35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Information Decoding in Microscopic Biological Processes

Tetsuya J. Kobayashi[1]

Abstract— **The cellular and intracellular dynamics are intrinsically stochastic and dynamic. However, whole biological system such as a cell or our body can function very robustly and stably even though they are composed of these stochastic reactions. To account for this riddling relation between macroscopic robustness and microscopic stochasticity, I propose a mechanism that information relevant for stable and reliable operation of a biological system is embedded in apparently stochastic and noisy behavior of their components. To show validity of this possibility, I demonstrates that information can actually be decoded from apparently noisy signal when it is processed by an appropriate dynamics derived by Bayes' rule. Next, I investigate biological relevance of this possibility by showing that several intracellular networks can implement this decoding dynamics. Finally, by focusing its dynamical properties, I show the mechanism how the derived dynamics can separate information and noise.**

I. Information Processing in Biological Systems

Innovation of bio-imaging technology revealed that cellular and intracellular dynamics are intrinsically stochastic and dynamic for various sub-processes such as signal transduction, cellular decision-making and gene expression [1]–[3]. This new experimental evidence apparently contradicts with our naive observations that whole biological system such as a cell or our body can function very robustly and stably even though they are composed of these stochastic reactions [4]. This fact sounds riddling from the engineering viewpoint in which stability of a system strongly depends on the stability and reliability of its building components, suggesting that there is unknown design principle in biological systems that leads to construction of stable systems out of unreliable components [5].

One possibility is to exploit stochasticity and noise. Stochastic resonance is a typical example of such noise-enhancement of biological functions, and has been investigate for several decades [6], [7]. However, stochastic resonance does not always explain stability and reliability of biological functions. In addition, the situation under which noise or stochasticity can be beneficial is limited mathematically except when cells need to use stochasticity to generate diversity in order to attenuate effect of environmental uncertainty [8].

Another possibility is that information relevant for stable and reliable operation of a biological system is embedded in apparently stochastic and noisy behavior of their components. Stability and reliability of a whole system may be realized if this information is exploited effectively by intracellular dynamics. This possibility has not yet been fully pursued despite its importance.

In this article, I address this problem by employing theory for Bayesian information processing as I did in [9]–[11]. Using this theory, I firstly derive a statistically optimal dynamics called here as information decoding dynamics for extraction of information from noisy signal. This demonstrates that information can actually be decoded from apparently noisy signal. Next, I investigate biological relevance of this possibility by showing that several intracellular networks can implement this decoding dynamics [9]–[11]. Finally, I also show the underlying mechanism how the derived dynamics can separate information and noise by focusing its dynamical properties. The extension of this approach is also discussed.

II. Modeling Binary Cellular Decision-Making

Binary decision-making is the simplest information processing in biological systems. In this process, a cell has to decide whether environment, $x(t)$, is in one state or the other, which are designated here with on and off. In general, the state of the environment changes over time stochastically. Thus, the sensing of environment is indispensable to appropriately respond to the current environmental state. Sensing by receptors is a major way to obtain information on $x(t)$. However, the sensing output $y(t)$ may not always provide accurate information on $x(t)$. One source to prevent accurate information transfer by $y(t)$ is low signal intensity of $y(t)$. For example, several environmental molecules exist at very low concentration. When the state of $x(t)$ is associated with the change in such molecule, $x(t)$-dependent change in $y(t)$ can be very subtle and hard to be identified. Another source is the stochasticity of receptor activation, which is induced either by random arrival of ligand related to $x(t)$ or by thermal fluctuation of receptor itself. Even when the signal intensity of $y(t)$ is large, information on $x(t)$ becomes ambiguous if fluctuation in $y(t)$ due to stochasticity is much larger than the signal intensity.

