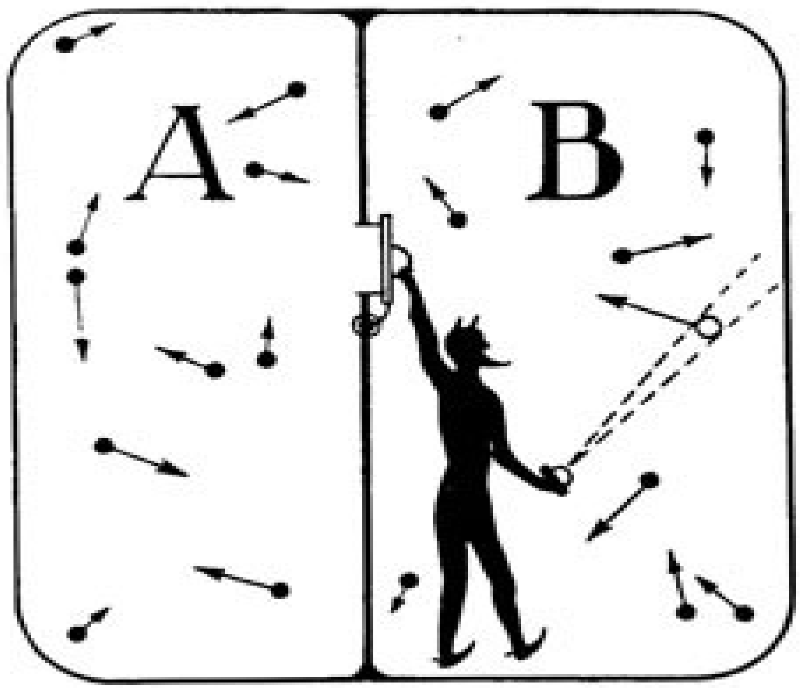
Information and Physics
Landauer Principle and Beyond

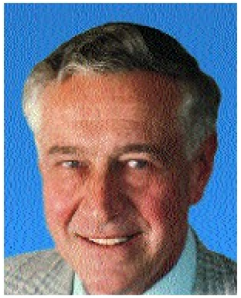
Ryoichi Kawai

Department of Physics
University of Alabama at Birmingham

Maxwell Demon



Lerner, 1975



Ralf Landauer
(1929-1999)

Landauer principle

"Computational steps which inevitably require a minimal *energy dissipation* are those which *discard information*."
(Landauer, 1961)

"Any logically irreversible manipulation of *information*, such as the *erasure* of a bit or the merging of two computation paths, must be accompanied by a corresponding *entropy increase* in non-information bearing degrees of freedom of the information processing apparatus or its environment."
(Bennett, 1982, Wikipedia)

The *erasure* of one bit of *information* is necessarily accompanied by a *dissipation* of at least $kT \ln 2$ heat.

Key words: information, erasure, dissipation

Confusion

If one were to imagine starting with those degrees of freedom in a thermalised state, there would be a real reduction in thermodynamic entropy if they were then re-set to a known state. This can only be achieved under information-preserving microscopically deterministic dynamics if the uncertainty is somehow dumped somewhere else — i.e. if the entropy of the environment (or the non information-bearing degrees of freedom) is increased by at least an equivalent amount, as required by the Second Law, by gaining an appropriate quantity of heat: specifically $kT \ln 2$ of heat for *every 1 bit of randomness erased*.
(Wikipedia, Information Theory and Physics)

Landauer's principle for information erasure is valid for a symmetric double-well potential, but not for an asymmetric one. The proof of

$$W_{\mathrm{erase}} \geq kT \ln 2$$

using statistical mechanics is valid only for the symmetric case.
(Sagawa and Ueda, 2009)

Key word: Dissipation

Which one is correct?

$$W \geq k\,T \ln 2$$

$$Q \geq k\,T \ln 2$$

$$\Delta S \geq k \ln 2$$

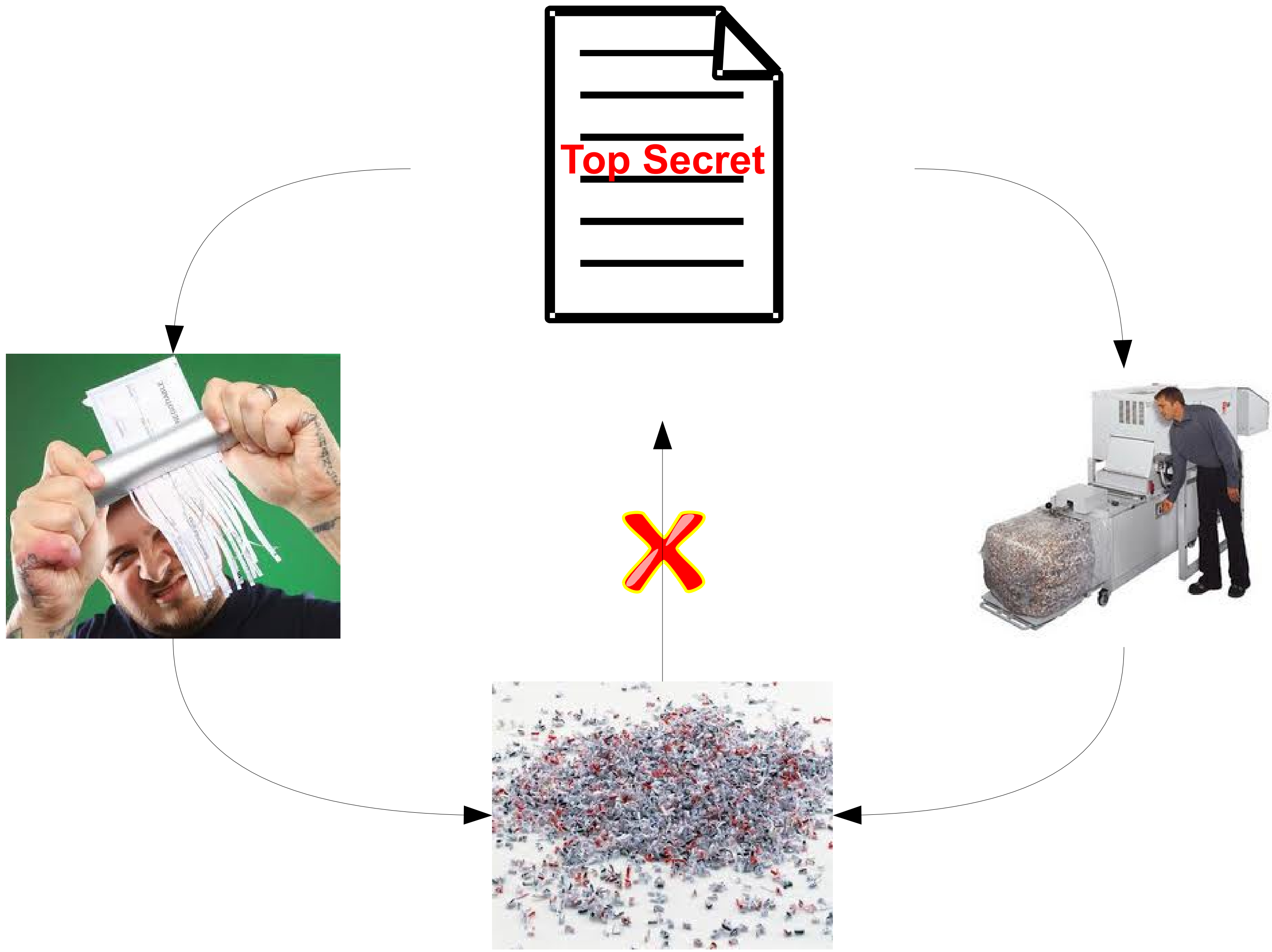
$$\Delta S = k\,\Delta H$$ (H = information entropy)

$$\Delta S_{\text{total}} \geq k \ln 2$$ (erasure of one bit)

$$\Delta S_{\text{total}} \geq k\,I$$ (erasure of mutual information I)

- Dissipation: Entropy production. $\Delta S_{total} = \Delta S_{sys} + \Delta S_{env}$
  Increase of total entropy in the universe.
- Information: ?
- Erasure: ?

