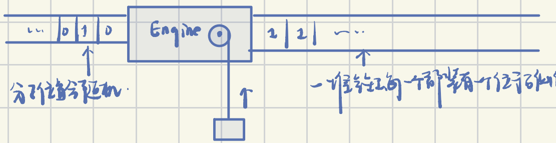
Day 5　Thermodynamics of Information

A Brief History

我们从另一个思想实验"Maxwell's demon"开始。假定将一个盒子分为A、B两部，由一个小妖精控制运动较快的分子进入B，而将较慢的分子送入A，从而B的温度升高，A反之。有一个类似的思想实验称为Szilard实验，它将一个只有1个分子的理想气体放在容器中，并把它放在了热库中。之后在容器中间插入一活塞，若分子在活塞左部，则将活塞向右移动直至抵达容器，反之亦然。此过程循环往复。在每一轮中，活塞("the Agent")将提取一份绝热膨胀功 $-W = k_BT\ln 2$。在这个实验中，活塞也可以被插入不同的位置，此时在一轮中可以提取的功为：

$-W = k_BT\cdot H(p)$，$H(p) = -\left(\frac{V_1}{V}\right)\ln\left(\frac{V_1}{V}\right) - \left(1-\frac{V_1}{V}\right)\ln\left(1-\frac{V_1}{V}\right)$，为粒子位置的香农熵。同理，你可以构想一个巨大的"信息引擎"：



以上看起来似乎违反了热二律，why? 因为在操作活塞的过程中，暂时地储存了关于分子位置的1bit信息，而循环中抹除这些信息需要做功。这就是所谓Landauer principle：当要逻辑不可逆地对信息进行操作时（例如对信息的擦除或computation path的合并），必然伴随系统或环境的熵增。更重要的是，从上"信息引擎"的例子可以大概知道，擦除一bit所需要的能量下界为 $\ln 2\cdot k_BT$。

Back to Noneq Free Energy

我们首先回顾一下Gibbs自由能的概念。在平衡态，我们有 $F=\langle\varepsilon\rangle^{eq}-TS$。这里S是对于一个系综直接定义的，而在随机热力学中，熵只对单条轨迹定义。

我们定义所谓非平衡自由能，$F^{neq}(p)=\langle\varepsilon\rangle-k_BT\,H(p)$。我们要说明在所有满足归一条件的分布中，$p=p^{eq}$ 时 F^{neq} 最小，证明如下：

$$
\begin{aligned}
\Delta F^{neq} &= F^{neq}(p)-F^{neq}(p^{eq}) \\
&= \sum_x\left[(p_x\varepsilon_x-p_x^{eq}\varepsilon_x)+k_BT\left(p_x\ln p_x-p_x^{eq}\ln p_x^{eq}\right)\right] \\
&= \left[\sum_x p_x(\varepsilon_x+k_BT\ln p_x)-\sum_x p_x^{eq}(\varepsilon_x-\ln p_x^{eq})\right] \\
&= \sum_x p_x(\varepsilon_x+k_BT\ln p_x)-F = \sum_x p_x(\varepsilon_x+k_BT\ln p_x-F) \\
&= \sum_x p_x\left(\varepsilon_x+k_BT\ln p_x-\varepsilon_x^{eq}-k_BT\ln p_x^{eq}\right) = k_BT\cdot D_{KL}(p\|p^{eq})\geq 0
\end{aligned}
$$

从而完成证明，并且我们发现两个分布间自由能之差为KL散度。

Information in Stochastic Thermodynamics

回顾一下Szilard Engine。我们将其视为要测量的物体(obj)和控制设备(dev)。在这个实验中，obj是个随机变量 $x\in\{L,R\}$，而dev也使用一个变量 $y\in\{L,R\}$ 来表示其测量结果。在测量完成后，dev的状态将与obj一致。在开始的情形下，假定系统完全随机，从而系统的熵为：$S^{sys}=k_B\ln 4=2\,k_B\ln 2$。

在完成测量后，系统的态只有(L,L)、(R,R)两种，从而系统熵减少了 $k_BT\ln 2$。这有两种情况：要么对系统做功 $-T\Delta S^{sys}=k_BT\ln 2$（假设系统已处于平衡态，则在最好的情形下，对系统做的功会变成热量），或者增加另一个系统的熵，例如一个memory。

那么来考虑一个更一般的情形。设obj可以处在 ε_x 上，以x标记其状态；而dev的状态以y表示，均满足 $\varepsilon_y=0$。且dev最初处在平衡态。

x,y之间的相关性可以由互信息表示：$I(obj;dev)=\sum_{x,y} p_{x,y}\ln\frac{p_{x,y}}{p_x p_y}$ 在测量开始时，假设完全处于平衡态，从而 $p_{x}^{eq}(t_0)=p_x^{eq}\cdot p_y^{0}$，在测量过程中 $p_x=\sum_y p_{x,y}$ 保持不变。

设系统在t_m时熵变为 $\Delta S^{sys}=S(t_m)-S(t_0)=-k_B\sum_{x,y}\left[p_{x,y}(t_m)\ln p_{x,y}(t_m)-p_{x,y}(t_0)\ln p_{x,y}(t_0)\right]$.