To account for the first one, I introduce two parameters, λ_{on} and λ_{off}, that designates the average frequency of receptor activation when $x(t) = \mathrm{on}$ and $x(t) = \mathrm{off}$, respectively. I define $\lambda(t)$ as $\lambda(t) := \lambda_{x(t)}$. The second source can be modeled by assuming that receptor activation occurs by following Poisson point process in which a receptor gets active n times during $[s,t]$ with the probability $\mathbb{P}_P(n;\Lambda_s^t) := (\Lambda_s^t)^n e^{-\Lambda_s^t}/n!$ where $\Lambda_s^t := \int_s^t \lambda(t')\mathrm{d}t'$. Let define $C_i(t)$ as the total number of activation of i-th receptor, and assume that a cell has total N_0 receptors. In addition, as in [9], the receptor is assumed to be get inactive after $\tau > 0$. Then $N(t) := \sum_{i=1}^{N_0} [C_i(t) - C_i(t-\tau)]$ represents the number

*This work was supported by JST, PRESTO

[1]T.J. Kobayashi is with Institute of Industrial Science, the University of Tokyo, 1538505 Tokyo, Japan. & Japan Science and Technology Agency, PRESTO. tetsuya@mail.crmind.net

978-1-4577-0216-7/13/$26.00 ©2013 IEEE

of active receptors at t. By normalizing $N(t)$, I define $y(t)$ as $y(t) := N(t)/(N_0\tau)$. In order to model stochastically changing environment, I also use two-state Markov model for the stochastic evolution of $x(t)$, i.e., $x(t)$ changes from off(on) to on(off) with the probability $r_{\rm on}\Delta t (r_{\rm off}\Delta t)$ within small interval $\Delta t > 0$. When $\lambda_{\rm on} - \lambda_{\rm off} \ll (\lambda_{\rm on} + \lambda_{\rm off})/2$, then $N(t)$ behaves very stochastically as if it does not reflect state of $x(t)$ as illustrated in Fig. 1 [9].

If we consider that there are different types of receptors that behave differently, then other sensing process such as gradient sensing can be modeled as an extension [10]. In gradient sensing, a cell obtains information of gradient direction in environment by using multiple receptors located spatially different positions on a cell [12]. Let $y_R(t)$ and $y_L(t)$ be the output of receptors on right and left membrane that follows Poisson point processes with intensity parameter, $\lambda^R(t)$ and $\lambda^L(t)$, respectively. When $x(t) = $ on corresponds to the situation that gradient is pointing to right, then $\lambda^R(t) = \lambda_{\rm on}$ and $\lambda^L(t) = \lambda_{\rm off}$ for $x(t) = $ on. When $x(t) = $ off corresponds to the situation that gradient is pointing to left, then $\lambda^R(t) = \lambda_{\rm off}$ and $\lambda^L(t) = \lambda_{\rm on}$ for $x(t) = $ off. Because of smallness of a cell, however, the difference of receptor signal at different positions, $\lambda_{\rm on} - \lambda_{\rm off}$, can be very small compared with the average absolute intensity, $(\lambda_{\rm on} + \lambda_{\rm off})/2$. Under this situation, $y_R(t)$ and $y_L(t)$ or corresponding $N_R(t)$ and $N_L(t)$ behave as in Fig. 2 where $N_R(t)$ and $N_L(t)$ are defined as $N(t)$ is.

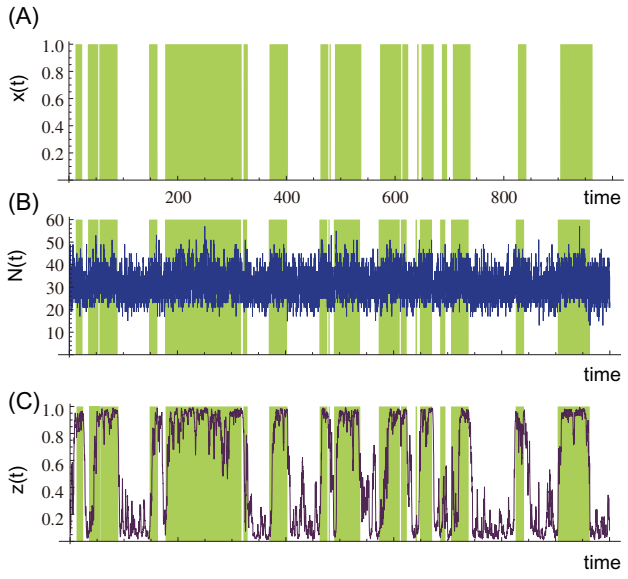

Fig. 1. Sample paths of binary cellular decision-making. (A):$x(t)$, (B):$N(t)$, and (C):$z(t)$. Green background indicates that $x(t) = $ on.