What is information Erasure?

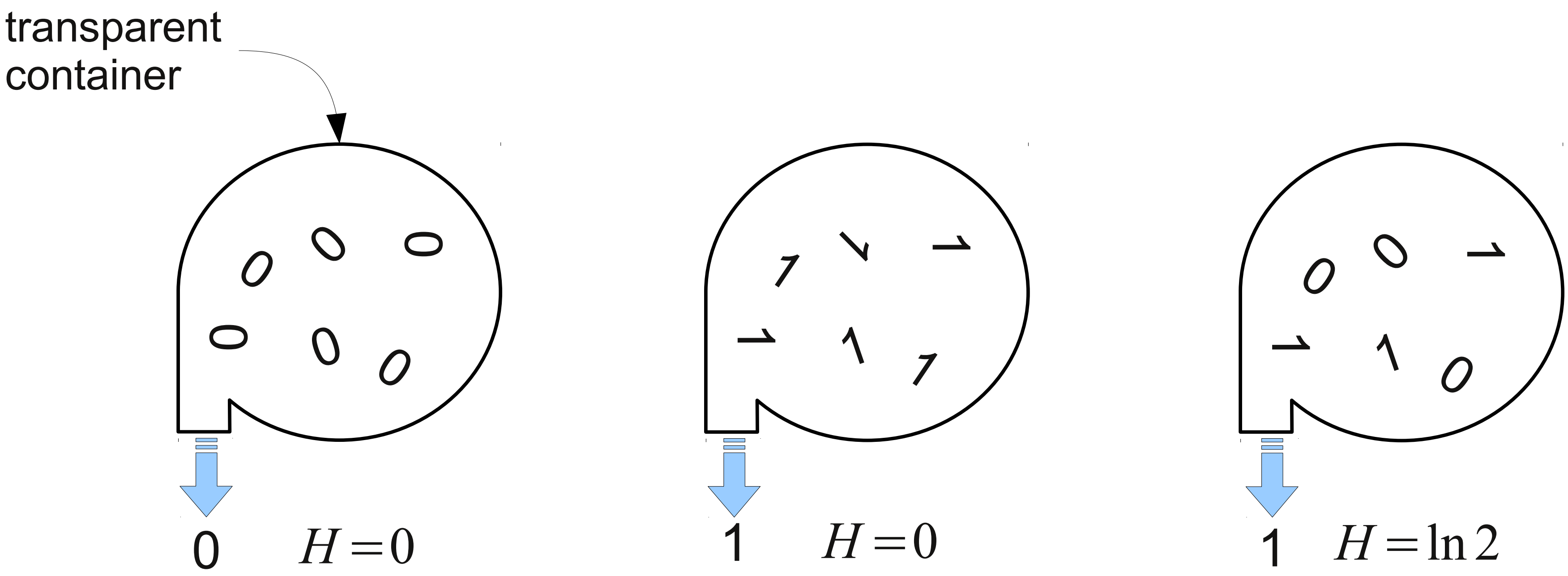


Bit Eraser

Which one erased the input information?

How much energy dissipates?

Information is probabilistic



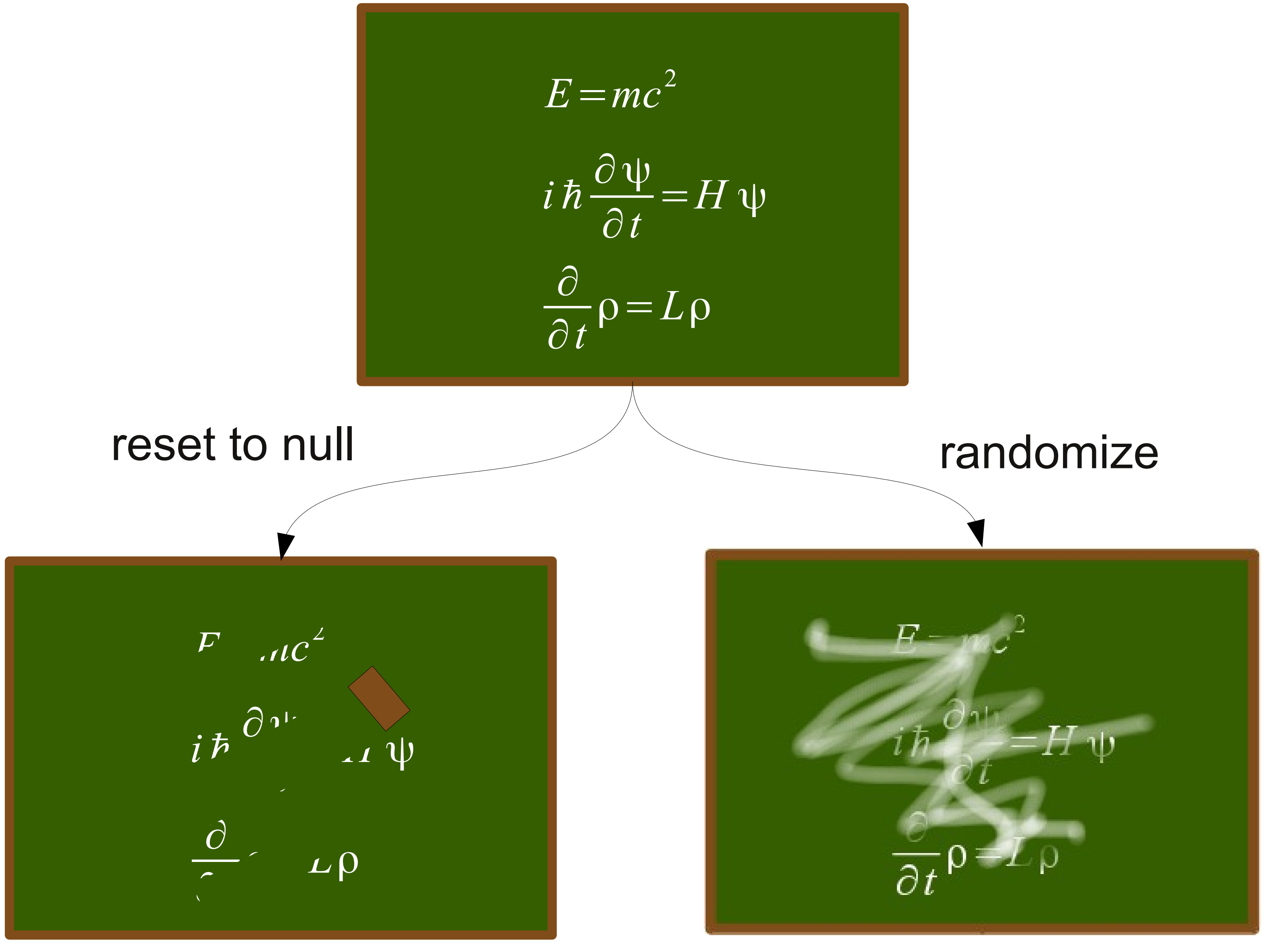
Information Entropy of one bit $H = -p_0 \ln p_0 - p_1 \ln p_1$

- Dissipation: Entropy production.
  Increase of total entropy in the universe.