有个结果：$-k_B I(obj;dev)=-\sum_{x,y}p_{x,y}\ln\frac{p_{x,y}}{p_x p_y}$

$$\Delta S^{sys}+k_B I=-k_B\sum_{x,y}p_{x,y}(t_m)\ln p_{x,y}(t_m)+k_B\sum_{x,y}p_{x,y}(t_0)\ln p_{x,y}(t_0)+k_B\sum_{x,y}p_{x,y}(t_m)\ln p_{x,y}(t_m)-k_B\sum_{x,y}p_{x,y}(t_m)\ln p_x(t_m)p_y(t_m)$$

$$=k_B\sum_{x,y}\left(p_{x,y}(t_0)\ln p_{x,y}(t_0)-p_{x,y}(t_m)\ln p_x(t_m)p_y(t_m)\right)$$

$$=k_B\sum_{x,y}\left(p_{x,y}(t_0)\ln p_x(t_0)p_y(t_0)-p_{x,y}(t_m)\ln p_x(t_m)p_y(t_m)\right)$$

其中 $\sum_{x,y}p_{x,y}(t_0)\ln p_x(t_0)=\sum_x p_x(t_0)\ln p_x(t_0)$

$\sum_{x,y}p_{x,y}(t_m)\ln p_x(t_m)=\sum_x p_x(t_0)\ln p_x(t_0)$. 所以这两项抵消，看看剩下的项就有：$\Delta S^{sys}=-k_B I(obj;dev)-\Delta S^{dev}$

所以系统的熵变有两部分，一部分来自dev在测量前后的熵减，另一部分则在于由于引入的信息。通常在测量过程中有$\Delta S^{sys}<0$，那么至少有$Q=-T\Delta S^{sys}$的热被耗散，从而也需要做功。

有两个特殊情形：1). 测量并不改变dev的边缘分布，从而$\Delta S^{dev}=0$，此时，所有功用于增加dev和obj之间的互信息。

2). $\Delta S^{sys}=0$，或称测量是"绝热的"，$k_B I(obj;dev)=-\Delta S^{dev}$，从而信息的提升要靠"排放"到dev的熵中，从而提升dev的熵。

The Sagawa–Ueda Relation

我们讨论Feedback Control，也就是根据测量结果改变manipulated protocol，接着 $\lambda=\lambda(t,y)$. 对于一条随机轨道，我们可以定义一条轨道上t时刻的随机互信息。

$i_{x;y}=\ln\frac{p_{x,y}(t)}{p_x(t)\cdot p_y(t)}=\ln\frac{p_{y|x}(t)\cdot p_x(t)}{p_x(t)\cdot p_y(t)}$ 同理，一条路径的概率也可以被写成条件概率的形式：

$\downarrow I=\langle i\rangle$

$\mathcal{P}_{y|x}(\vec{x})=\underbrace{p_{y|x}(\vec{x})}_{\text{表示一条路径在}t_m\text{时刻测量结果的概率}}\,p_{\vec{x}}(\vec{x})$ 从而我们可以有Feedback Control的前提下的涨落关系。

$S^{tot}(\vec x)/k_B=\ln\left[\frac{p_{\vec x}(\vec x)}{\tilde p_{\vec x}(\vec x)}\right]$ ⇒ $-S^{tot}(\vec x)/k_B=\ln\left[\frac{\tilde p_{\vec x}(\vec x)}{p_{\vec x}(\vec x)}\right]$ ⇒ $-S^{tot}(\vec x)/k_B-i_{x;y}(t_m)=\ln\left[\frac{\tilde p_{\vec x}(\vec x)}{p_{\vec x}(\vec x)}\right]-\ln\frac{p_{y|x}(t_m)}{p_y(t_m)}$

⇒ $\exp\left(-S^{tot}(\vec x)/k_B-i_{x;y}(t_m)\right)=\frac{\tilde p_{\vec x}(\vec x)}{p_{\vec x}(\vec x)}\cdot\frac{p_y(t_m)}{p_{y|x}(t_m)}$ 于是有：

