九州大学学術情報リポジトリ
Kyushu University Institutional Repository

Maxwell’s Demon: More about MATH

Narikiyo, Osamu
Department of Physics, Faculty of Sciences, Kyushu University : Associate Professor

https://hdl.handle.net/2324/7157980

出版情報：2023-11-24
バージョン：
権利関係：

KYUSHU UNIVERSITY

Maxwell's Demon: More about MATH

Osamu Narikiyo

Department of Physics, Kyushu University

We discuss more about the MATH part of the digest on Maxwell's demon (https://hdl.handle.net/2324/7157421).

Let us consider the probability distributions $\vec{p} = (p_1, p_2, p_3, \cdots)$ for the system and $\vec{q} = (q_1, q_2, q_3, \cdots)$ for the demon. They are the marginals of the joint probability distribution $\hat{P} = \{P_{ij}\}$: $p_i = \sum_j P_{ij}$ and $q_j = \sum_i P_{ij}$. In terms of the conditional probabilities $p_{i/j}$ and $q_{j/i}$, p_i, q_j and P_{ij} are related as $P_{ij} = p_{i/j} \cdot q_j = q_{j/i} \cdot p_i$.

The entropies for $\vec{p}$, $\vec{q}$ and $\hat{P}$ are given by $H(\vec{p}) = -\sum_i p_i \log p_i$, $H(\vec{q}) = -\sum_j q_j \log q_j$ and $H(\hat{P}) = -\sum_{ij} P_{ij} \log P_{ij}$.

Now let us consider the case where the demon gets some information on the system by some measurement. If the measurement outcome for the demon is j, then the probability of i-state of the system becomes $p_{i/j}$ for the demon. Before the measurement the expectation value of the unknown information on the system for the demon is $H(\vec{p})$. After the measurement it reduces to $H_j = -\sum_i p_{i/j} \log p_{i/j}$. Averaging H_j over measurements the expectation value of the unknown information becomes $\overline{H} = \sum_j q_j H_j$. Since

$$\overline{H} = -\sum_{ij} P_{ij} \log \frac{P_{ij}}{q_j} = H(\hat{P}) - H(\vec{q})$$

the additivity* (chain rule) of the entropies holds:

$$H(\hat{P}) = H(\vec{q}) + H(\hat{P}/\vec{q})$$

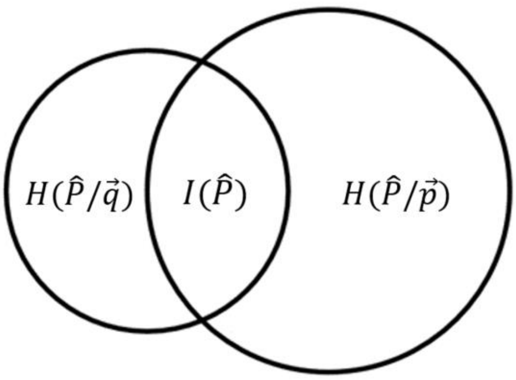
where $H(\hat{P}/\vec{q}) \equiv \bar{H}$. By exchanging the roles of $\vec{p}$ and $\vec{q}$ the chain rule becomes $H(\hat{P}) = H(\vec{p}) + H(\hat{P}/\vec{p})$. Here $H(\hat{P}/\vec{p}) = \sum_i p_i H_i$ with $H_i = -\sum_j q_{j/i} \log q_{j/i}$. Since the information gain** of the demon $I(\hat{P})$ is given by $I(\hat{P}) = H(\vec{p}) - \bar{H}$, we obtain

$$I(\hat{P}) = H(\vec{p}) + H(\vec{q}) - H(\hat{P}).$$

Moreover, $I(\hat{P})$ is symmetric under the exchange of the roles of $\vec{p}$ and $\vec{q}$:

$$I(\hat{P}) = H(\vec{p}) - H(\hat{P}/\vec{q}) = H(\vec{q}) - H(\hat{P}/\vec{p}).$$

Equivalently $H(\vec{p}) = I(\hat{P}) + H(\hat{P}/\vec{q})$ and $H(\vec{q}) = I(\hat{P}) + H(\hat{P}/\vec{p})$:



where the area of the left circle is $H(\vec{p})$ and the area of the right circle is $H(\vec{q})$. The total area of the jointed circles is $H(\hat{P}) = H(\vec{p}) + H(\hat{P}/\vec{p}) = H(\vec{q}) + H(\hat{P}/\vec{q})$.

* Since $P_{ij} = q_j \cdot p_{i/j}$, the following additivity holds for the information:

$$-\log P_{ij} = -\log q_j - \log p_{i/j}.$$

Namely, the multiplication of the probabilities leads to the addition of the information. By averaging this equation with the probability distribution $\{P_{ij}\}$, we obtain $H(\hat{P}) = H(\vec{q}) + H(\hat{P}/\vec{q})$.

** $I(\hat{P})$ is known as the mutual information. The probability for the pair of events (i, j) is P_{ij} and the unknown information for the pair is $-\log P_{ij}$. If the events are dependent, the total number of the allowed pair events is reduced so that $P_{ij} > p_i \cdot q_j$ where $p_i \cdot q_j$ is the probability for the independent pair. This inequality for the probabilities is translated into the following information inequality

$$-\log P_{ij} < -\log p_i - \log q_j.$$

By averaging this equation with the probability distribution $\{P_{ij}\}$, we obtain $H(\hat{P}) < H(\vec{p}) + H(\vec{q})$. Thus, we can say $H(\vec{p}) + H(\vec{q}) = H(\hat{P}) + I(\hat{P})$ with $I(\hat{P}) > 0$. Since $H(\hat{P}) = H(\vec{q}) + H(\hat{P}/\vec{q})$, we obtain $H(\vec{p}) = I(\hat{P}) + H(\hat{P}/\vec{q})$. Namely, $H(\vec{p})$ consists of the dependent information $I(\hat{P})$ and the independent information $H(\hat{P}/\vec{q})$. By the same way $H(\vec{q}) = I(\hat{P}) + H(\hat{P}/\vec{p})$. $I(\hat{P})$ is common to the system and the demon because it reflects the character of the pairs (i, j). In other words, the measurement of i and the measurement of j for a dependent pair give us the same information.

$H(\hat{P}) < H(\vec{p}) + H(\vec{q})$ is equal to $H(\hat{P}/\vec{q}) < H(\vec{p})$, since $H(\hat{P}) = H(\vec{q}) + H(\hat{P}/\vec{q})$. Thus, another explanation of $H(\hat{P}) < H(\vec{p}) + H(\vec{q})$ can be carried out by the explanation of $H(\hat{P}/\vec{q}) < H(\vec{p})$. Putting aside the mathematical proof, the meaning of $H(\hat{P}/\vec{q}) < H(\vec{p})$ is evident: unknown information on the system reduces after the measurement.