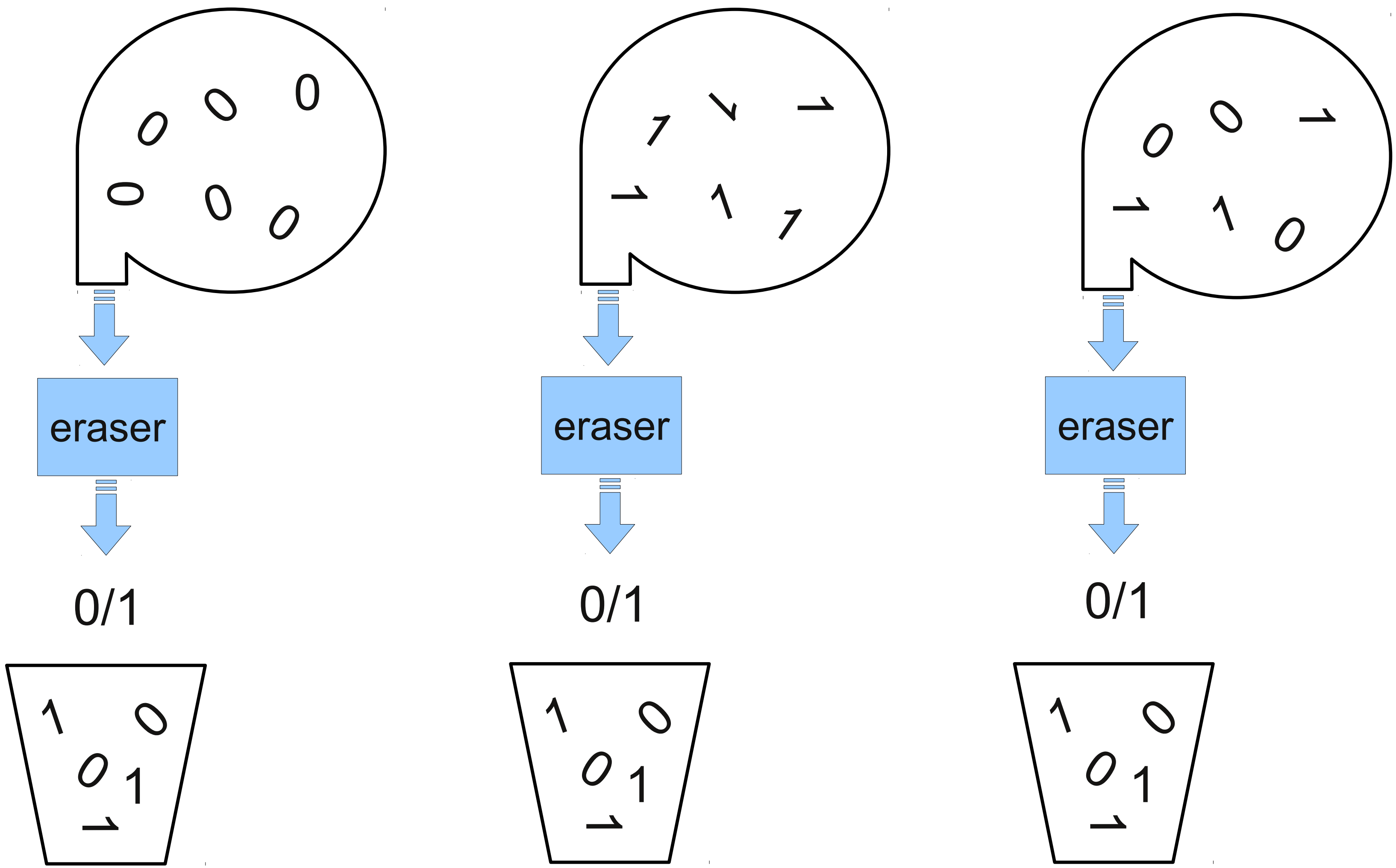
- Information: Uncertainty of bits
  Ensemble of bits

- Erasure: ?

Information Erasure

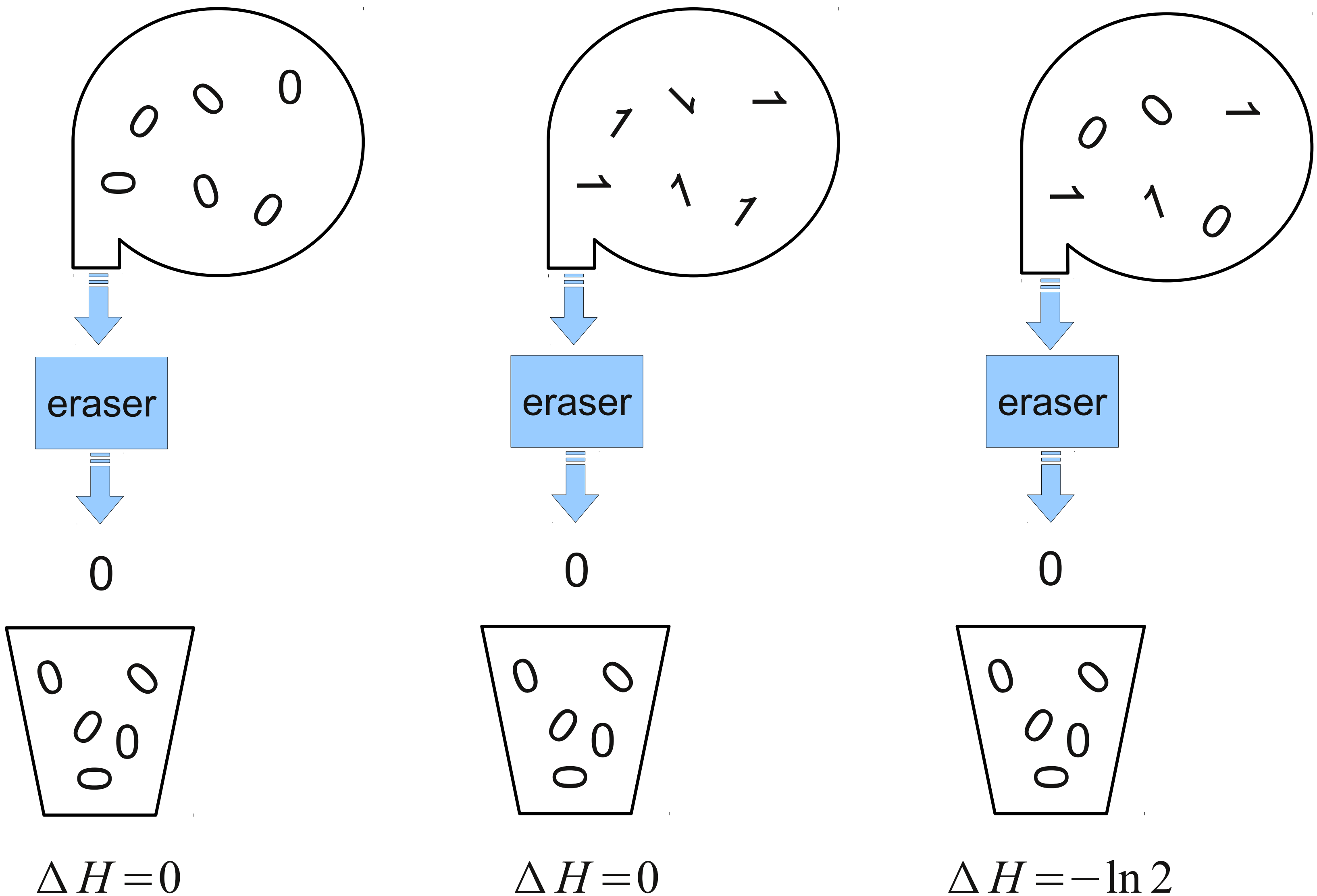


Bit Eraser (Randomize)



$\Delta H = \ln 2$ $\Delta H = \ln 2$ $\Delta H = 0$

Bit Eraser (Reset to Zero)