III. Extraction of Information by Bayesian Decoding Dynamics

As shown in Figs. 1 and 2, $y(t)$ or $y_R(t)$ and $y_L(t)$ are apparently very noisy as if they do not convey information on $x(t)$. However, information on $x(t)$ is contained in these noisy signals, and that information can be extracted when $y(t)$ are processed subsequently by appropriate dynamics. To

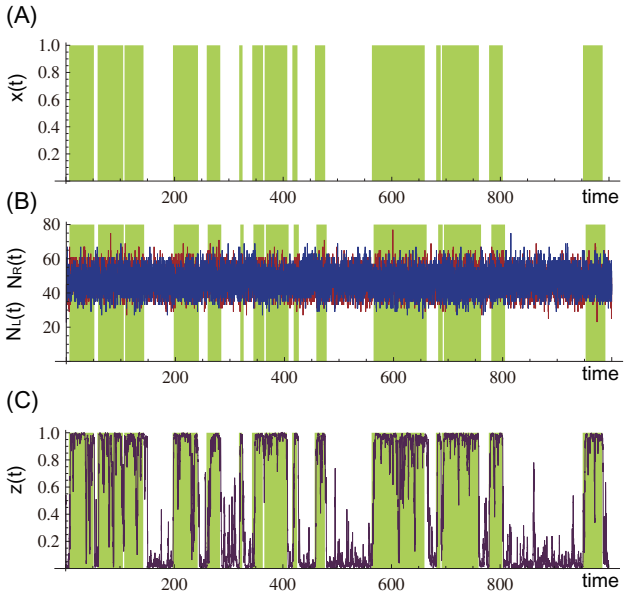

Fig. 2. Sample paths of gradient direction sensing. (A): $x(t)$, (B): $N_R(t)$, $N_L(t)$, and (C):$z(t)$. Green background indicates that $x(t) = $ on, meaning that gradient is pointing to right. Red and Blue curves are $N_R(t)$ and $N_L(t)$, respectively.

derive such dynamics, I here use Bayes' rule. Let $z(t)$ ($\tilde{z}(t)$) be the posterior probability that $x(t) = $ on ($x(t) = $ off) given the history of $y(t)$ as $z(t) := \mathbb{P}(t, x = \text{on}|Y(t))$ and $\tilde{z}(t) := \mathbb{P}(t, x = \text{off}|Y(t))$ where $Y(t) := \{y(t') : t' \in [0,t]\}$. Then for small $\Delta t > 0$, we can have

$$\mathbb{P}(t', x'|Y(t')) \propto \mathbb{P}^+_{\Delta t}(\Delta y|x) \sum_x \mathbb{P}_{\Delta t}(x|x') \mathbb{P}(t, x|Y(t)), \qquad (1)$$

where $\mathbb{P}_{\Delta t}(x'|x)$ is the probability that environment changes from x to x' within the time interval Δt. Because $x(t)$ follows two-state Markov process, $\mathbb{P}_{\Delta t}(x'|x) = r_{x'}\Delta t$ when $x' \neq x$ and $\mathbb{P}_{\Delta t}(x'|x) = 1 - r_x \Delta t$ when $x' = x$. In contrast, $\mathbb{P}^+_{\Delta t}(\Delta y|x)$ represents stochastic activations of receptors. Because each receptor activation is modeled by Poisson point process whose intensity is $\lambda(t)$, $\mathbb{P}^+_{\Delta t}(\Delta y|x) = \mathbb{P}_P(\Delta y; \lambda_x \Delta t)$ for sufficiently small Δt. By using these relation and taking the limit $\Delta t \to 0$ for sufficiently small $\tau > 0$, we obtain

$$\begin{aligned} \frac{dz(t)}{dt} &= z(t)\tilde{z}(t)\tfrac{\lambda_r}{\tau}\left[N(t) - \tau\lambda_m N_0\right] + r_{\rm on}\tilde{z}(t) - r_{\rm off} z(t), \\ &= z(t)\tilde{z}(t)\lambda_r N_0 \left[y(t) - \lambda_m\right] + r_{\rm on}\tilde{z}(t) - r_{\rm off} z(t), \end{aligned} \qquad (2)$$

where $\lambda_r := \log \lambda_{\rm on}/\lambda_{\rm off}$, $\lambda_d := \lambda_{\rm on} - \lambda_{\rm off}$, and $\lambda_m := \lambda_d/\lambda_r$ [9]. In addition, $z(t) + \tilde{z}(t) = 1$ holds. It should be noted that this equation is an approximation of Bayesian update of the posterior probability for finite τ [10].

