Published in final edited form as:

Integr Biol (Camb). 2012 March ; 4(3): 310–317. doi:10.1039/C2IB90009B.

A framework for designing and analyzing binary decision-making strategies in cellular systems†

Joshua R. Porter, Burton W. Andrews, and Pablo A. Iglesias*

Department of Electrical and Computer Engineering, The Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218

Abstract

Cells make many binary (all-or-nothing) decisions based on noisy signals gathered from their environment and processed through noisy decision-making pathways. Reducing the effect of noise to improve the fidelity of decision-making comes at the expense of increased complexity, creating a tradeoff between performance and metabolic cost. We present a framework based on rate distortion theory, a branch of information theory, to quantify this tradeoff and design binary decision-making strategies that balance low cost and accuracy in optimal ways. With this framework, we show that several observed behaviors of binary decision-making systems, including random strategies, hysteresis, and irreversibility, are optimal in an information-theoretic sense for various situations. This framework can also be used to quantify the goals around which a decision-making system is optimized and to evaluate the optimality of cellular decision-making systems by a fundamental information-theoretic criterion. As proof of concept, we use the framework to quantify the goals of the externally triggered apoptosis pathway.

1 Introduction

Many cellular systems make binary (all-or-nothing) decisions based on information gathered from the environment. Some well-studied examples of such binary decisions are apoptosis in multicellular organisms^{1–3}, oocyte maturation in *Xenopus laevis*⁴, mating in *S. cerevisiae*^{5,6}, and induction of the *lac* operon in *E. coli*^{7–10}. The stakes in these decisions are high, as the survival of individuals and populations can depend on correct decision-making¹¹. However, cells' ability to accurately sense their environment to inform these decisions is limited by stochastic molecular fluctuations¹². Thus, error-free responses are impossible, and even minimizing error comes at the cost of maintaining complex biochemical pathways^{13,14}.

To understand binary decision-making systems, we need a way to evaluate their ability to make accurate judgments based on noisy, uncertain observations. How close to optimal is a decision-making strategy in a given situation? Alternatively, for what situation is a given strategy optimized? How do cells balance the metabolic cost of complex decision-making equipment with the benefit of accurate decisions? Finally, how do we design synthetic

†Electronic Supplementary Information (ESI) available: Rate distortion function and optimal decision-making strategies computed using a Gaussian distribution of stimuli (Fig. S1) and distortion function computed from apoptosis pathway using a Gaussian distribution of stimuli (Fig. S2). See DOI:10.1039/b000000x/

pi@jhu.edu; Fax: +1-410-516-5566; Tel: +1-410-516-6026.

biological systems to make optimal binary decisions? Here we present a framework based on information theory to answer these and related questions. Information theory, which was originally developed for use in telecommunications, has in recent years been increasingly applied to analyzing biological signaling pathways^{15–21}. Our framework uses rate distortion theory, a branch of information theory that deals with communication systems in which error-free transmission is impossible or undesirable due to its high cost²². Rate distortion theory quantifies the information cost associated with improved performance in such systems and identifies the mechanisms used to balance this tradeoff^{23,24}. In biology, rate distortion theory has been employed to understand the evolution of the genetic code²⁵; we previously applied it to finding optimal strategies for chemotaxis¹⁵.

Here we expand on our previous work by using rate distortion theory to design and evaluate strategies for making binary cell fate decisions. We regard a cellular decision-making system as a noisy communication channel with an environmental stimulus as its input and a decision as its output. We can then design the optimal channel behavior for a given situation or assess the optimality of a given channel. Using this framework, we show that a cell's optimal decision-making strategy depends on its prior knowledge about its environment, its goals for the decision, and how much metabolic cost it is willing to “pay” for an accurate decision. Moreover, this framework explains experimentally observed characteristics of biological decision-making pathways such as bet-hedging, hysteresis, and irreversibility as being optimal for various purposes.