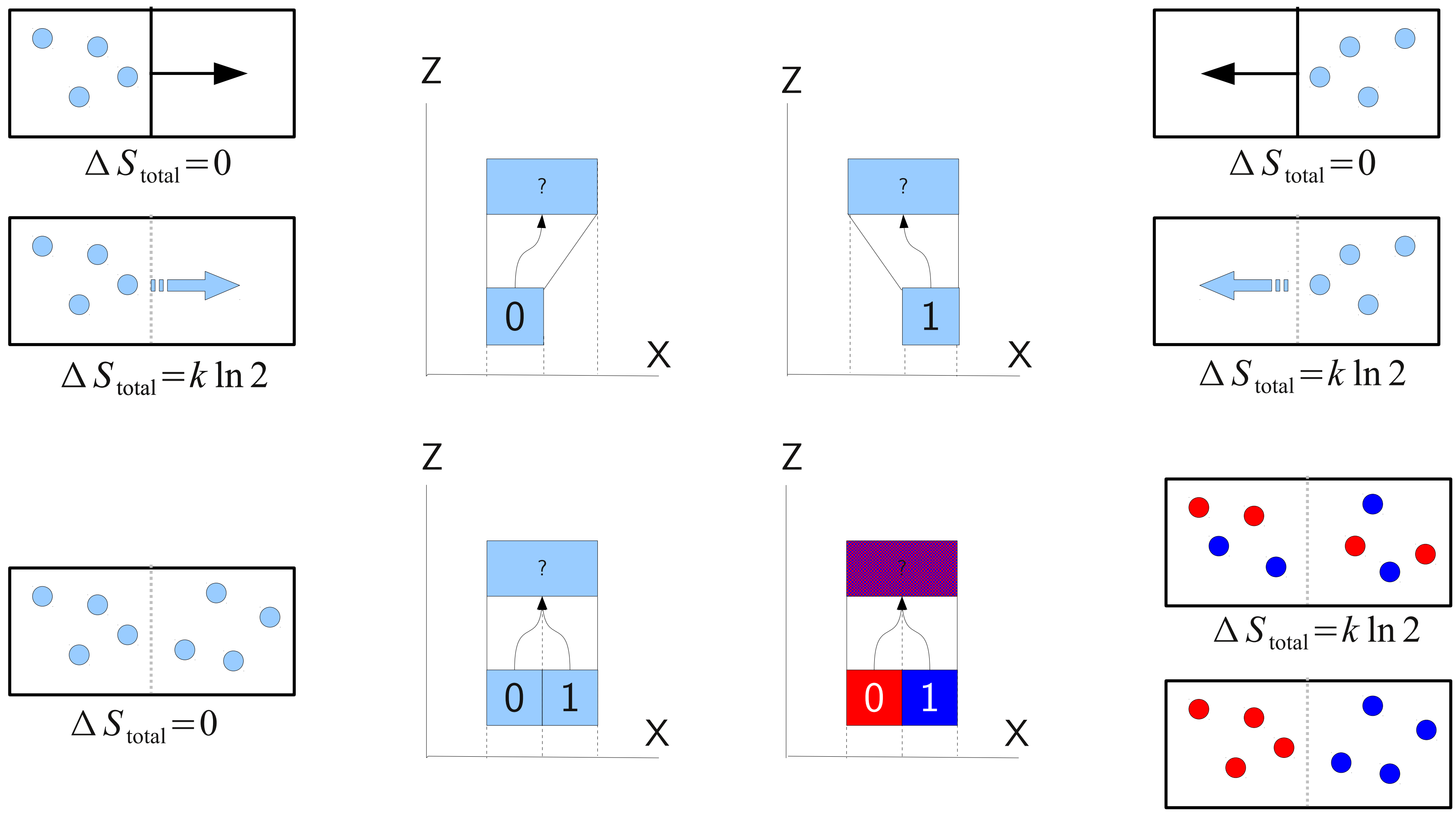
- Dissipation: Entropy production.
  Increase of total entropy in the universe.

- Information: Uncertainty of bits
  Ensemble of bits

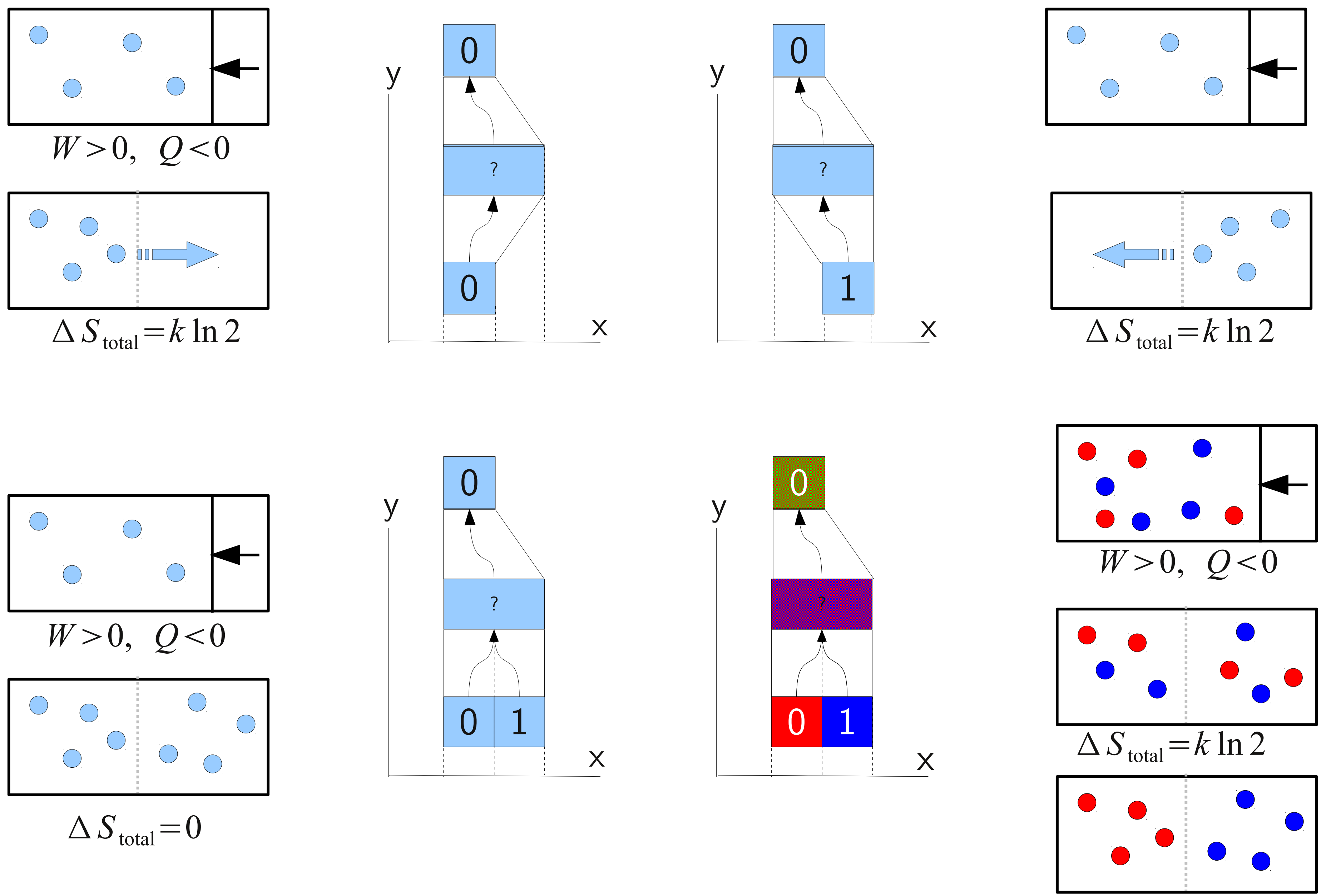
- Erasure: For all possible input information,
  the outcome is the same.
  Any ensemble of bits is mapped to
  a *standard* ensemble.
  (There is a constraint to the mapping procedure.)