This equation is driven only by noisy $y(t)$. Nonetheless, it can reconstruct behavior of $x(t)$ with high fidelity as demonstrated in Fig. 1. This result clearly indicates that information can be extracted even from apparently noisy signal when the noisy signal is processed appropriately.

When a cell receive input from multiple receptors $y_R(t)$ and $y_L(t)$ to know the state of environmental gradient, similar

equation can be obtained as

$$\begin{aligned}\frac{\mathrm{d}z(t)}{\mathrm{d}t} &= z(t)\tilde{z}(t)\tfrac{\lambda_r}{\tau}\left[N_R(t)-N_L(t)\right]+r_{\mathrm{on}}\tilde{z}(t)-r_{\mathrm{off}}z(t),\\ &= z(t)\tilde{z}(t)\lambda_r N_0\left[y_R(t)-y_L(t)\right]+r_{\mathrm{on}}\tilde{z}(t)-r_{\mathrm{off}}z(t). \quad (3)\end{aligned}$$

Fig. 2 illustrates that this equation also extracts the state of gradient direction from noisy $y_R(t)$ and $y_L(t)$ even when their difference is extremely small.

Since the process to reconstruct original $x(t)$ from noisy signal $y(t)$ as $z(t)$ is equivalent to the information decoding in Shannon's communication theory, the dynamics derived here is called as information decoding dynamics in this article [13].

IV. Biochemical Implementation of Bayesian Information Decoding

While information decoding dynamics can extract information on $x(t)$ from noisy $y(t)$, this does not mean that cellular systems can do the same thing. Because of physical properties of intracellular dynamics, only subset of dynamics can be implemented biochemically. Next question, therefore, is whether the information decoding dynamics can be implemented biochemically or not. To demonstrate that (2) and (3) are actually implementable biochemically, I first rearrange (2) as

$$\frac{\mathrm{d}z(t)}{\mathrm{d}t}=\left[\frac{\lambda_r}{\tau}N(t)z(t)+r_{\mathrm{on}}\right]\tilde{z}(t)-\left[\frac{\lambda_r}{\tau}\lambda_m N_0\tilde{z}(t)+r_{\mathrm{off}}\right]z(t).$$

If we consider $z(t)$ and $\tilde{z}(t)$ as the ratio of phosphorylated and unphosphorylated intracellular protein, then $R_{\mathrm{on}}=\left[\frac{\lambda_r}{\tau}N(t)z(t)+r_{\mathrm{on}}\right]$ and $R_{\mathrm{off}}=\left[\frac{\lambda_r}{\tau}\lambda_m N_0\tilde{z}(t)+r_{\mathrm{off}}\right]$ can be regarded as rate constants for phosphorylation and dephosphorylation, respectively. In addition, because these rates depend on the ratio of phosphorylated and unphosphorylated proteins, $z(t)$ and $\tilde{z}(t)$, these reactions can be biochemically realized by a positively auto-regulative reaction such as an auto-catalytic reaction. Therefore, (2) is shown to be biochemically implementable in principle, for example, by an auto-catalytic phosphorylation and de-phosphorylation cycle in which the noisy receptor signal $y(t)$ works to accelerate phosphorylation reaction [9].

Similarly, (3) can also be implementable by the following polarity formation reaction [11]:

$$\begin{aligned}\frac{\mathrm{d}B_R}{\mathrm{d}t} &= B_C\left[\nu B_R N_R(t)+k_a\right]-k_d B_R,\\ \frac{\mathrm{d}B_L}{\mathrm{d}t} &= B_C\left[\nu B_L N_L(t)+k_a\right]-k_d B_L,\\ \frac{\mathrm{d}B_C}{\mathrm{d}t} &= -B_C\left[\nu B_R N_R(t)+B_L N_L(t)+2k_a\right]+k_d(B_R+B_L).\end{aligned}$$