2 Results and discussion

2.1 A simple example illustrates the framework

In our generic model of binary decision-making, a cell senses a stimulus X and chooses a decision Y , which can be “high” or “low,” based on the stimulus (Fig. 1a). The stimulus can be any measurable signal from the environment, including but not limited to chemical concentrations, temperature¹⁴, and forces on the cell²⁶. We assume that the probability distribution of stimuli $p_X(x)$ is known; that is, stimulus x occurs with probability $p_X(x)$. For example, consider a cell whose decision to mate is based on the level of pheromone in its environment. The cell senses the level of pheromone (X) and chooses either to mate ($Y = \text{high}$) or not to mate ($Y = \text{low}$). We assume that the probability distribution of pheromone levels $p_X(x)$ is exponential with finite support (Fig. 1b), *i.e.* the environment is more likely to contain lower concentrations of pheromone, and there is zero probability of pheromone concentrations above a certain level.

We characterize the quality of a cellular decision with a distortion function $d(x, y)$. The distortion function describes the goals of a decision-making pathway by quantifying how disadvantageous, or “distorted,” a decision y is in response to a stimulus x . In our example, suppose that when the pheromone concentration is below a threshold x_{th} , the cell should not mate ($Y = \text{low}$); for a concentration greater than x_{th} , the cell should mate ($Y = \text{high}$). To describe the goals of this decision, we use the Hamming distortion function (Fig. 1c):

$$d(x, y) = \begin{cases} & x \leq x_{th} & x > x_{th} \\ y = \text{low} & 0 & 1 \\ y = \text{high} & 1 & 0 \end{cases}$$

That is, if the concentration of pheromone is below x_{th} , the decision not to mate ($Y = \text{low}$) has zero distortion, and the decision to mate ($Y = \text{high}$) carries one (arbitrary) unit of distortion. The opposite is true if the concentration of pheromone is higher than x_{th} .

To quantify the cost of a cellular decision-making strategy, we measure its mutual information $I(X; Y)$, the reduction of uncertainty in the decision Y given the stimulus X . Mutual information, measured in bits, quantifies the degree to which the stimulus influences the decision. A noisy decision-making pathway will have low mutual information, as noise increases the randomness of a decision with respect to the stimulus. A pathway will have higher mutual information when it is designed to suppress noise, which generally requires more complex mechanisms for sensing and signal transduction^{13,14}. Thus, the mutual information between the stimulus and the decision is directly related to the metabolic cost of a decision-making pathway. The advantage of using mutual information over a more direct measure of metabolic cost is that it is mechanism-independent: it quantifies the fidelity of decision-making without specifying how that fidelity is achieved²⁷.

Given a stimulus distribution $p_X(x)$ and distortion function $d(x, y)$, we can compute the rate distortion function $R(D)$ (Materials and Methods), which gives the minimum amount of mutual information between the stimulus and decision (the “information rate”) that a cell must have to make a decision with expected distortion $E[d] = D$. The rate distortion function forms a fundamental bound on the tradeoff between cost (mutual information) and performance (expected distortion). Since less mutual information is required to make decisions with more distortion, $R(D)$ is a decreasing function of D . For our example of the mating decision, Fig. 1d shows the rate distortion function, which gives the minimum mutual information between the pheromone concentration and the mating decision required for the process to have expected distortion $E[d] = D$.

We can assess the optimality of a decision-making strategy by finding how closely its cost-performance point approaches the theoretical bound of the rate distortion function. We represent a decision-making strategy as a conditional probability distribution $p_{Y|X}(y|x)$, which specifies the likelihood that the cell makes decision $Y = y$ when the stimulus is $X = x$. With the distribution of stimuli $p_X(x)$, we can compute a strategy’s expected distortion D and mutual information $I(X; Y)$ (Materials and Methods). We can then measure the distance of the distortion-information point $(D, I(X; Y))$ from the rate distortion function. For example, consider the “suboptimal” strategy () in Fig. 1e as a way of making a mating decision based on measured pheromone concentration. When we compute the expected distortion and mutual information for this strategy (Fig. 1d), we find that it has 43% (0.19 bits) more mutual information than is necessary to achieve its expected distortion. Alternatively, the “suboptimal” strategy has 88% (0.056 units) higher distortion than could optimally be achieved with the same amount of mutual information.