Implementation of Information in Physics: A model

A Physical Model: Randomizing



A Physical Model: Resetting to Zero

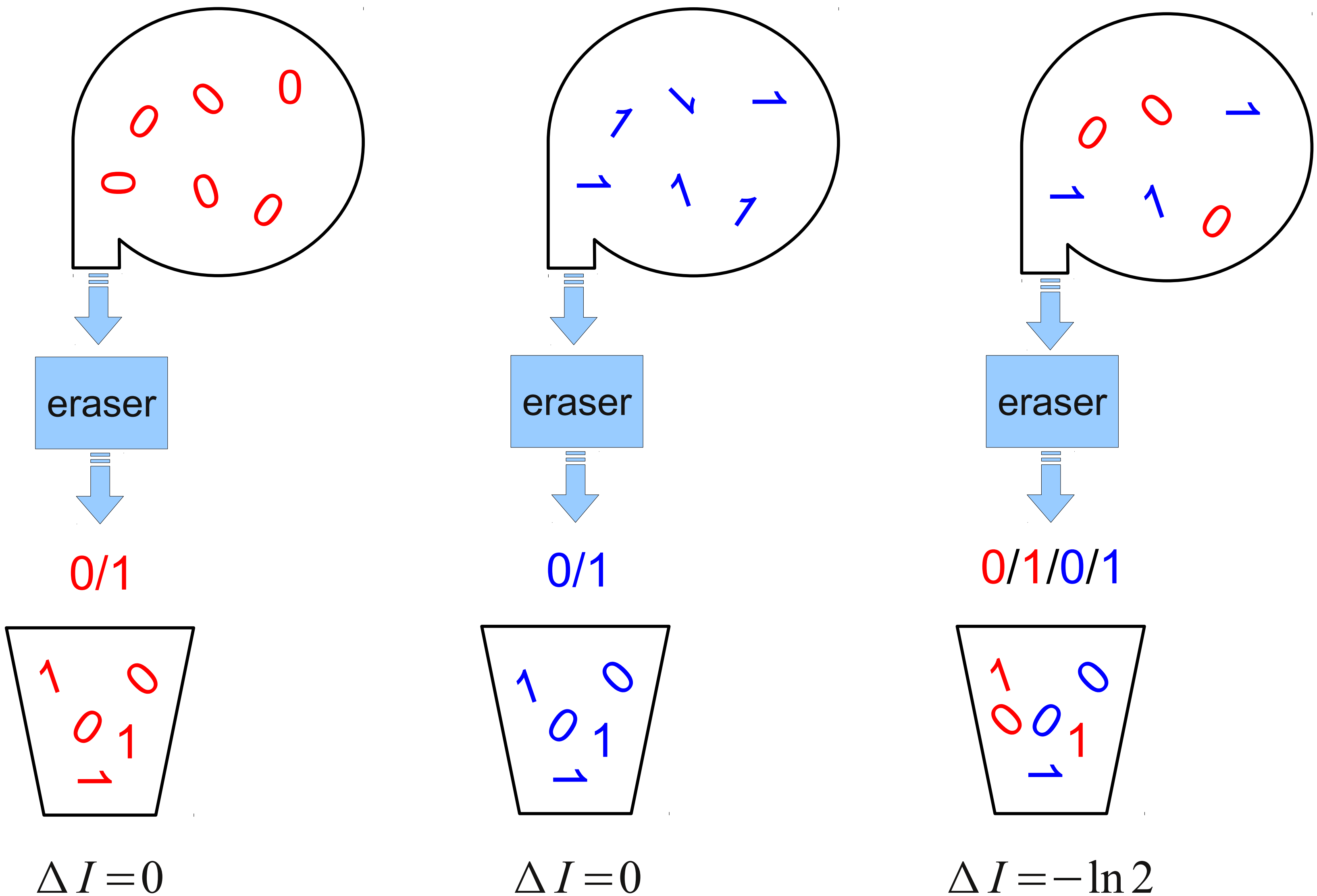


- Dissipation: Entropy production.
  Increase of total entropy in the universe.

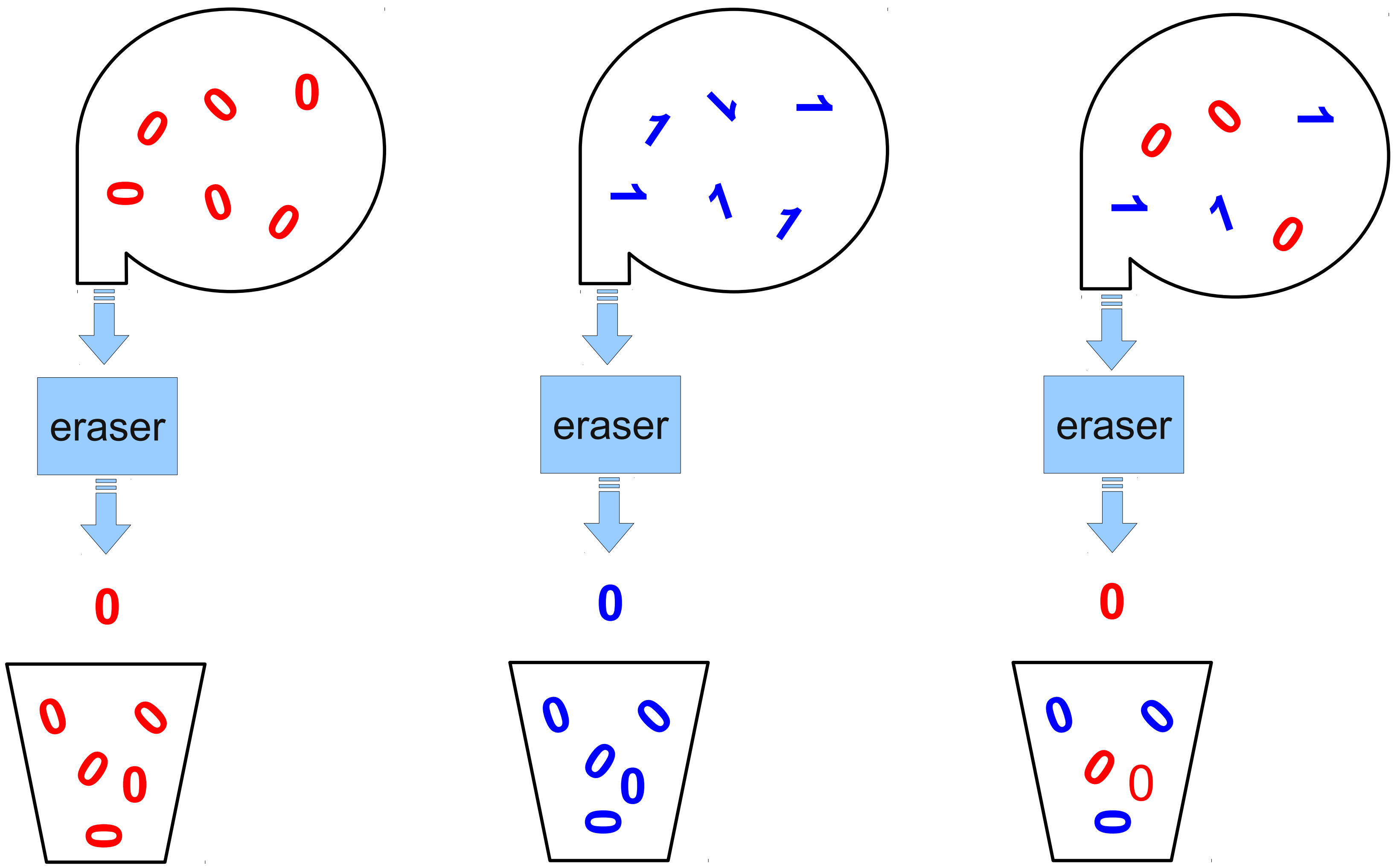
- Information: Uncertainty of bits
  Ensemble of bits

- Erasure: For all possible input information,
  the outcome is the same.
  Any ensemble of bits is mapped to
  a *standard* ensemble.
  The erasing procedure must be
  autonomous. (No feedback control.)

