaxis, the directed movement of migrating cells in response to chemical cues, in part because the input/output signal ensembles could be clearly defined [5]–[10]. Advances in high-throughput experimental techniques have facilitated gathering enough data to quantitatively measure the capacity of specific signaling pathways [11]–[13]. At the same time, interest has grown among information theorists in a constellation of microbiological signaling problems, for instance the problem of a population of agents collectively estimating local concentration of chemical species (the consensus problem [14], [15]), the problem of memory effects in diffusion-mediated communication [16], [17], bacterial quorum sensing [18], [19], and communication via bacterial cables [20]. In this paper we discuss recent work on the capacity of a simple intercellular signal transduction system based on the cAMP receptor of the social amoeba Dictyostelium discoideum [21].

II. MARKOV MODELS OF SIGNAL TRANSDUCTION

As we show in this section, signal transduction can be modelled as a finite state Markov chain. Thus, we can take advantage of the rich toolbox of Markov chains [22] to analyze signal transduction.

A. Biological model

Throughout this paper, we use cyclic adenosine monophosphate (cAMP) as a motivating example to discuss signal transduction. This molecule is used in many biological processes, for example by the amoeba Dictyostelium discoideum in determining its social behaviour. Under normal circumstances, Dictyostelium act as unicellular individuals, but under stress, numerous individuals gather together (in response to cAMP) and form a multicellular “slug” in which the former independent cells take on specialized roles. Dictyostelium is well studied as a model organism for signal transduction.

cAMP has a simple (yet illustrative) model. The receptor can be in one of two states: unbound (U) or bound (B). In state U, the receptor awaits the arrival of a cAMP molecule; once one arrives, the receptor enters state B. In state B, the receptor cannot bind to other cAMP molecules (rendering it insensitive to the signal), while some processing time is required before returning to state U. We describe continuous-time and discrete-time mathematical models for this process below.

B. Signal transduction as a Poisson process

In continuous time, the binding process is represented as a Poisson process, transitioning between states U and B. The transition rate from U to B is \( r_{UB}(t) \), proportional to ligand concentration \( c(t) \); the transition rate from B to U is \( r_{BU} \), independent of ligand concentration.

The transition rates are given as a time-varying matrix \( R(t) \), as follows:

\[
R(t) = \begin{bmatrix}
-r_{UB}(t) & r_{UB}(t) \\
r_{BU} & -r_{BU}
\end{bmatrix}.
\]

(Note that the rows of \( R \) must sum to zero.) Let \( p(t) = [p_U(t), p_B(t)] \) represent the time-varying row vector of state occupancy probabilities; these probabilities are related to the Poisson rates via the differential equation

\[
\frac{d}{dt} p(t) = p(t) R(t).
\]

Equation (2) forms the master equation of the system.
C. Signal transduction as a Markov process

We are interested in representing the binding process as a discrete-time Markov chain. We can discretize time into steps of length $\Delta t$. Now, if $I$ is the $2 \times 2$ identity matrix, the Poisson process becomes a discrete-time Markov chain with state transition probability matrix (at the $i$th step)

$$P_i = I - \Delta t R(i \Delta t) = \begin{bmatrix} 1 - \Delta t r_{UB} c(i \Delta t) & \Delta t r_{UB} c(i \Delta t) \\ \Delta t r_{BU} & 1 - \Delta t r_{BU} \end{bmatrix}. \quad (3)$$

For consistency with notation in [23], we let

$$c_i = c(i \Delta t) \quad (5)$$
$$\alpha_{c_i} = \Delta t r_{UB} c(i \Delta t) \quad (6)$$
$$\beta = \Delta t r_{BU} \quad (7)$$

and (4) becomes

$$P_i = \begin{bmatrix} 1 - \alpha_{c_i} & \alpha_{c_i} \\ \beta & 1 - \beta \end{bmatrix}. \quad (8)$$

The state transition diagram is given in Figure 1.

The ligand concentration $c_i$ exists in a range, from the minimum allowed concentration $c_i = L$ to the maximum $c_i = H$. (Physically, there must exist a finite maximum concentration. The minimum may be zero.) For simplicity, we will use only the extreme concentrations $c_i = L$ and $c_i = H$; however, this restriction does not limit our analysis in the next section.

Thus, given the concentration, the signal transduction process is a time-inhomogeneous Markov chain with either

$$P_H = \begin{bmatrix} 1 - \alpha_H & \alpha_H \\ \beta & 1 - \beta \end{bmatrix} \quad (9)$$

or

$$P_L = \begin{bmatrix} 1 - \alpha_L & \alpha_L \\ \beta & 1 - \beta \end{bmatrix}, \quad (10)$$

selected by the ligand concentration.

III. CAPACITY OF SIGNAL TRANSMISSION

Here we review some of our main results on the capacity of signal transduction; more details can be found in [23], [24].