In addition to assessing decision-making strategies, we can use the rate distortion framework to design optimal strategies, *i.e.* those for which the distortion-information point lies on the rate distortion function. The algorithm to compute the rate distortion function (Materials and Methods) also generates a collection of optimal decision-making strategies $p_{Y|X}(y|x)$ with various levels of expected distortion. We simply pick from this collection to find an optimal strategy with expected distortion D . For the example of the mating decision, Fig. 1e shows three optimal strategies ($\circ$, $\triangle$, $\square$) that achieve different levels of expected distortion with minimal mutual information between the pheromone concentration and the mating decision (points in Fig. 1d). In each case, the conditional probability of mating ($Y = \text{high}$) takes one of two values depending on whether the pheromone concentration exceeds the threshold. For distortion requirements above zero, the probability of mating is neither zero nor one, indicating that the optimal strategy involves a certain amount of randomness. Thus, experimentally, we expect to see some cells mating and others not mating in a population exposed to a given level of pheromone. As the allowable distortion decreases and approaches zero, the required mutual information between the pheromone concentration and the mating decision increases, and the optimal strategy approaches that of a perfect switch in which all cells mate or do not mate depending on whether the pheromone concentration exceeds x_{th} (not shown). Fig. 1f shows the probability of error for each optimal strategy. These results do not change qualitatively when a Gaussian stimulus distribution is used instead of an exponential distribution (Fig. S1a–e online).

Random decision-making strategies like that prescribed here have been observed in a number of cellular systems^{28,29}. For a cell, a random strategy is “cheaper” in that it reduces the cost of processing information from the environment. This has been recognized as the optimal strategy for a situation in which making high-fidelity decisions is not worth the cost of the associated sensing equipment³⁰. Such a situation corresponds to the low-information, high-distortion region of the rate distortion function. On the level of populations, random or “bet-hedging” strategies create heterogeneity, which can improve the overall fitness of a population^{31,32}.

2.2 Graded distortion functions lead to graded pathways

The Hamming distortion function strictly penalizes incorrect responses to stimuli around the threshold x_{th} . In practice, this threshold may not be so clear; a cellular system could be forgiven for making either decision in response to a stimulus close to the threshold. To represent this situation, we use a graded distortion function (Fig. 1h). With the same stimulus distribution as before, we computed the rate distortion function $R(D)$ (Fig. 1i) and, for several values of D , the optimal strategy that achieves D with minimal mutual information (Fig. 1j). For each value of D , the optimal conditional probability of deciding “high” is a continuous, sigmoidal function of the stimulus level. To quantify the steepness of these functions, we fitted them with a Hill-type function of the form $f(x) = (x/x_{\text{th}})^n / ((k_{50}/x_{\text{th}})^n + (x/x_{\text{th}})^n)$ (Materials and Methods), where k_{50} is the stimulus level at which the probability of deciding “high” is 50%. Increasing the fidelity of the decision ($D = 0.12 \rightarrow 0.06 \rightarrow 0.02$) yields a sharper transition between the “low” and “high” decisions, with increasing exponents ($n = 4.3 \rightarrow 6.0 \rightarrow 10.0$) and effective thresholds k_{50} closer to the actual threshold x_{th} ($k_{50}/x_{\text{th}} = 1.21 \rightarrow 1.12 \rightarrow 1.06$). Fig. 1k shows the probability of error

for each optimal strategy. As before, these results remain qualitatively the same when a Gaussian distribution of stimuli is used (Fig. S1f–j online).

For the most stringent fidelity requirements (Fig. 1j), stimuli demonstrably above or below x_{th} elicit a nearly deterministic optimal response – the probability of a “high” decision is close to zero or one. However, for stimuli around x_{th} , the optimal response is such that “high” and “low” decisions are both expected to occur with significant probability. Thus, a population of cells following this strategy would demonstrate three types of behavior: At low and high stimulus levels far from the threshold, nearly all cells exhibit “low” and “high” decision phenotypes, respectively, leading to unimodal population distributions. However, for stimulus levels near the threshold, the distribution of decisions should be bimodal, with both “low” and “high” decision phenotypes observed in the population. This strategy has been seen experimentally in several cell fate decision processes. For example, the steady-state activity of JNK, a MAP kinase involved in cellular stress responses, is high in Jurkat T cells treated with a high concentration of PMA (10 nM) and low in untreated cells. In a population of cells treated with an intermediate concentration of PMA (0.5 nM), both activity levels are observed, and individual cells are found in either the high or low response group³³. A similar bimodal population response has been seen in the progesterone-induced maturation of *Xenopus laevis* oocyte extracts exposed to intermediate concentrations of progesterone (10–100 nM). In this case, the conditional probability of a “high” decision (high MAPK phosphorylation) in a given oocyte was observed to be a graded function of the stimulus (progesterone concentration), with exponent $n \approx 1$ ³⁴.