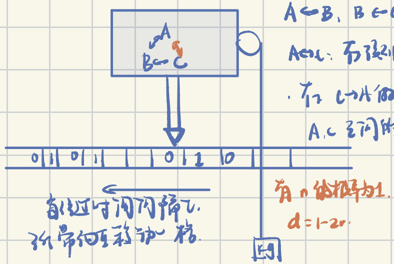
$\vec{x}$含y

$$\left\langle\exp\left(-S^{tot}(\vec x)/k_B-i_{x;y}(t_m)\right)\right\rangle=\int D\vec x\cdot dy\cdot p_{y|x}(t_m)\cdot p_{\vec x}(\vec x)\cdot\frac{\tilde p_{\vec x}(\vec x)}{p_{\vec x}(\vec x)}\cdot\frac{p_y(t_m)}{p_{y|x}(t_m)}=\int D\vec x\cdot dy\cdot\tilde p_{\vec x}(\vec x)\cdot p_y(t_m)=1.$$

令$I=\langle i\rangle$，从而上面的方程可以直接得出 $W-\langle\Delta F\rangle\geq -k_B T I$.

The Mandal–Jarzynski Machine.

M-J机器是一种可以被解析求解的将信息转化为自由能的机器，它有三个状态：A、B、C；机器与一条纸带交互，纸带有 1/0 两态，并能读、和写入。



A↔B，B↔C 有相同的速率。

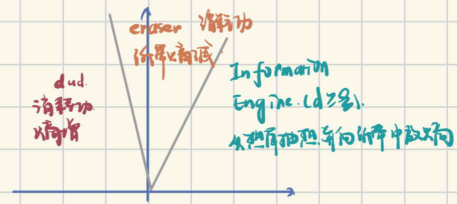
A↔C：若读到"1"时不能跳，同时纸带翻转。到0时。

在 C→A 的时候，需要将重物升 Δh，反之下降 Δh。

A、C 之间的跳转服从细致平衡。$k_{AC}/k_{CA} = \exp(-\frac{1}{k_BT} mg\Delta h)$.

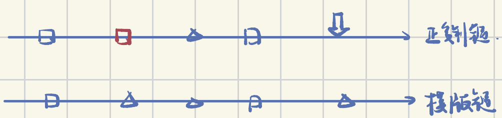
A、B、C 三个态被认为是简并的。重物被作为外部 Agent。换言之，A→C 的时候，Agent 对 machine 做功 $mg\Delta h$。

为方便起见，记 $k_{AC} = 1-\delta$，$k_{CA} = 1+\delta$，$\delta = \tanh\left(mg\Delta h/2k_BT\right)$。这东西有相图：



Copying Information.

在生命体内，通常存在对于信息高速且准确的复制行为，这通常包括 DNA 复制、转录。这一过程往往是可逆的。例如考虑 $(x, \ ) \to (x, y)$。$x' = x \text{ XOR } y$。所以这种可逆的信息复制可以是无耗散的，不需要 Landauer Bound 中的功的限制。具体来说，我们考虑如下模型：



在新链上的每一端，可以加入或移除新的单体（但也可能加入错误单体）。加入(+)／移除(-)，正确(r)／错误(w)单体速率为 $k^{(\pm)}_{r/w}$ 等的复制。从而，系统的 jump network 应该用一棵无限深的树建模。$0 \to r \to rr/rw$，$0 \to w \to wr/ww$，……

设在复制中有化学驱动 δ，从而速率设置为：$k_r^+ = \omega_r e^{\delta/k_BT}$，$k_r^- = \omega_r e^{\Delta\varepsilon_r/k_BT}$，$k_w^+ = \omega_w e^{\delta/k_BT}$，$k_w^- = \omega_w e^{\Delta\varepsilon_w/k_BT}$。从而我们有递归的主方程：

$$\begin{cases} \frac{d}{dt}p_{\ldots r} = k_r^+ p_{\ldots} + k_r^- p_{\ldots rr} + k_w^- p_{\ldots rw} - (k_r^+ + k_w^+ + k_r^-) p_{\ldots r} \\ \frac{d}{dt}p_{\ldots w} = k_w^+ p_{\ldots} + k_r^- p_{\ldots wr} + k_w^- p_{\ldots ww} - (k_r^+ + k_w^+ + k_w^-) p_{\ldots w} \end{cases}$$

由于这个方程是递归的，它的平稳分布并不好精确给出。由于我们有前面单体化的假设，所以我们可以猜测一条链出现的概率 $p_{\ldots} \propto (1-\eta)^{N_r} \eta^{N_w}$。

从而可以得到错误的比率：$\frac{\eta}{1-\eta} = \frac{k_w^+ - \eta k_w^-}{k_r^+ - (1-\eta) k_r^-}$。从而，复制链的延伸速度 $v = k_r^+ + k_w^+ - (1-\eta) k_r^- - \eta k_w^-$。