A. Signal transduction as a communication channel

To define a communication channel, we must define inputs, outputs, and channel input-output relationship:

- **Input.** The input $X$ is the concentration of ligands in the environment. We have a binary channel input: $X \in \{L, H\}$.
- **Output.** The output $Y$ is the state of the receptor. We have a binary channel output: $Y \in \{U, B\}$.
- **Input-output relationship.** As a Markov channel, the state of the output at time $i$, $Y_i$, depends on the current input $X_i$ and the previous channel output $Y_{i-1}$. We can write

$$p_{Y_i|X_i,Y_{i-1}}(y_i| x_i, y_{i-1}),$$

where the correct probability is given by an entry in either (9) or (10). In particular, if $X_i = H$, we have

$$p_{Y_i|X_i,Y_{i-1}}(B| H, U) = \alpha_H, \quad (11)$$

selected from (9) since $X_i = H$.

From now on, we will omit the subscripts for probability mass functions where unambiguous, e.g. $p_Y(y)$ becomes $p(y)$.

The conditional probability of a vector of outputs $Y^n_i$ given a vector of inputs $X^n_i$ is written

$$p(y^n_i | x^n_i) = \prod_{i=1}^n p(y_i | x_i, y_{i-1}), \quad (12)$$

where $y_0$ is null, and where each $p(y_i | x_i, y_{i-1})$ is selected from (9) or (10), as above.

B. Capacity and IID capacity

Logarithms are base 2 throughout. We will use the following special functions: let

$$\phi(p) = \begin{cases} 0, & p = 0 \\ -p \log p, & p \neq 0 \end{cases} \quad (13)$$

represent the partial entropy function, and let

$$\mathcal{H}(p) = \phi(p) + \phi(1 - p) \quad (14)$$

represent the binary entropy function.

For any communication system with inputs $X_i$ and outputs $Y_i$, the mutual information rate is given by

$$I(X^n_i;Y^n_i) = \frac{1}{n} \lim_{n \to \infty} I(X^n_i;Y^n_i) \quad (15)$$

where $I(X^n_i;Y^n_i)$ is the mutual information between vectors $X^n_i$ and $Y^n_i$. In our specific case, it can be shown that

$$I(X^n_i;Y^n_i) = \lim_{n \to \infty} H(Y_n | Y_{i-1}^n) - H(Y_n | X_n, Y_{i-1}^n). \quad (16)$$

The capacity of the system is given by

$$C = \max_{p(x^n_i)} \mathcal{I}(X^n_i;Y^n_i), \quad (17)$$

where the maximum is taken over all possible input distributions $p(x^n_i)$.

If we restrict the input distributions in (17) to be independent and identically distributed (IID), then we have the IID capacity, $C_{\text{IID}}$. Since the maximizing distribution in (17) is possibly (but not necessarily) IID, it should be clear that $C_{\text{IID}} \leq C$. 

\[120\]
To calculate $C_{\text{ID}}$, first note that the process $Y^n_1$ is a time-homogeneous Markov chain if $X^n_1$ is unknown but IID: (12) becomes

$$p(y^n_i) = \sum_{x_1} \cdots \sum_{x_n} \prod_{i=1}^n p(y_i | x_i)p(x_i)$$

$$= \sum_{i=1}^n p(y_i | x_i)p(x_i).$$

Letting

$$p(y_i | y_{i-1}) = \prod_{i=1}^n p(y_i | x_i)p(x_i),$$

and letting

$$\alpha = p_X(L)\alpha_L + p_X(H)\alpha_H,$$

we have that $Y^n_1$ is a Markov chain with transition probability matrix

$$P = p_X(L)P_L + p_X(H)P_H$$

$$= \begin{bmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{bmatrix}.$$ (22)

Moreover, as a two-state Markov chain, the steady-state distribution of the process $Y$ is given by

$$p_Y(y) = \begin{cases} \frac{\beta}{\alpha + \beta}, & y = U \\ \frac{\alpha}{\alpha + \beta}, & y = B \end{cases}.$$ (24)

In (16), $H(Y_n | Y^{n-1})$ becomes $H(Y_n | Y_{n-1})$, since the process $Y$ is a Markov chain. This quantity is given by

$$H(Y_n | Y_{n-1})$$

$$= -E[\log p(y_n | y_{n-1})]$$

$$= -\sum_{y_{n-1}} p(y_{n-1}) \sum_{y_n} p(y_n | y_{n-1}) \log p(y_n | y_{n-1})$$

$$= \frac{\beta}{\alpha + \beta}H(\alpha_L) + \frac{\alpha}{\alpha + \beta}H(\beta).$$ (27)

Similarly,

$$H(Y_n | X_n, Y_{n-1})$$

$$= -E[\log p(y_n | x_n, y_{n-1})]$$

$$= \frac{\beta}{\alpha + \beta} \left( p_X(L)H(\alpha_L) + p_X(H)H(\alpha_H) \right) + \frac{\alpha}{\alpha + \beta}H(\beta).$$ (29)

Substituting back into (16),

$$I(X; Y)$$

$$= \frac{\beta}{\alpha + \beta} \left( H(\alpha) - p_X(L)H(\alpha_L) - p_X(H)H(\alpha_H) \right).$$ (30)

Finally, $C_{\text{ID}}$ is found by maximizing (30) with respect to $p_X(L)$, recalling that $p_X(H) = 1 - p_X(L)$. Since $C_{\text{ID}} \leq C$, this provides us with a lower bound on capacity.