Here, B_R and B_L are the amounts of polarity protein localized on right and left membrane, respectively. In contrast, B_C is the amount of that protein in cytoplasm. As equations indicate, the localization to right (left) membrane is enhanced by the receptor activity on right (left) membrane $N_R(t)(N_L(t))$. In addition, this process is positively regulated by already localized protein. Let define $z(t)$ and $\tilde{z}(t)$ as $z(t):=B_R(t)/B_M(t)$ and $\tilde{z}(t):=B_L(t)/B_M(t)$ where $B_M(t):=(B_R(t)+B_L(t))$. Then, we can obtain equations for $z(t)$ and $\tilde{z}(t)$ as

$$\begin{aligned}\frac{\mathrm{d}z}{\mathrm{d}t} &= z(t)\tilde{z}(t)(\nu B_C)\left[N_R(t)-N_L(t)\right]+\frac{k_a B_C}{B_M}(\tilde{z}(t)-z(t)),\\ \frac{\mathrm{d}\tilde{z}}{\mathrm{d}t} &= z(t)\tilde{z}(t)(\nu B_C)\left[N_L(t)-N_R(t)\right]+\frac{k_a B_C}{B_M}(z(t)-\tilde{z}(t)).\end{aligned}$$

For $|\lambda_R(t)-\lambda_L(t)|\ll(\lambda_R(t)+\lambda_L(t))/2$, B_C and B_M can be regarded approximately constant as shown in [11]. Thus, by identifying $\nu B_C=\lambda_r/\tau$ and $r_{\mathrm{on}}=r_{\mathrm{off}}=k_a B_C/B_M$, the equation for $z(t)$ derived here becomes identical to (3), meaning that (3) is biochemically implementable.

These examples illustrate that information decoding dynamics are biochemically implementable and there are several ways to implement such dynamics because of the simplicity of (2) and (3).

V. Principle in Information Decoding Dynamics

While the impact of information decoding dynamics is illustrated by Figs.1 and 2, it is still uncertain why (2) and (3) are so efficient in terms of extraction of information. To reveal the underlying mechanism of information decoding, let break (2) into three pieces as

$$\frac{\mathrm{d}z(t)}{\mathrm{d}t}=\underbrace{z(t)\tilde{z}(t)}_{(B)}\underbrace{\lambda_r N_0\left[y(t)-\lambda_m\right]}_{(A)}+\underbrace{r_{\mathrm{on}}\tilde{z}(t)-r_{\mathrm{off}}z(t)}_{(C)}.$$

First, I focus on (A). By definition, $\lambda_m=\lambda_d/\lambda_r=\frac{\lambda_{\mathrm{on}}-\lambda_{\mathrm{off}}}{\log\lambda_{\mathrm{on}}/\lambda_{\mathrm{off}}}$. When $\Delta\lambda:=\lambda_{\mathrm{on}}-\lambda_{\mathrm{off}}$ is small, then $\lambda_m\approx(\lambda_{\mathrm{on}}+\lambda_{\mathrm{off}})/2+O(\Delta\lambda^2)$ holds. Thus, (A) is simply checking whether $y(t)$ is higher or lower than a threshold which is located at almost the middle of λ_{on} and λ_{off}. The more $y(t)$ deviates from λ_m, the more strongly $y(t)$ drives $z(t)$. Intuitively and qualitatively, this behavior sounds reasonable because small deviation is induced by noise more frequently than by actual change in $x(t)$ whereas large deviation may be induced by change in $x(t)$. However, (A) is linear with respect to $y(t)$, which cannot be automatically guaranteed by these qualitative arguments. From the statistical viewpoint, (A) is proposional to log-likelihood ratio between two hypothesis that $x(t)$ is on and off as

$$(A)=\left[\frac{\lambda_r}{\tau}N(t)-\lambda_d N_0\right]=\log\frac{\mathbb{P}_P(N(t);\lambda_{\mathrm{on}}\tau N_0)}{\mathbb{P}_P(N(t);\lambda_{\mathrm{off}}\tau N_0)}.$$

Thus, (A) quantitatively relates extent of deviation with extent of likelihood to observe such deviation under alternative hypothesis on the state of $x(t)$.

