

MEMOIR ON INFORMATION OF CRN

FILIPPE

1. CASE OF STUDY (MICHAELIS-MENTEN MECHANISM)

The mechanism is



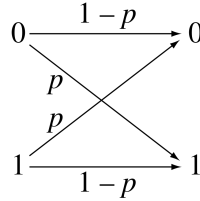
- We assume S and P to be chemostated (fixed concentration/copy numbers).
- There's a single molecule of each enzyme (E) or enzyme complexed (ES).
- The system has 2 states, either in E or ES .

1.1. Posing the problem.

- What is information in the scenario of the enzyme processing?
- Can this CRN be seen as an information channel?

2. INSIGHTS FROM INFORMATION THEORY

A similar system is the Binary Symmetric Channel (BSC), which we'll try to construct an analogy with.



2.1. Binary Symmetric Channel. Suppose a channel that transmits one single bit per usage (cycle).

- The probability of sending either 0 or 1 and receiving the same thing is

$$Pr\{Y = 0|X = 0\} = Pr\{Y = 1|X = 1\} = 1 - p.$$

- Otherwise we obtain an error in the message sent

$$Pr\{Y = 1|X = 0\} = Pr\{Y = 0|X = 1\} = p.$$

(!) The channel is called symmetric because there's no preference for an error or a message to be sent (equiprobable events).

The BSC models the behavior of a channel which transmits the output of a system with two states. Let's define

- **(System/Source)**: It has two states $X = 0, 1$. Assuming the probability of them being equal¹:

$$Pr\{X = x\} = 1/2, x \in X.$$

- The Shannon entropy of the source is

$$\mathcal{H}(X) = - \sum_{x \in X} Pr\{X = x\} \log Pr\{X = x\}.$$

If we assume base 2 for the log we obtain that $\mathcal{H}(X) = 1$ bit. That is, we expect to obtain 1 bit of information by measuring the system.

- The base 2 was chosen because the phase space of the system has size $|X| = 2$.²

- **(Channel)**: It transmits the information from the source. If the channel has a Markovian property, its current state only depends in the previous state, e.g. the BSC.

- The Mutual Information quantifies the information expected to be acquired about the system X or the output of the channel Y from the observation of Y or X

$$\mathcal{I}(X; Y) = D_{KL}(Pr\{X, Y\} || Pr\{X\}Pr\{Y\})$$

$$= \sum_{x \in X, y \in Y} Pr\{X = x, Y = y\} \log \left(\frac{Pr\{X = x, Y = y\}}{Pr\{X\}Pr\{Y\}} \right)$$

being $Pr\{X = x, Y = y\} = Pr\{Y = y | X = x\}Pr\{X = x\}$.

- If the channel is the BSC, the mutual information is

$$\mathcal{I}(X; Y) = \log 2 + (1 - p) \log(1 - p) + p \log p.$$

- For $p = 0$ there's no error and the message gets through as well as transmitted.
- For $p = 1$, we assume that it is known the channel always “flip” the bit, thus the information at the end of the channel is known, but flipped.
- $\mathcal{I}(X; Y)$ is minimized for $p = 1/2$ and the expected information to be acquired with the measurement is null.

3. WHAT MEANS INFORMATION IN THE SCENARIO OF THE MICHAELIS-MENTEN MECHANISM?

We know the system can be in the states $X = \{E, ES\}$.

- The reaction of the system will transform (or not) the species X into something else.
- We can understand this as the action of the channel over the message. The channel can flip (or not) the bit as the reaction can switch (or not) the enzyme to the complexed enzyme.

¹We may also say the system splits the *phase space* in two, this might be useful later.

²What would happen if $|X| \rightarrow \infty$?

Translating this into probabilities, we understand that:

- The system can be in $X = \{E, ES\}$.
- A reaction occurs, and the states at the end of the reaction are $Y = \{E, ES\}$.

Let's suppose that τ is what is known as the *time for the next reaction*, i.e. the time in which only a single reaction fires.

- Given the system is initially at X_0 , when a reaction fires the system goes to a state X_τ .
- The probability of the system being in a state $\{E, ES\}$ after the reaction is given by $Pr\{X_\tau = x | X_0 = x\}$.
- The solution of the Chemical Master Equation gives us the transition probability of an arbitrary time t after the initial state

$$\frac{\partial}{\partial t} Pr\{X_t = x | X_0 = x\} = \text{TO-DO} \Rightarrow Pr\{X_t | X_0\} = Pr\{X_0 | X_0\} e^{Gt},$$

where G is the *generator matrix*.

3.1. Information of the Michaelis-Menten. Once posed the mathematics, we are in position to question what the calculations give us.