In [23], it is shown that even if inputs at intermediate concentrations other than $L$ and $H$ are allowed, the input distribution that maximizes $C_{\text{ID}}$ has all probability mass on $L$ and $H$, and none on any intermediate concentration.

### C. Feedback capacity

Suppose the transmitter has causal feedback of all past outputs $Y_{1}^{n-1}$, and uses these outputs to determine its strategy in setting $X_i$; the best such strategy gives the feedback capacity $C_{\text{FB}}$.

In our model of the signal transduction channel, the output is the same as the Markov state of the channel. Such channels are sometimes called Previous Output is the State (POST) channels [25], or unit output memory (UOM) channels [26]. For our specific case, $C_{\text{FB}}$ is given by the maximum of the directed information [27] over all causal-conditional input distributions of the form $\prod_{i=1}^{n} p(x_i | y_i^{n-1}, x_i^{n-1})$. Since the transmitter has the option of disregarding the feedback (resulting in the regular capacity $C$), we now have the sequence of bounds

$$C_{\text{ID}} \leq C \leq C_{\text{FB}}.$$ (31)

To obtain $C_{\text{FB}}$, it seems a daunting task to specify the correct input distribution for all settings of $y_i^{n-1}$ and $x_i^n$. However, from [26], we have a key simplification: in most UOM channels (including ours), the feedback-capacity-achieving input distribution is stationary, with the form

$$p(x^n_1 | y^n_1) = \prod_{i=1}^{n} p(x_i | y_{i-1}).$$ (32)

That is, $x_i$ is selected by considering only the two possibilities of $y_{i-1} \in \{U, B\}$. Moreover, in our specific case, it can be shown that the directed information rate has the same form as $I(X; Y)$ in (16); we need only to maximize over the input distributions in (32) to obtain $C_{\text{FB}}$.

Consider $H(Y_i | Y_{i-1})$ from (16). From the above discussion, we need to specify only $p_{X,Y|Y_{i-1}}(L | U)$ and $p_{X,Y|Y_{i-1}}(L | B)$. Since the input $x_i$ depends only on $y_{i-1}$, the process $Y$ with feedback is still a time-homogeneous Markov chain: we can substitute

$$\alpha = p_{X,Y|Y_{i-1}}(L | U)\alpha_L + (1 - p_{X,Y|Y_{i-1}}(L | U))\alpha_H$$ (33)

into $P$ from (23). Note that $\beta$ (the transition rate in state $B$) is independent of the input $x_i$. Thus, the transition probability matrix $P$ is independent of $p_{X,Y|Y_{i-1}}(L | B)$, as is $H(Y_i | Y_{i-1})$. By a similar argument, $H(Y_i | X_i, Y_{i-1})$ is also independent of $p_{X,Y|Y_{i-1}}(L | B)$.

Since the directed information rate is constant with respect to $p_{X,Y|Y_{i-1}}(L | B)$, it makes no difference to $C_{\text{FB}}$ if we set

$$p_{X,Y|Y_{i-1}}(L | B) = p_{X,Y|Y_{i-1}}(L | U) = p_X(L),$$ (34)

which is an IID input distribution.

Since $C_{\text{FB}}$ is satisfied with an IID input distribution, we have $C_{\text{ID}} = C_{\text{FB}}$; thus, from (31),

$$C_{\text{ID}} = C = C_{\text{FB}},$$ (35)

and the (non-feedback) capacity-achieving input distribution is IID.

Please see [23] for a more formal and complete description of this argument.
Fig. 2. Illustration of the information rate per time step $\Delta t$ as a function of $p_r(L)$. In this figure, $\alpha_L = 0.01$, $\beta = 0.1$, and $\alpha_M$ varies from 0.1 (bottom curve) to 0.9 (top curve) in increments of 0.1.

IV. DISCUSSION

In Figure 2, we illustrate the information rates of our system for some sample parameter values. The capacity $C$ may be determined by taking the maximum of each curve.

The techniques employed in this paper may be generalized to receptors with larger state spaces. In forthcoming work [28], we perform similar capacity calculations on the channelrhodopsin-2 (ChR2) receptor, with three states; and the acetylcholine (ACh) receptor, with five states. If there is exactly one state transition that is sensitive to the input signal, then the capacity-achieving input distribution is IID. This is the case for ChR2, but not for ACh. In the latter case, the IID capacity $C_{\text{IID}}$ remains a lower bound on the true capacity $C$.

REFERENCES