这里我们讨论两个极限：

1). Eq. limit：链不再增长的精确解给出。此时用前面方程解前面未知数，得到 $\begin{cases} \text{平衡分布}：\eta^{eq} = \dfrac{\exp(-\Delta\varepsilon_w/k_BT)}{\exp(-\Delta\varepsilon_r/k_BT) + \exp(-\Delta\varepsilon_w/k_BT)} \\ \text{"失速"能量 } \delta^{stall} = \varepsilon_r + k_BT \ln(1-\eta^{eq}) \end{cases}$。当 $\delta < \delta^{stall}$，系统将无法维持在平衡态上。

2). Fully Irreversible Limit：在非常大 driving 的情形下，$\eta = k_w^+/(k_r^+ + k_w^+) = \omega_w/(\omega_r + \omega_w)$。

在 δ 从 δ^{stall} 逐步升至 $+\infty$ 的过程中，系统将从一个极限移动到另一个。自然地，链复制的速度也应上升（注意：两个极限下的错误率无必然关系）。

另外可以导出，在非平衡态情形下的熵产生：(原来不仅有三项贡献，但根据我们的选择，不仅可以转换为熵产生：$(\cdots)$ $(\cdots\eta)$ $(\cdots w)$。从而熵产生速率为：

$$T\dot{S}^{tot} = k_BT\left[J^{+}\ln\left(\frac{k_+'}{k_-'}(1-\eta)\right) + J^{-}\ln\left(\frac{k_+''}{k_-''}\eta\right)\right] = v\left[\Delta W - \Delta F - k_BT\, D_{KL}(\eta\|\eta^{eq})\right] \geq 0.$$

其中 J_p 为从所有…到…的净速率之和。$\Delta W = \delta$ 为驱动。$\Delta F = k_BT\ln\left(e^{-\frac{\epsilon_1}{k_BT}} + e^{-\frac{\epsilon_2}{k_BT}}\right)$。

Information cost in sensing

在实验中，我们通常需要仪器测量快速变化的环境在一段时间内的平均状态，并保存在内存中。这样的系统被称为 sensory adaption System.

我们考虑简单的情形，信号 $e \in \{0,1\}$。系统的状态由两个 $\{0,1\}$ 变量描述：a，m。它们要跟踪 signal，但 m 的动力学慢于 a，所以 a 会先跟踪上，而 m 更慢。

能量用 $x = (a,m)$，$\epsilon = \epsilon_{a,m,e}$ 表示，它应有性质：· 在 $m=e$ 时，能量应当使系统倾向于 a 翻转朝向 e；· 在 $m \neq e$ 时，$a=e$ 的态比 $a \neq e$ 的态有更低的能量，所以可以这样选择：

$$\epsilon_{a,m,e} = |e-m|(\Delta_m + |a-e|\Delta_a)$$ 并记 $\Delta_{a/m} = -k_BT\ln\delta_{a/m}$。为简化起见，我们认为 a/m 的跳转不会同时发生，从而速率为

$$\begin{cases} k_{(1-a,m)\to(a,m)} = w_a\exp\left(\epsilon_{a,m,e}/k_BT\right) & (w_a \gg w_m) \\ k_{(a,1-m)\to(a,m)} = w_m\exp\left(\epsilon_{a,m,e}/k_BT\right) \end{cases}$$

$$P_{x|e} \propto \exp(-\epsilon_{a,m,e}/k_BT)$$

系统与环境的"相关性"仍应用互信息度量：$I(sys:env) = \sum_{x,e} p_{x,e}\ln\frac{p_{x,e}}{p_x p_e} = \sum_{x,e} p_{x|e}p_e\ln\frac{p_{x|e}}{p_x}$。利用互信息的性质可将其分解为：

$$I(sys:env) = \sum_{m,e} p_{m|e}p_e\ln\frac{p_{m|e}}{p_m} + \sum_{a,m,e} p_{a,m|e}p_e\ln\frac{p_{a|m,e}}{p_{a|m}} = I(mem:env) + I(act:env\,|\,mem).$$

考虑一个环境初态以等概率处于 $\{0,1\}$ 上，在 t 时刻以 $1/2$ 概率翻转。给定 $P_{x|e(t_0),e(t_f)}(t)$，我们定义两个量来分别代表对 t_f 时信息的获取和对 t_0 时信息的擦除。

$$\begin{cases} \Delta I^{meas}(t) = I(sys(t):env(t_f)) - I(sys(t_0):env(t_f)), \\ \Delta I^{eras}(t) = I(sys(t_0):env(t_0)) - I(sys(t):env(t_0)\,|\,env(t_f)), \end{cases}$$