- We can obtain the mutual information about the states of the enzyme at the beginning and after a time t .
- This will give us the expected information to be gained about the initial state by measuring the system at a time t , or *vice-versa*. Then

$$\mathcal{I}(X_0; X_t) = \sum_{x \in X_0, y \in X_t} Pr\{X_0 = x, X_t = y\} \log \left(\frac{Pr\{X_0 = x, X_t = y\}}{Pr\{X_0 = x\} Pr\{X_t = y\}} \right),$$

where $Pr\{X_t = y, X_0 = x\} = Pr\{X_t = y | X_0 = x\} Pr\{X_0 = x\}$ and $Pr\{X_t = y\} = \sum_{x \in X_0} Pr\{X_t = y, X_0 = x\}$.

3.1.1. Evaluation of the quantities. Let's define:

- The transition matrix $[P]_{ij} = Pr\{X_t = j | X_0 = i\}$;
- The probability of the initial state $[p_0]_i = Pr\{X_0 = i\}$;
- The above gives us the joint probability $Pr\{X_t = j, X_0 = i\} = [\text{diag}(p_0)P]_{ij}$.
- The marginalization over X_0 is:

$$Pr\{X_t = j\} = \left[\mathbb{1}^\top \text{diag}(p_0)P \right]_j = \left[p_0^\top P \right]_j, \quad \mathbb{1} = [1, 1, \dots, 1]^\top.$$

Recap the probability transition of the continuous-time Markov chain (CTMC). Then:

- The transition probability matrix for a time t as $[P_t]_{ij} = Pr\{X_t = j | X_0 = i\}$;
- P_t is obtained by the solution of an ODE, being $P_t = e^{Gt}$.

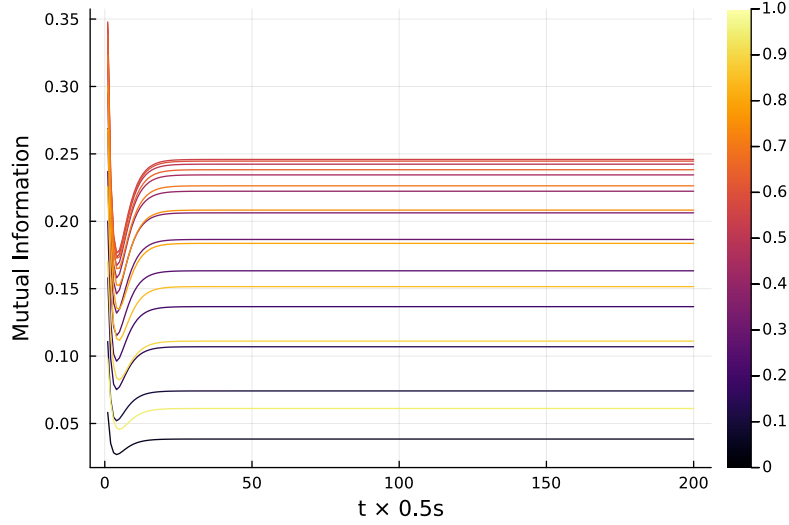
TO-DO

3.1.2. *Channel capacity.* It is defined as the probability distribution of X_0 , say $Pr\{X_0\}$, which maximizes the expected information to be acquired from the observation of X_t , which reads

$$C = \max_{Pr\{X_0\}} \mathcal{I}(X_0; X_t).$$

- **(BSC):** Which $Pr\{X\}$ maximizes the mutual information for the BSC?
 - Let's assume that $Pr\{X = 0\} = f$ and $Pr\{X = 1\} = 1 - f$. If the BSC has an error probability of $p = 0.15$, then we can verify that $\mathcal{I}(X; Y)$ is maximized for $f = 1/2$, the uniform distribution (see D.J.C. MacKay).

In Information Theory, this definition makes sense to help us in finding a coding for the source which mitigates the errors of the channel. Such coding scheme has a probability distribution over its words $Pr\{X\}$. In the case of reaction channels we would have initial distributions which would mitigate the information lost about the initial state when reacting³. In the figure below it is plotted how $\mathcal{I}(X_0; X_t)$ varies with f for $Pr\{X_0 = E\} = f$ and $Pr\{X_0 = ES\} = 1 - f$. By inspection we note that C is maximized when $f \rightarrow 1/2$.



This is equivalent to the maximum entropy of the source $\mathcal{H}(X_0)$.

³Think about the definition of the mutual information in terms of the KL-divergence, maximizing the later is the same as increase the “difference” in a KL sense between the joint probability of (X_0, X_t) and the product of their marginals, what would imply that they are not independent.