2.3 Different *a priori* assumptions lead to hysteretic pathways

So far we have assumed that the stimuli are more likely to be below the threshold x_{th} than above it. We now consider how stimuli more likely to be above x_{th} affect the cell’s optimal decision-making strategy. For both “likely high” and “likely low” distributions of stimuli (Fig. 2a, e, i), we computed $R(D)$ (Fig. 2c, g, k) and the optimal strategy achieving $D = 0.02$ (Fig. 2d, h, l). When the stimulus is “likely high,” the optimal conditional probability of deciding “high” is a sigmoidal function of the stimulus level, much like that for a “likely low” stimulus distribution. However, the transition point of the decision occurs at a stimulus level below x_{th} . The intermediate region, in which different optimal responses are predicted depending on whether the stimulus distribution is “likely low” or “likely high,” becomes wider as the distributions become more disjoint.

The difference in optimal responses for the “likely low” and “likely high” stimulus distributions shows how prior knowledge about the stimulus affects the optimal response strategy. Given the same stimulus, cells conditioned to “low” stimuli are more reluctant to make a “high” decision than cells conditioned to “high” stimuli. This optimal strategy is implemented as a decision-making pathway with hysteresis. Cells that begin in the “low” decision state conditioned to a “likely low” distribution of stimuli will transition to the “high” decision state given a sufficiently large stimulus. At this point, the cells are conditioned to the “likely high” distribution of stimuli, having “learned” from the stimulus they just saw, and a much lower stimulus is required to send the cells back to the “low” decision state. One example of a pathway like this is the *lac* operon in *E. coli*, which

determines whether cells will express genes necessary to import and metabolize lactose. Cells grown in 1 mM TMG to induce expression of the *lac* gene required treatment with TMG below 3 μ M to fully turn off gene expression, while cells grown without TMG required treatment with TMG above 30 μ M to fully turn on gene expression⁸.

2.4 Asymmetric distortion functions lead to irreversible pathways

The distortion functions we have been using so far penalize incorrect “low” and “high” decisions equally. In real systems, a false “high” decision may be more disadvantageous than a false “low” decision or vice versa. To analyze this scenario, we constructed a distortion function that penalizes a false “low” decision more heavily than a false “high” decision (Fig. 2n). Using this distortion function, we computed $R(D)$ (Fig. 2o) and the optimal strategy achieving $D = 0.02$ (Fig. 2p) for both “likely low” and “likely high” stimulus distributions. The optimal response curves are no longer symmetric. While a “likely low” stimulus distribution leads to an optimal response similar to that seen before, a “likely high” stimulus leads to an optimal strategy of always choosing “high” regardless of the stimulus. The lower distortion for an incorrect “high” decision pushes the rate distortion function for the “likely high” stimulus to the left – that is, a given expected distortion can be achieved with less mutual information between the stimulus and decision. In this case, $D = 0.02$ can be achieved with no mutual information at all, while more mutual information and a more complex decision-making pathway are required to achieve the same D when the stimulus is “likely low.”

A population of cells following this optimal strategy would demonstrate irreversible behavior: Cells conditioned to “likely low” stimuli remain in the “low” decision state until they experience a sufficiently large stimulus. When the population makes the transition to the “high” decision state, the cells become conditioned to the “likely high” distribution of stimuli, and even a low stimulus cannot reverse the “high” decision. This irreversibility is seen in many cell cycle transitions. For example, when *Xenopus* oocytes were incubated in 600 nM progesterone to induce maturation, washing the cells for 10 h to remove progesterone could not undo their commitment to maturation³⁵. Similarly, the process of apoptosis, or programmed cell death, in multicellular organisms is known to be irreversible, helping to prevent uncontrolled cell proliferation³. From a rate distortion perspective, irreversible cell transitions arise because it is worse for cells to not transition when they should (false “low” decision) than to transition when they shouldn’t (false “high” decision).

2.5 Finding the goals of an optimized pathway

So far we have postulated the form of distortion functions to design optimal decision-making strategies with the rate distortion framework. One could think of evolution as performing a similar task – designing optimal pathways whose goals are characterized by a particular distortion function. This raises the question: for what distortion function, or quantified set of goals, is a particular decision-making pathway optimized?