其中

$$\begin{cases} I(sys(t):env(t_f)) = \sum_{x,e_i,e_f} p_{x|e_i,e_f}(t)\cdot p_{e_i,e_f}\ln\frac{p_{x|e_f}(t)\cdot p_{e_f}(t)}{p_x(t)\,p_{e_f}}, \\ I(sys(t):env(t_0)\,|\,env(t_f)) = \sum_{x,e_i,e_f} p_{x|e_i,e_f}(t)\, p_{e_i,e_f}\ln\frac{p_{x|e_i,e_f}(t)\,p_{e_i,e_f}}{p_{x|e_f}(t)\,p_{e_i,e_f}} = \sum_{x,e_i,e_f} p_{x|e_i,e_f}(t)\,p_{e_i,e_f}\ln\frac{p_{x|e_i,e_f}(t)}{p_{x|e_f}(t)}. \end{cases}$$

$p_{x|e_f}(t) = \sum_{e_i} p_{x|e_i,e_f}(t)\,p_{e_i}$，$p_x(t) = \sum_{e_i,e_f} p_{x|e_i,e_f}(t)\,p_{e_i}\,p_{e_f}$。

显然，系统初始的能量为0（此时 mem 和 env 对齐）。在系统翻转的过程中，外界将对系统做功，这完全是在 manipulation 的一刻做功。

所以做功总为 $W = \sum_{e_i,e_f}\langle\epsilon_{x,e_f}\rangle_{p^{eq}_{x|e_i}}\, p_{e_i,e_f} \geq 0$。而在后续弛豫至平衡的过程中，有一部分也要耗散为热。

考虑熵产生：$S^{tot}_{e_i,e_f}(t) = \Delta S^{sys}_{e_i,e_f}(t) + S^{res}_{e_i,e_f}(t)$（给定 e_i,e_f）。

$$\Delta S^{sys}_{e_i,e_f}(t) = k_B\left[H(p_{x|e_i,e_f}(t)) - H(p_{x|e_i,e_f}(t_0))\right]$$

$$S^{res}_{e_i,e_f}(t) = \frac{1}{T}Q^{res}_{e_i,e_f}(t) = \frac{1}{T}\left(\langle\epsilon_{x,e_f}\rangle_{p^{eq}_{x|e_i}} - \langle\epsilon_{x,e_f}\rangle_{p_{x|e_i,e_f}(t)}\right)$$

注意，由热二定律，$S(t)$ 是单增的。

熵产生也可被拆成与测量(measure)有关的部分和与擦除(erase)有关的部分。具体而言：

$$H(sys(t)\,|\,env(t_0),env(t_f)) = H(sys(t)\,|\,env(t_f)) - I(sys(t):env(t_0)\,|\,env(t_f))$$
$$= \underbrace{H(sys(t)) - I(sys(t):env(t_f))}_{measure} - \underbrace{I(sys(t):env(t_0)\,|\,env(t_f))}_{erase}.$$

计算可以得到：$d\Delta S^{meas}/dt = -k_B\sum_{e_f} p_{e_f}\frac{d}{dt}D_{KL}(p_{x|e_f}(t)\,\|\,p^{eq}_{x|e_f}) \geq 0$。其物理意义为[illegible] $\Delta S^{eras}\uparrow$

推论 $\lim_{t\to t_c} Q = W - \langle\Delta\varepsilon\rangle \geq 0$. $\Delta F = \langle\Delta\varepsilon\rangle - k_B T\Delta H$，从而有，

$W - \langle\Delta\varepsilon\rangle \geq 0$，$\langle\Delta\varepsilon\rangle = \Delta F + k_BT\Delta H$ ⇒ $W - \Delta F - k_BT\Delta H \geq 0$ ⇒ $W - \Delta F \geq k_BT\Delta H$. 从而有：

$$\begin{cases} W - \langle\Delta F\rangle \geq k_BT\left[I(sys(t_m);env(t_f)) - I(sys(0);env(t_f))\right] = k_BT\,\Delta I^{meas} \\ T\,\Delta S^{eras} \geq k_BT\left[I\left(sys(0);env(\tau)\,\middle|\,env(t_f)\right) - I\left(sys(t_m);env(t_m)\,\middle|\,env(t_f)\right)\right] = k_BT\,\Delta I^{eras} \geq 0 \end{cases}$$

依据上面的表示，我们可以得到解释：对系统做测量需要做功，在调节signal的时候相当于调节了系统各个能级的能量，从而对系统做了功；而擦除原有的信息则会有熵的产生。

对于很多真实的系统，尤其是生物体内的系统，它们通常位于不满足精细平衡条件的非平衡稳态。对于这样的系统，以上不等式也是成立的，因为可以将 ΔS 分解为 ΔS^{na}，ΔS^{a}，再分别证明。

Information Reservoirs.