Bit Eraser (Randomize)



Bit Eraser (Reset to Zero)



$\Delta I = 0$ $\Delta I = 0$ $\Delta I = -\ln 2$

Erasure of Mutual Information

The bit eraser deleted the correlation between two bits of information.

Input: (0, R) and (1, B) $\quad p_{0R}=\frac{1}{2},\quad p_{1B}=\frac{1}{2},\quad p_{1R}=p_{0B}=0$

Output: (0,R), (0,B), (1,R), and (1,B) $\quad p_{0R}=p_{1R}=p_{0B}=p_{1B}=\frac{1}{4}$

Mutual Information between information X and Y:

$$I(X,Y)=H(X)+H(Y)-H(X,Y)=D(X\|Y)=\sum_{XY} p_{XY}\ln\frac{p_{XY}}{p_X p_Y}$$

where $p_X=\sum_Y p_{XY},\quad p_Y=\sum_X p_{XY}$

$$I_{\text{in}}=\ln 2,\quad I_{\text{out}}=0$$

When mutual information $I(X,Y)$ is erased,
$\Delta S_{\text{total}}\geq k\,I(X,Y)$ of entropy is generated.

More detailed analysis

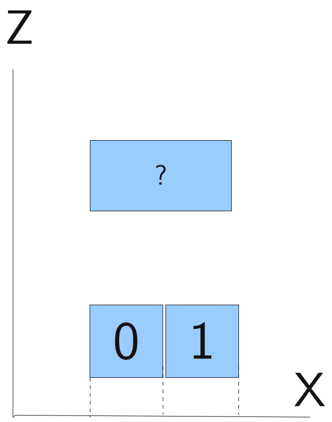
$$\Delta S_{\rm total} = k\, D(\rho_\Gamma(t) || \tilde{\rho}_\Gamma(t)) = k \int \rho_\Gamma(q,p,t) \ln \frac{\rho_\Gamma(q,p,t)}{\tilde{\rho}_\Gamma(q,-p,t)}$$

(Kawai et al., 2007)

Information bearing coordinates: X

Other coordinates: Z

$$\rho(X,t) = \int dZ\, \rho_\Gamma(q,p,t)$$



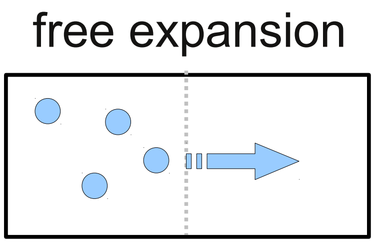
Relation with information

$$p_i = \int dX\, \rho(X,t)\, \chi_i(X), \qquad \chi_i(X) = \begin{cases} 1 & X \in R_i \\ 0 & \text{otherwise} \end{cases}$$

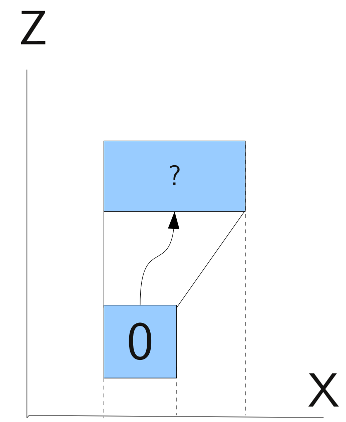
$$\Delta S_{\rm total} \geq D(\rho || \tilde{\rho}) \geq D(p || \tilde{p}), \quad D(p || \tilde{p}) = \sum_i p_i \ln \frac{p_i}{\tilde{p}_i}$$

Landauer principle via free expansion

Forward process: $\rho(X)=\begin{cases}\frac{1}{\Omega} & X\in R_0\\ 0 & X\in R_1\end{cases}$



Backward process: $\tilde{\rho}(X)=\begin{cases}\frac{1}{2}\Omega & X\in R_0\\ \frac{1}{2}\Omega & X\in R_1\end{cases}$

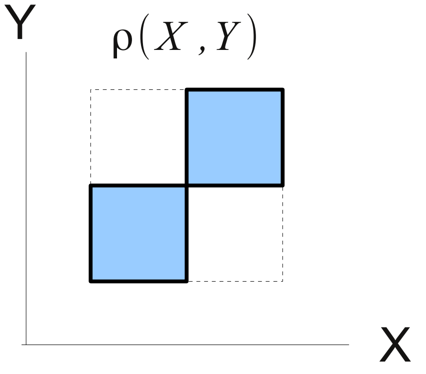


$$\Delta S_{\text{total}}\geq D(\rho||\tilde{\rho})=\ln 2$$

Thermodynamics textbook says: Entropy increases during free expansion because information is lost.

Landauer principle: When information is lost, entropy increases.

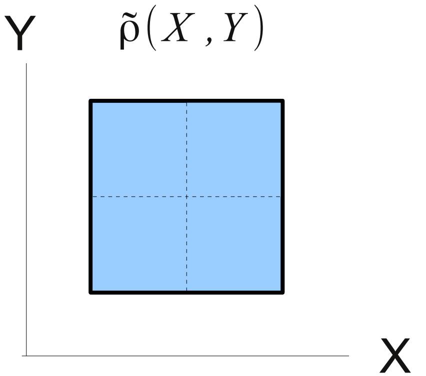
Proof of $\Delta S_i \geq I(X,Y)$



Hamiltonian: $H(X,Y) \to H_x(X) + H_y(Y)$

Forward: $\rho_{(X,Y)} \neq \rho_x(X)\rho_y(Y)$

$$\rho_x(X) = \int dY\, \rho_{\mathrm{eq}}(X,Y), \quad \rho_y(Y) = \int dX\, \rho_{\mathrm{eq}}(X,Y)$$



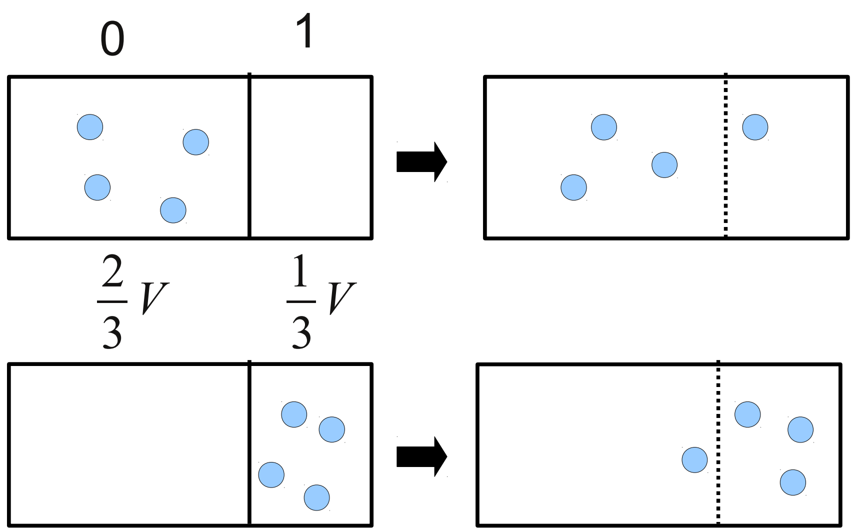
Backward: $\tilde{\rho}(X,Y) = \tilde{\rho}_x(X)\tilde{\rho}_y(Y)$

$$\Delta S_{\mathrm{total}} = D(\rho||\tilde{\rho}) = \int dX \int dY\, \rho(X,Y) \ln \frac{\rho(X,Y)}{\tilde{\rho}(X,Y)}$$

$$= \int dX \int dY\, \rho(X,Y) \ln \frac{\rho(X,Y)}{\rho_x(X)\rho_y(Y)} + \int dX \int dY\, \rho(X,Y) \ln \frac{\rho_x(X)\rho_y(Y)}{\tilde{\rho}_x(X)\tilde{\rho}_y(Y)}$$

$$= I(X,Y) + D(\rho_x||\tilde{\rho}_x) + D(\rho_y||\tilde{\rho}_y) \;\geq\; I(X,Y)$$