As a test case, we looked at the process of extrinsically triggered apoptosis, or programmed cell death, in multicellular organisms. In this decision-making pathway, stimulation of death receptors leads to the formation of a death-inducing signaling complex, which activates

3 Materials and methods

3.1 Mutual information

The entropy $H(Y)$ of the random variable Y , which represents the decision, quantifies the uncertainty in the decision²⁴:

$$H(Y) = - \sum_y p_Y(y) \log_2 p_Y(y) \quad (1)$$

The mutual information $I(X;Y)$ between the stimulus X and the decision Y is the reduction of uncertainty in Y given X ²⁴:

$$I(X;Y) = H(Y) - H(Y|X) = \sum_x \sum_y p_{Y|X}(y|x) p_X(x) \log_2 \frac{p_{Y|X}(y|x)}{p_Y(y)} \quad (2)$$

3.2 Computing the rate distortion function

Computation of $R(D)$ can be performed using the method of Lagrange multipliers²³. We solve the problem computationally using the Blahut-Arimoto algorithm²³. The basic steps are as follows:

1. Initialize the probability distribution of the decision Y , $p_Y(y)$, as a uniform distribution.
2. Given $p_Y(y)$, compute the conditional probability distribution $p_{Y|X}(y|x)$ that minimizes the mutual information $I(X;Y)$ while satisfying the distortion constraint:

$$p_{Y|X}(y|x) = \frac{p_Y(y) e^{-\lambda d(x,y)}}{\sum_y p_Y(y) e^{-\lambda d(x,y)}} \quad (3)$$

3. Given $p_{Y|X}(y|x)$, compute the marginal probability distribution $p_Y(y)$ that minimizes the mutual information $I(X;Y)$:

$$p_Y(y) = \sum_x p_X(x) p_{Y|X}(y|x) \quad (4)$$

4. Repeat steps 2 and 3 until $p_{Y|X}(y|x)$ and $p_Y(y)$ converge.

As these steps are carried out, the limiting mutual information $I(X;Y)$ has been shown to be $R(D)$, where $E[d] = D$ is determined by the Lagrange multiplier λ chosen in the minimization. Repeating this process with different values of λ generates a collection of optimal strategies $p_{Y|X}(y|x)$ for which the distortion-information point $(D, I(X;Y))$ lies on $R(D)$.

3.3 Computing the distortion function from optimal probability characteristics

If the probability characteristics $p_{Y|X}(y|x)$ of a given decision-making pathway are optimal, *i.e.* they minimize the mutual information $I(X;Y)$ subject to a distortion constraint, then they satisfy (3) and (4) above. We can use this fact to compute the distortion function around which a given pathway $p_{Y|X}(y|x)$ is optimized. We begin by looking at (3) with $y = \text{high}$:

$$p_{Y|X}(y=\text{high}|x) = \frac{p_Y(y=\text{high})e^{-\lambda d(x,y=\text{high})}}{p_Y(y=\text{low})e^{-\lambda d(x,y=\text{low})} + p_Y(y=\text{high})e^{-\lambda d(x,y=\text{high})}} \quad (5)$$

Algebraic rearrangement of (5) yields:

$$d(x, y=\text{high}) = \frac{1}{\lambda} \log \left[\frac{p_{Y|X}(y=\text{low}|x)p_Y(y=\text{high})}{p_{Y|X}(y=\text{high}|x)p_Y(y=\text{low})} \right] + d(x, y=\text{low}) \quad (6)$$

For consistency, we pick a function $d(x, y = \text{low}) = g(x)$ such that $\min[d(x, y = \text{low}), d(x, y = \text{high})] = 0$ for all x . The choice of scaling factor λ is arbitrary and affects the scale of the distortion function but not the shape; we used $\lambda = 1$ in our computations.

3.4 Hill-function fitting

The optimal conditional probability curves for a “high” decision were fitted with a Hill-type function of the form $f(x) = (x/x_{\text{th}})^n / ((k_{50}/x_{\text{th}})^n + (x/x_{\text{th}})^n)$ by minimizing the mean squared distance between the curves and the function using the `fminsearch` command in MATLAB (MathWorks, Natick, MA).