前面，我们已经讨论了两种不同的情形，分别为Szilard思想实验和M-D机器。现在，我们引入一个统一框架，来说明测量设备(dev)和纸带都可以被视作某种信息库。

简单起见，我们考虑一个二能级系统，$\varepsilon_{down} = 0$，$\varepsilon_{up} = \varepsilon > 0$，它处在温度为T的热库中，从而它到有：$p_d^{eq} = \frac{1}{1+\exp(-\varepsilon/k_BT)}$ $p_u^{eq} = \frac{\exp(-\varepsilon/k_BT)}{1+\exp(-\varepsilon/k_BT)}$，对应跃迁 $k^\uparrow/k^\downarrow = \exp(-\varepsilon/k_BT)$.

现在，除了热库之外，还让系统与一个信息库接触，系统与信息库交互的速率为 γ。若系统位于u态并发生交互，则系统会跳至d态并对外做功 ε，反之亦然。

具体来说，"翻转"的过程是从信息库中抽一个态，并且将系统的态变为抽取的这个态。所以假设信息库中u态含量为r，d态为1-r，从而由信息库引发的状态跳跃速率：$\tilde{k}^\uparrow = \gamma r$ $\tilde{k}^\downarrow = \gamma(1-r)$.

与信息库接触后，系统的主方程和平衡态均发生变化，有：（此时系统在外部/信息库的驱动下处于非平衡稳态）.

$$\begin{cases} dp_u/dt = (k^\uparrow + \tilde{k}^\uparrow)\,p_d - (k^\downarrow + \tilde{k}^\downarrow)\,p_u \\ dp_d/dt = (k^\downarrow + \tilde{k}^\downarrow)\,p_u - (k^\uparrow + \tilde{k}^\uparrow)\,p_d \end{cases} \Rightarrow \begin{cases} p_u^{st} = \frac{k^\uparrow + \tilde{k}^\uparrow}{k+\gamma} \\ p_d^{st} = \frac{k^\downarrow + \tilde{k}^\downarrow}{k+\gamma} \end{cases}$$

其中 $k = k^\uparrow + k^\downarrow$，$\gamma = \tilde{k}^\uparrow + \tilde{k}^\downarrow$.

下面考虑如何使用这样一套建模方法的问题.

> Measurement of Feedback 考虑一个最简单的反馈控制系统，系统的状态被以 γ 的速率测量，$y \in \{0,1\}$，1对应u，0对应d，r为测量错误率。换言之，$p_{1|u} = p_{0|d} = 1-r$；$p_{0|u} = p_{1|d} = r$.

这时候我们就会改变jump rate还原至该模型。dev测出1的概率(也即在一次交互后，系统立刻处于u态的概率)为 $p_1^{dev} = (1-r)\,p_u^{st} + r\,p_d^{st} = p_u^{st} + r - 2r\,p_u^{st}$

而单位时间内外界对系统做功：$\dot W = \varepsilon\gamma\left[r\,p_d^{st} - (1-r)\,p_u^{st}\right] = -\gamma\varepsilon\,(p_u^{st} - r)$。当系统处于非平衡稳态时，$\dot W$ 作为单位时间内系统向热库放热 $\dot Q_{res}$，或 $\dot W = T\dot S_{res}$

对于处于非平衡稳态的系统，其dev和sys的互信息 $I = H(dev) - H(dev|sys) = H(dev) - H(r)$。对于任一等温过程有 $\langle\Delta F\rangle = 0$，从而有 $W + k_BTI \geq 0$。对于任意无限小的过程都成立，从而 $\dot W + k_BT\dot I = 0$.

> TAPE. 在M-D机器中，信息库由纸带充当，以 γ 的速率纸带被移动。若纸带上的1被读取则系统以u表示。若系统在某一时刻通过与热库交互转换了状态，则纸带上的bit也被反转。从而这时候，系统的动力学和前面的模型也完全相同。*在这样系统下进入纸带的/离开的bit为1的概率分别为 r 和 p_u^{st}。从而，纸带上的熵产生率 $\gamma k_B\left[H(p^{st}) - H(r)\right]$.

$$T\dot S^{tot}_{TAPE} = \dot W + \gamma k_BT\left[H(p^{st}) - H(r)\right] \geq 0, \quad = \gamma k_B T\left[D_{KL}(r\|p^{eq}) - D_{KL}(p^{st}\|p^{eq})\right] \geq 0.$$

从这里可以看出若 $H(p^{st}) - H(r) < 0$，此时系统像eraser一样 ⇒ $\dot W > 0$.