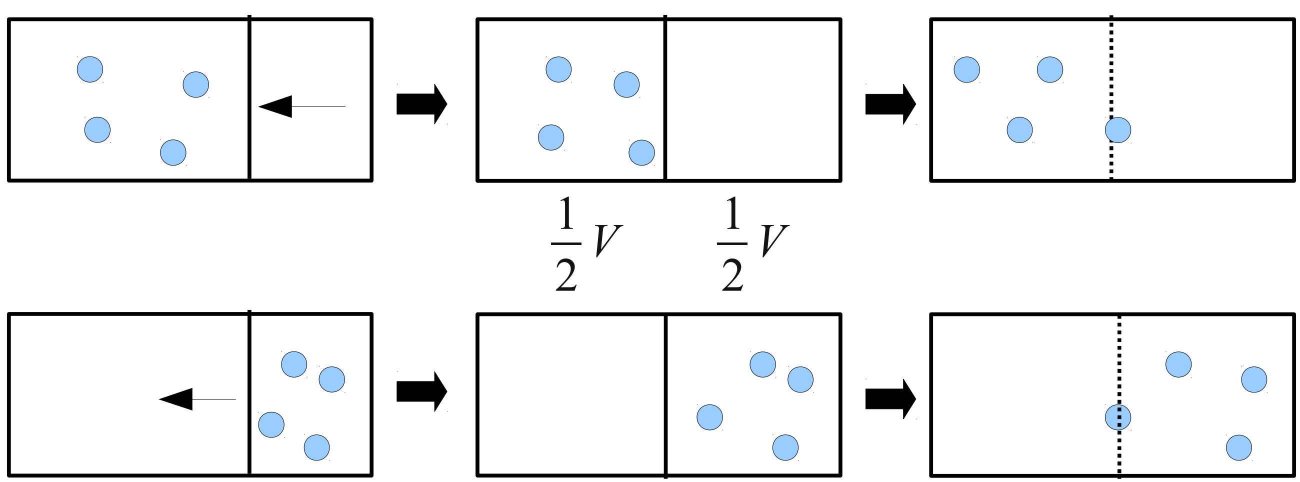
The issue of asymmetric domain size



$\Delta S_{\text{total}} = k \ln 3/2$

average: $\ln 3 - \frac{1}{2}\ln 2 > \ln 2$

$\Delta S_{\text{total}} = k \ln 3$



$W = kT \ln \frac{4}{3}, \quad \Delta S_{\text{total}} = k \ln 2$

$W = kT \ln \frac{2}{3}, \quad \Delta S_{\text{total}} = k \ln 2$

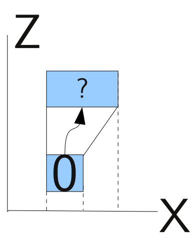
Sagawa and Ueda (2009)

About $\Delta S = k \Delta H$

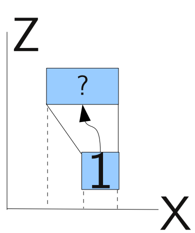
"When information entropy changes by ΔH , the system entropy changes by $\Delta S_{sys} = k\Delta H$. Since ΔS_{total} = ΔS_{sys} + $\Delta S_{env} > 0$, $\Delta S_{env} > -k\Delta H$."

This information entropy is different from the one used to measure the uncertainty in information.

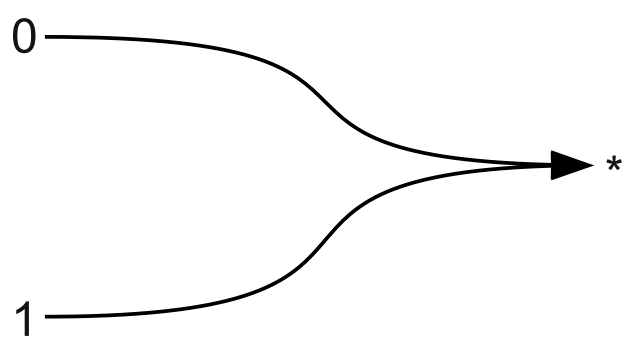
Information flow diagram

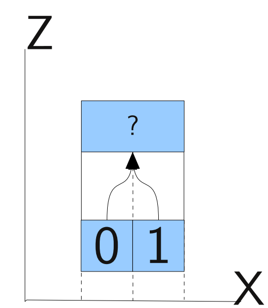


$$P_0 = \frac{1}{2}$$



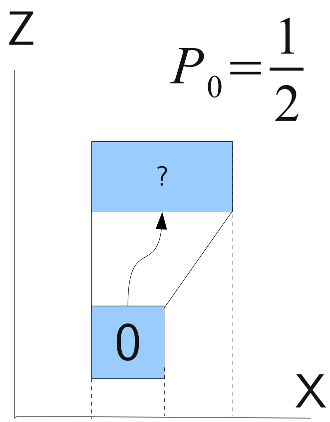
$$P_1 = \frac{1}{2}$$



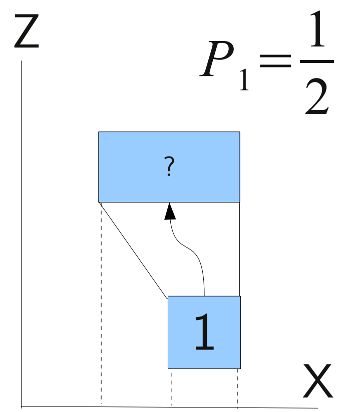


$$p_0 = \frac{1}{2} \quad p_1 = \frac{1}{2}$$

$$H = -P_0 \ln P_0 - P_1 \ln P_1$$



$P_0 = \frac{1}{2}$



$P_1 = \frac{1}{2}$

Input has two possible perfectly certain states. $H_{\text{in}} = \ln 2$

Output is unique. $H_{\text{out}} = 0$

$$\Delta H = -\ln 2 \quad \rightarrow \quad \Delta S = -k \ln 2$$

$$\Delta S_{\text{env}} \geq k \ln 2$$

Entropy is transferred from information world to physical world.

In this sense, the statement in Wikipedia is correct!

This was Landauer's original idea but no physical basis has been offered.