3.5 Stochastic simulations

To estimate the conditional probability of apoptosis as a function of initial C8* levels, we performed stochastic simulations of the model of externally triggered apoptosis developed by Eissing and colleagues^{2,36}. Using the next reaction algorithm³⁷ running in MATLAB (MathWorks, Natick, MA), we simulated the model 5000 times for each of 81 initial C8* levels uniformly distributed between 300 and 1100 molecules. Published values were used for all model parameters and all initial conditions other than that of C8*^{2,36}. Each simulation terminated when apoptosis had occurred, as measured by active caspase 3 reaching at least 8000 molecules, or when 24 hours of simulation time had passed. Results were consistent with those previously published³⁶. For each initial level of C8*, we estimated the conditional probability of apoptosis as the fraction of these simulations that ended in apoptosis.

4 Conclusion

In biology much attention is focused on mechanisms: what is the mechanism by which a biological system works, or what behavior is produced by a given mechanism? The rate distortion framework, by contrast, addresses questions about goals: what is the goal toward which a biological system works, or what system behavior best achieves a given goal? The distortion function provides an elegant way of quantifying the goals of a binary decision-

making process. Because of stochastic fluctuations in cellular systems, it is impossible or impractical to achieve these goals perfectly, so it is important to understand how they may be achieved imperfectly in the most efficient way possible. The rate distortion function, derived from the distortion function and the distribution of stimuli, provides a fundamental limit on how well a process can achieve its goals at a given cost in mutual information. Furthermore, the process of calculating the rate distortion function generates optimal decision-making strategies that achieve this fundamental limit. The beauty of this approach is that the results are independent of the biological mechanism by which such a strategy is implemented²⁷. While we have tailored the techniques described here to understand binary decision-making based on a single stimulus, the framework could be easily extended to handle multiple stimuli by replacing the random variable X with a random vector $\mathbf{X}$ and modifying the mathematics accordingly. Moreover, rate distortion theory can be recast to consider more complex cellular processes¹⁵.

Practically, the rate distortion framework can be used for both the design and analysis of decision-making systems with respect to their goals. Experimenting with rate distortion as a design tool shows that certain observed characteristics of binary decision-making pathways, including bet-hedging, hysteresis, and irreversibility, are optimal in an information-theoretic sense to achieve various goals. In the context of synthetic biological design, the framework can be used to determine what decision-making strategy optimally meets a decision fidelity requirement when responding to a noisy stimulus. As an analysis tool, rate distortion can quantify the goals of a pathway assumed to be optimal or evaluate the optimality of a pathway with known goals. For example, we can regard the apoptosis pathway as a system that has been optimized to achieve the goals described by the distortion function in Fig. 3c. While this analysis requires knowing the probability distribution of stimuli encountered by the pathway, an estimate of the stimulus distribution can produce an approximation of the distortion function around which the pathway is optimized.

The rate distortion framework suggests a way of thinking about the evolution of cellular decision-making strategies. One impetus for evolution is the need for efficiency: organisms that can make decisions just as well using less information, and therefore less energy, can out-compete their neighbors. This corresponds to a pathway's distortion-information point moving closer to the rate distortion function. This sort of evolution might occur in response to a change in the distribution of stimuli in an organism's environment or a change in the goals of the pathway as defined by the distortion function. Either change would move the rate distortion function and the distortion-information point of the pathway, creating a need for optimization. A second impetus for evolution is the need for better performance: as an organism and its environment increase in complexity, more accurate decision-making yields an increasing advantage. In this case, lower distortion is worth its cost in information, and a pathway's evolution from right to left along the rate distortion function becomes favorable. This idea of evolution in response to the changing value of low distortion has been proposed to explain the emergence of molecular codes²⁵.

Perkins and Swain²⁹ characterized cellular decision-making as having three main tasks: a cell must 1) estimate the state of its environment by sensing stimuli; 2) make a decision informed by the consequences of the alternatives; and 3) perform these functions in a way

that maximizes the fitness of the population. The rate distortion framework provides a complementary perspective on decision-making, regarding these three tasks as a single process. The distortion function defines accurate sensing (task 1) by how heavily it penalizes small mistakes, and it quantifies the disadvantages of alternative decisions (task 2). The expected distortion describes how accurate sensing must be (task 1) and how much disadvantage the cell can afford in making a decision (task 2). The resulting optimal strategies fulfill task 3 by making choices with decisiveness proportional to the information available. When there is little information, the optimal strategy incorporates randomness to generate biological variation, which improves the fitness of a population in an uncertain environment. With more information, the optimal strategy eschews randomness and simply makes a better decision. By rolling several facets of decision-making into one overall process, the rate distortion framework enables design and evaluation of that process with a fundamental optimality criterion.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Sanjeev Khudanpur, Andre Levchenko, and Saurabh Paliwal for fruitful discussions. This work was supported by NIGMS grant 72024.