In contrast to (A), (B) is determined solely by the state of $z(t)$. Because (A) is multiplied by (B), (B) determines the sensitivity of $z(t)$ to respond to the receptor signaling. Since $(B)=z(t)(1-z(t))$, (B) is maximal when $z(t)=1/2$ whereas (B) approaches to zero when $z(t)$ is close to either 0 or 1. In other words, (2) is sensitive to input $y(t)$ when $z(t)$ is producing ambiguous output $z(t)=1/2$, but it is insensitive to $y(t)$ when $z(t)$ is producing distinctive output. This adaptive change in sensitivity also has an obvious statistical

meaning and an intuitive explanation. By definition, $z(t)$ is the posterior probability that $x(t) =$ on given the history of $y(t)$ up to t, $Y(t)$. Thus, $z(t)(1-z(t))$ is equivalent to the variance of this posterior probability. This fact means that the sensitivity of $z(t)$ is controlled by the uncertainty (variance) of its estimation of $x(t)$. In other words, $z(t)$ is sensitive to $y(t)$ when it estimates the state of $x(t)$ with low confidence whereas it is insensitive when with high confidence, which is understandable from our intuition.

Finally, (C) corresponds to the process to forget past information on $x(t)$. Since $x(t)$ changes over time, $y(t')$ of long past does not reflect current state of $x(t)$. Forgetting is indispensable to reduce the effect of past observation by $y(t)$ for estimation of $x(t)$. Quite interesting is that these intuitively important factors emerge auttomatically from Bayes' rule and are integrated minimally as in (2). It should be noted that the same interpretation is valid for (3).

VI. SUMMARY & DISCUSSION

In order to address the question how stable and robust behaviors emerge from stochastic and unreliable components in biological systems, in this article, I firstly showed that relevant information can be extracted from apparently noisy signal when appropriate dynamics processes that signal. This result implies that apparent stochasticity and noisiness in signal does not always mean that the signal does not contain information.

Secondly, I demonstrated that such dynamics called information decoding dynamics here can be implemented by using several biochemical reactions. This indicates that such information decoding dynamics may be embedded in actual biological systems especially when the systems produce robust output by processing very noisy signal.

Finally, I clarified the mechanism how the decoding dynamics can extract information from noisy signal. Three factors play important roles. One is quantification of probability of deviation in signal. Second one is adaptive sensitivity change depending on the uncertainty of estimate of information. Last one is the forgetting of past information. These factors are integrated within the very simple equation (2) or (3), and works as statistically optimal decoding dynamics.

While I treated very simple situation in which x_t has only two states and $x(t)$-dependent intensity of $y(t)$ is fixed, this approach can be extended for more complicated and biologically realistic situations. Such extension is going to be discussed in the near future.

ACKNOWLEDGMENT

I would like to thank Dr. Atsushi Kamimura for discussion and comment on this work.

REFERENCES

[1] M.B.Elowitz *et al*., Science, vol. 297, pp. 1183-1186, 2002 .
[2] N. Rosenfeld *et al*.,Science, vol. 307, pp. 1962-1965, 2005.
[3] Y.Taniguchi *et al*., Science, vol. 329, pp. 533-538, 2010.
[4] D. Bray, "Wet ware: A computer in every living cell", Yale Univ. Press, 2009.
[5] http://www.nature.com/nature/focus/biologicalnoise/index.html, 2010.
[6] Y. Morishita, T. J. Kobayashi, *et al*., Biophys, J. vol. 91, pp. 2072-2081, 2006.
[7] P. Hnggi, Chemphyschem, vol. 3, pp. 285—90, 2002.
[8] T. J. Kobayashi and A. Kamimura, Adv. Exp. Med. Biol., vol. 736, pp. 275-291, 2012.
[9] T. J. Kobayashi, Phys. Rev. Lett. vol. 104, pp. 228104, 2010.
[10] T. J. Kobayashi, and A. Kamimura, Phys. Biol. vol. 8, no.5, 2011.
[11] T. J. Kobayashi, Phys. Rev. Lett. vol. 106, pp. 228101, 2011.
[12] M. Ueda & T. Shibata, Biophys. J., vol. 93, pp. 11-20, 2007.
[13] T. M. Cover & J. A. Thomas, " Elements of Information Theory ", Wiley-Interscience, 2006.