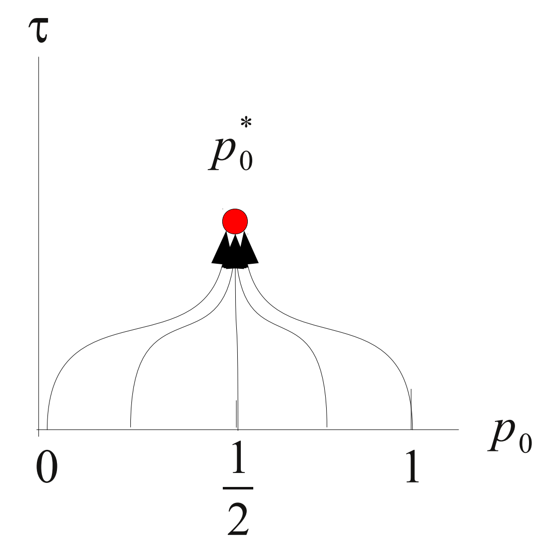
Summary on Information Erasure

- Information is a profile of probability distribution: $\{p_i\}$
- Relation between information and physics: $p_i = \int dX \, \rho(X) \chi_i(X)$
- Information erasure: $\forall \{p_i\} \to \{p_i^*\}$

  using an autonomous procedure (no feedback).

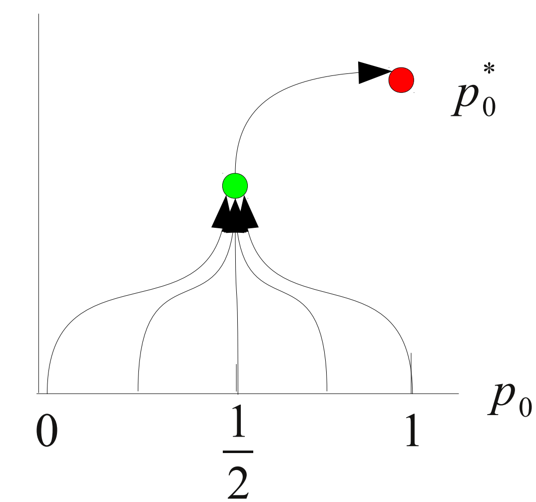
  (A single Hamiltonian for all initial conditions.)

  Remark: $\forall \rho(X, t_0) \to \rho^*(X, \tau)$ is not necessary



- Amount of required entropy production depends on the initial condition (input information).

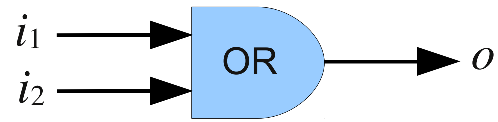
  The less uncertain is input, the larger is dissipation.
  If the input has no uncertainty, $\Delta S_{\text{total}} \geq k \ln 2$
  If the input is completely uncertain, $\Delta S_{\text{total}} \geq 0$



- Erasure of mutual information requires dissipation:

  $\Delta S_{\text{total}} \geq k I$

Simple Logical Gates

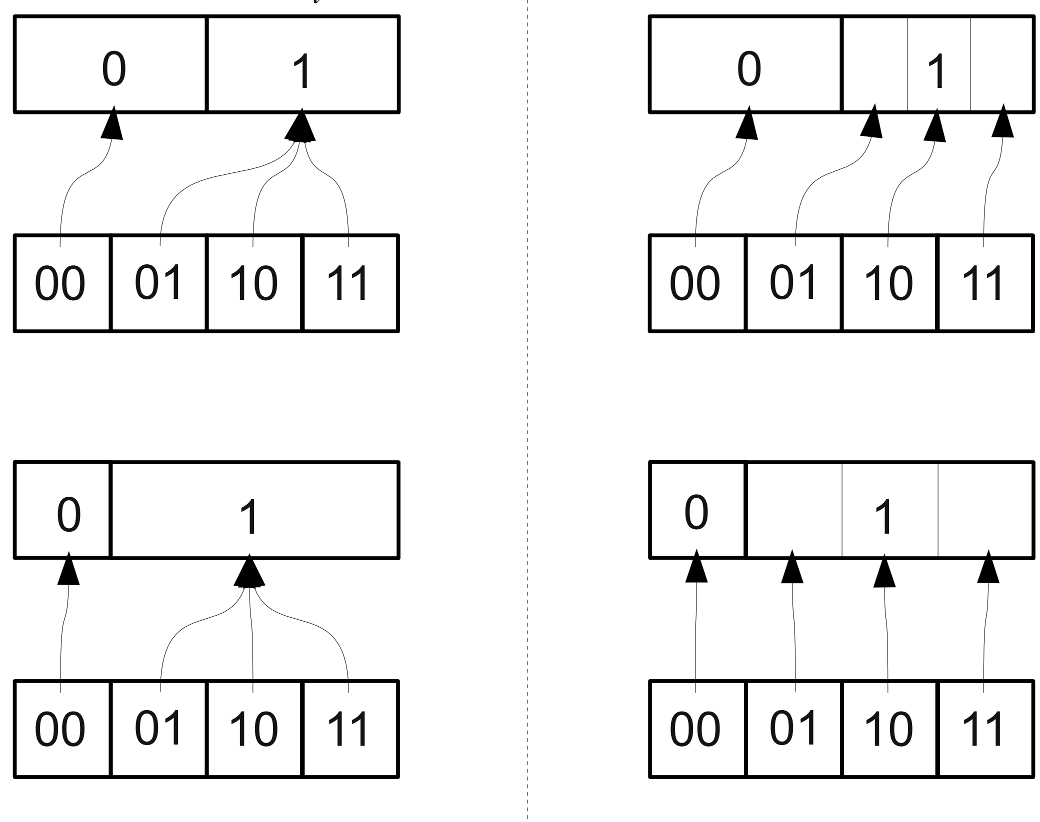


Input: Two bits ; Output: One bit $\rightarrow$ Erasure of one bit $\rightarrow$ $\Delta S_i \geq k \ln 2$

dissipative $\Delta S_i > ?$ | non-dissipative $\Delta S_i \geq 0$

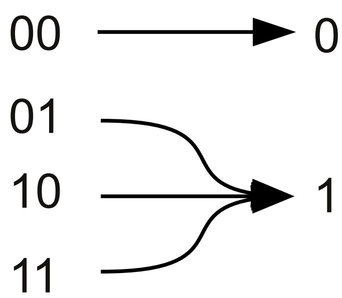
Need work

No work



Dissipation in an dissipative OR gate (information theory)

Common argument in information theory using information flow



$$P_{00}=P_{01}=P_{10}=P_{11}=\frac{1}{4} \quad \rightarrow \quad H_{\text{in}}=2\ln 2$$

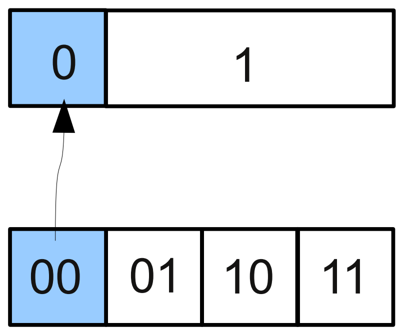
$$P_{0}=\frac{1}{4}, \quad P_{1}=\frac{3}{4} \quad \rightarrow \quad H_{\text{out}}=2\ln 2-\frac{3}{4}\ln 3$$

$$\Delta H=-\frac{3}{4}\ln 3$$

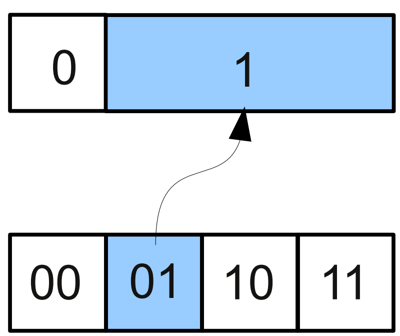
$$\Delta S_{\text{total}}=-k\,\Delta H=k\,\frac{3}{4}\ln 3$$

Dissipation in an dissipative OR gate (physics)

Assume that the input information has no uncertainty.



$$\Delta S_{\text{total}} \geq 0$$



$$\Delta S_{\text{total}} \geq k \ln 3$$

$$\overline{\Delta S_i} \geq \frac{1}{4} \times 0 + \frac{3}{4} k \ln 3 = k \frac{3}{4} \ln 3$$