References

1. Fussenegger M, Bailey JE, Varner J. *Nat. Biotechnol.* 2000; 18:768–774. [PubMed: 10888847]
2. Eissing T, Conzelmann H, Gilles ED, Allgöwer F, Bullinger E, Scheurich P. *J. Biol. Chem.* 2004; 279:36892–36897. [PubMed: 15208304]
3. Legewie S, Blüthgen N, Herzog H. *PLoS Comput. Biol.* 2006; 2:e120+. [PubMed: 16978046]
4. Ferrell JE, Pomeroy JR, Kim SY, Trunnell NB, Xiong W, Huang C-YF, Machleder EM. *FEBS Lett.* 2009; 583:3999–4005. [PubMed: 19878681]
5. Colman-Lerner A, Gordon A, Serra E, Chin T, Resnekov O, Endy D, Pesce CG, Brent R. *Nature.* 2005; 437:699–706. [PubMed: 16170311]
6. Paliwal S, Iglesias PA, Campbell K, Hilioti Z, Groisman A, Levchenko A. *Nature.* 2007; 446:46–51. [PubMed: 17310144]
7. Novick A, Weiner M. *Proc. Natl. Acad. Sci. U. S. A.* 1957; 43:553–566. [PubMed: 16590055]
8. Ozbudak EM, Thattai M, Lim HN, Shraiman BI, van Oudenaarden A. *Nature.* 2004; 427:737–740. [PubMed: 14973486]
9. Mettetal JT, Muzzey D, Pedraza JM, Ozbudak EM, van Oudenaarden A. *Proc. Natl. Acad. Sci. U. S. A.* 2006; 103:7304–7309. [PubMed: 16648266]
10. Julius AA, Halász A, Sakar MS, Rubin H, Kumar V, Pappas GJ. *IEEE Trans. Autom. Control.* 2008; 53:51–65.
11. Dekel E, Alon U. *Nature.* 2005; 436:588–592. [PubMed: 16049495]
12. Paulsson J. *Nature.* 2004; 427:415–418. [PubMed: 14749823]
13. Lauffenburger DA. *Proc. Natl. Acad. Sci. U. S. A.* 2000; 97:5031–5033. [PubMed: 10805765]
14. El-Samad H, Kurata H, Doyle JC, Gross CA, Khammash M. *Proc. Natl. Acad. Sci. U. S. A.* 2005; 102:2736–2741. [PubMed: 15668395]
15. Andrews BW, Iglesias PA. *PLoS Comput. Biol.* 2007; 3:e153+. [PubMed: 17676949]
16. Ziv E, Nemenman I, Wiggins CH. *PLoS ONE.* 2007; 2:e1077+. [PubMed: 17957259]
17. Tostevin F, ten Wolde PR. *Phys. Rev. Lett.* 2009; 102:218101+. [PubMed: 19519137]