> Generalized Detailed Balance. 另外一种方式，我们可以使用前文结果计算熵产生率.

$$\dot S^{tot}_{GDB} = k_B\left(k^\downarrow p_u^{st} - k^\uparrow p_d^{st}\right)\ln\frac{k^\downarrow}{k^\uparrow} + k_B\left(\tilde{k}^\downarrow p_u^{st} - \tilde{k}^\uparrow p_d^{st}\right)\ln\frac{\tilde{k}^\downarrow}{\tilde{k}^\uparrow}$$

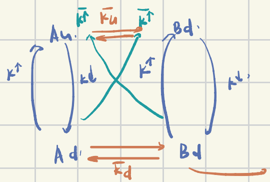
在这个不同的情形下，我们得到了三个不同的熵产生，这是因为MF和TAPE中我们其实忽略了一部分熵产生。e.g. 在MF中，我们忽略了dev获得和擦除信息的熵，在TAPE中，我们忽略了产生纸带初态的熵。

在TAPE模型中，不妨考虑 $\dot{W}<0$ 的情形，此时有 $|\dot{W}| \geq T k_B T [H(p^{st}) - H(\cdot)]$ 这是将"燃尽了"的纸带恢复回初始状态的最小功。

→ Fluctuation Relations with Information reservoirs.

前面有一些不等式，这些不等式可以从涨落关系推出。之前，我们知道有些交互并不会导致跳跃。$\bar{k}_d = \sigma(1-\eta)$，$\bar{k}_u = \gamma\sigma$。我们现在考虑

我们考虑系统的两个copy，A、B。每次与信息库交互时，我们就从一个copy转至另一个。从而系统有四个态。$x \in \{d, u\}$，$\alpha \in \{A, B\}$。



从而在这种情况下，系统的稳态 $p^{st}_{x,\alpha} = \frac{1}{2} p^{st}_x$，$\forall x, \alpha$。

现在我们计算一下涨落关系。在之前情形下，我们有

为什么要有两套系统？因为有些系统是不可观测的，与信息库的交互也有不可观测信息库的熵，例如改写纸带上的内容。

$$\frac{P[\vec{x}|x_0(\omega)]}{P[\tilde{\vec{x}}|x_0(\omega)]} = \prod_{i=1}^{n} \frac{k_{x_i, x_{i-1}}}{k_{x_{i-1}, x_i}}$$

现在，我们认为系统与信息库的交互也会导致信息库的熵变，对信息库熵变的描述（或对信息库的不同描述），将导致这一部分的熵有着不同结果，设在路径反向后，$k^{\uparrow}$ 与保持不变。

但 $\bar{k}^{\uparrow}$、$\bar{k}_u$、k_u、k_d 要发生变化。且在这里的模型中，我们设置，$\hat{k}^{\uparrow} + \hat{k}_d = \hat{k}^{\downarrow} + \hat{k}_u = \gamma$。从而由正反向轨迹的概率之比，我们可以求解熵变。

$$\frac{P_{\vec{x}}(k, \bar{k}, k_{u/d})}{P_{\tilde{\vec{x}}}(k, \hat{k}, k_{u/d})} = \frac{P_{x_0}(t=0)}{P_{x_n}(t_f)} \exp\left(\frac{1}{k_B} S^{neq}[\vec{x}]\right) = \exp\left(\frac{1}{k_B} S^{total}[\vec{x}]\right)$$

$$S^{tot} = \underbrace{\frac{1}{2} k_B \sum_{x \neq x'} J_{xx'}(\vec{x}) \ln \frac{k_{xx'}}{k_{x'x}}}_{\text{热库的熵变（与之前一样）}} + \underbrace{k_B \sum_{x, x'} R_{xx'}(\vec{x}) \ln \frac{\bar{k}_{xx'}}{\hat{k}_{x'x}}}_{\text{信息库的熵变（依 }\hat{k}_{x'x}\text{ 的不同选法代表了对信息库的不同处理）}}$$

$J_{xx'}$ 只统计在同一套系统内部的跳跃，因为这些跳跃是通过与热库的交互进行的。而 $R_{xx'}$ 统计系统的两个copy之间的跳跃，因为这些跳跃是通过与信息库的交互进行的。

$$J_{xx'}(\vec{x}) = \sum_{l=0}^{n} \sum_{\alpha} \left(\delta_{x\alpha, x_{l+1}\alpha_{l+1}} \delta^{K}_{x'\alpha, x_l\alpha_l} - \delta^{K}_{x'\alpha, x_{l+1}\alpha_{l+1}} \delta^{K}_{x\alpha, x_l\alpha_l} \right)$$

$$R_{xx'} = \sum_{l=0}^{n} \left(\delta^{K}_{xA, x_{l+1}\alpha_{l+1}} \delta^{K}_{x'B, x_l\alpha_l} + \delta^{K}_{xB, x_{l+1}\alpha_{l+1}} \delta^{K}_{x'A, x_l\alpha_l} \right)$$

不存在系数 $\frac{1}{2}$，因为正/反向的跳跃只会被统计一次，而并非像 $J_{xx'}$ 中被统计二次，所以这里没有系数 $\frac{1}{2}$。

我们还要解释为何这样写出来的东西确实是正反路径的概率之比。众所周知，一条路径的概率由跳跃概率和在各个态上保持概率连乘得到。由于 $\bar{k}_{\uparrow} + k_d = \hat{k}_{\uparrow} + \hat{k}_d = \bar{k}^{\downarrow} + k_u = \hat{k}^{\downarrow} + \hat{k}_u = \gamma$，所以，在正/反向路径上，从某个态离开的速率是一样的，这样我们仍只需考虑跳跃概率。

$$\frac{P_{\vec{x}}(k, \bar{k}, k_{u/d})}{P_{\tilde{\vec{x}}}(k, \hat{k}, \hat{k}_{u/d})} = \frac{P(\quad)}{P(\quad)} (在同一套系统间跳跃) \cdot \frac{P(\ )}{P(\ )} (在两个copy间跳跃)$$

$= \prod_{l=0}^{N\,(总的步数)} \frac{k_{x_{l+1},x_l}}{k_{x_l,x_{l+1}}}$ (若 x_l, x_{l+1} 同属一个 copy)，$\frac{\bar{k}_{x_{l+1}x_l}}{\hat{k}_{x_l,x_{l+1}}}$ (若 x_l, x_{l+1} 不属同 copy)。

所以很容易看出上面的形式确实给出了一条路径的概率比，从而我们有涨落关系：$\langle \exp(-S^{tot}/k_B) \rangle = 1$，从而 $\langle S^{tot} \rangle \ge 0$。

特别地，对于以上描述的系统，在与热库、信息库交换的过程中，不断有熵产生。⇒

$$\dot{S}^{tot} = \lim_{T\to\infty} \frac{\langle S_{res} \rangle}{T} = \frac{1}{2} k_B \sum_{x\neq x'} J_{xx'} \ln \frac{k_{xx'}}{k_{x'x}} + k_B \sum_{xx'} R_{xx'} \ln \frac{\bar{k}_{xx'}}{\hat{k}_{x'x}} \ge 0$$

其中 $J_{xx'}$、$R_{xx'}$ 都是稳态时的概率流。

$$\begin{cases} J_{xx'} = k_{xx'} p^{st}_{x'} - k_{x'x} p^{st}_{x} \\ R_{xx'} = \bar{k}_{xx'} p^{st}_{x'} \end{cases}$$

(注意这里 $R_{xx'}$ 只统计了单向的流量)。

下面我们说明，对于 $\hat{k}$ 的不同选取对应于对信息库的不同解读，由此对应了不同的熵产率。

> 对于 Measurement and Feedback，选取 $\hat{k}_{ud} = \hat{k}_{du} = \gamma p_1^{dev}$，$\hat{k}_{uu} = \hat{k}_{dd} = \gamma(1-p_1^{dev})$。从而立刻得到其熵产率。

$$\dot{S}^{tot}_{MF} = \frac{1}{2} k_B \sum_{x\neq x'} J_{xx'} \ln \frac{k_{xx'}}{k_{x'x}} + \gamma k_B [H(dev) - H(Y)] \ge 0$$

> 对于 TAPE，取 $\hat{k}_{uu} = \hat{k}_{ud} = \gamma p^{st}_u$；$\hat{k}_{du} = \hat{k}_{dd} = \gamma p^{st}_d$，从而得到其熵产率。$\dot{S}^{tot}_{TAPE} = \frac{1}{2} k_B \sum_{x\neq x'} J_{xx'} \ln \frac{k_{xx'}}{k_{x'x}} + k_B \gamma [H(p^{st}) - H(Y)] \ge 0$

这就是一个例子：由于不能同时看到信息库的状态，而且不能可以翻转纸带上的状态，所以可以视作每一次和信息库交互都在改变信息库的熵。

> 对于 Generalized Detailed Balance，实际上就是正常的取法，$\hat{k}_{xx'} = k_{xx'}$，从而导出

$$\dot{S}^{tot}_{GDB} = \frac{1}{2} k_B \sum_{x\neq x'} J_{xx'} \ln \frac{k_{xx'}}{k_{x'x}} + k_B J' \ln \frac{H}{p} \ge 0, \quad J' = \bar{k}\, p^{st}_u - \bar{k}\, p^{st}_d$$

三种定义下的熵产率有关系：

$$\begin{cases} \dot{S}^{tot}_{GDB} - \dot{S}^{tot}_{TAPE} = k_B \gamma D_{KL}(p^{st} \| Y) \ge 0 \\ \dot{S}^{tot}_{MF} - \dot{S}^{tot}_{TAPE} = k_B \gamma [H(dev) - H(p^{st})] \ge 0 \end{cases}$$

在信息库上的影响状态等价的部分状态时，以上三种核也是可以使用的。