18. Mehta P, Goyal S, Long T, Bassler BL, Wingreen NS. *Mol. Sys. Biol.* 2009; 5:325.
19. Hu B, Chen W, Levine H, Rappel W-J. *J. Stat. Phys.* 2011; 142:1167–1186. [PubMed: 21643513]
20. Tkacik G, Callan CG, Bialek W. *Proc. Natl. Acad. Sci. U. S. A.* 2008; 105:12265–12270. [PubMed: 18719112]
21. Tkacik G, Walczak AM. *J. Phys.: Condens. Matter.* 2011; 23:153102+. [PubMed: 21460423]
22. Shannon, CE. Coding theorems for a discrete source with a fidelity criterion. Machol, RE., editor. New York: McGraw-Hill; 1960. p. 93-126.
23. Berger, T. Rate Distortion Theory: A Mathematical Basis for Data Compression. Englewood Cliffs, NJ: Prentice-Hall; 1971.
24. Cover, TM.; Thomas, JA. Elements of Information Theory. 2nd edn.. Wiley-Interscience; 2006.
25. Tlusty T. *Phys. Life Rev.* 2010; 7:362–376. [PubMed: 20558115]
26. Engler AJ, Sen S, Sweeney HL, Discher DE. *Cell.* 2006; 126:677–689. [PubMed: 16923388]
27. Mian IS, Rose C. *Integr. Biol.* 2011; 3:350–367.
28. Veening JW, Smits WK, Kuipers OP. *Annu. Rev. Microbiol.* 2008; 62:193–210. [PubMed: 18537474]
29. Perkins TJ, Swain PS. *Mol. Sys. Biol.* 2009; 5:326.
30. Kussell E, Leibler S. *Science.* 2005; 309:2075–2078. [PubMed: 16123265]
31. Kærn M, Elston TC, Blake WJ, Collins JJ. *Nat. Rev. Genet.* 2005; 6:451–464. [PubMed: 15883588]
32. Rao CV, Wolf DM, Arkin AP. *Nature.* 2002; 420:231–237. [PubMed: 12432408]
33. Bagowski CP, Besser J, Frey CR, Ferrell JE. *Curr. Biol.* 2003; 13:315–320. [PubMed: 12593797]
34. Ferrell JE, Machleder EM. *Science.* 1998; 280:895–898. [PubMed: 9572732]
35. Xiong W, Ferrell JE. *Nature.* 2003; 426:460–465. [PubMed: 14647386]
36. Eißing T, Allgöwer F, Bullinger E. *IEE Proc. Syst. Biol.* 2005; 152:221–228.
37. Gibson MA, Bruck J. *J. Phys. Chem. A.* 2000; 104:1876–1889.

Fig. 1.

Optimal decision-making is specified by the distortion function and stimulus distribution. (a) The generic model for binary decision-making involves a stimulus X that is sensed and transmitted through a signaling pathway, leading to a decision Y , which can be “low” or “high.” The pathway implements a decision-making strategy represented as the conditional probability of a specific decision y given a specific stimulus x . (b–f) The probability distribution of stimuli (b) and the Hamming distortion function (c) are used to compute the rate distortion function (d) and optimal decision-making strategies (e). The distortion function specifies how disadvantageous a given decision is in response to a given stimulus. The rate distortion function $R(D)$ specifies, in units of bits, the minimum amount of mutual information between the stimulus X and the decision Y required for a decision-making strategy to have expected distortion $E[d] = D$. For each decision-making strategy in (e), we can compute a distortion-information point (d) to evaluate its optimality with respect to the rate distortion function. The probability of error for the three optimal strategies ($\circ$, $\triangle$, $\square$) is shown in (f). (g–k) Using a graded distortion function (h) leads to a slightly different rate distortion function (i) and optimal strategies in which the probability of a “high” decision is a sigmoidal function of the stimulus level (j). The probability of error for these optimal strategies is shown in (k).

Fig. 2.

Complementary “likely low” and “likely high” stimulus distributions lead to hysteretic and irreversible optimal responses. (a–l) As the “likely low” and “likely high” stimulus distributions become more disjoint (a → e → i), the optimal response curves grow further apart (d → h → l). Such an optimal response would be implemented as a hysteretic pathway. (m–p) An asymmetric distortion function that penalizes an incorrect “low” response more than an incorrect “high” response (n) produces asymmetric optimal response curves for the “likely low” and “likely high” stimulus distributions (p). This type of response would be implemented as an irreversible decision-making pathway.

Fig. 3.

Rate distortion theory can be used to analyze the strategies implemented by cellular decision-making pathways. (a–c) Given a distribution of stimuli (a), for a decision-making pathway such as that which controls apoptosis (b), we can find the distortion function around which that pathway is optimized (c), which quantifies the goals of the pathway. The “jumps” in the distortion function are an artifact that arises when probabilities are estimated from a finite number of simulations. (d–e) For a given set of goals, we can compute the rate distortion function (d), which describes the information cost of low distortion. Points along

this curve correspond to different optimal strategies (e), forming a trajectory for the evolution of a pathway to better achieve a fixed set of goals. The probability of error for these optimal strategies is shown in (f).