随机热力学（书版）笔记重制-Day5：信息的热力学

#StochasticThermaldynamics

有一个有趣的思想实验：一个容器中只有 1 个气体分子，热浴在温度为 T 的恒温热库中。在容器中点插入厚度可忽略的活塞，则若分子位于活塞左侧，活塞将向右移动；反之亦然。在每一轮中，活塞都从绝热膨胀中提取了功 $-w = k_B T \ln 2$（注意这里使用 w 代表外界对系统做功）。这个做法看似是违反热力学第二定律的 （热二定律的一个等价表述：从单一热源吸热并完全转化为功），但是，要进行这个操纵，热机必须有一个存储器存储关于分子位置的 1 Bit 信息，并在每轮运行后擦除。通过这个思想实验，兰道尔原理指出：擦除一个比特的信息 （或者可以说得更广泛一些：对 1 Bit 信息进行不可逆操纵时）所需的能量下界是 $k_B T \ln 2$。

为了进行下文中的研究，我们先将自由能推广到非平衡的情形下，此时，它是概率密度的泛函：

$$f^{neq}[p_x] = \epsilon - Ts = \int \epsilon_x p_x \mathrm{d}x + T \int p_x \ln p_x \mathrm{d}x$$

我们检查非平衡情形下的自由能和平衡情形下自由能的差距：

$$\begin{aligned}
\Delta f^{neq} &= f^{neq} - f^{eq} \\
&= \int \epsilon_x p_x \mathrm{d}x + T \int p_x \ln p_x \mathrm{d}x - f^{eq} \\
&= \int (-T \ln p_x^{eq} + f^{eq}) p_x \mathrm{d}x + T \int p_x \ln p_x \mathrm{d}x - f^{eq} \\
&= T \int (\ln p_x - \ln p_x^{eq}) p_x \mathrm{d}x \\
&= T D_{KL}(p_x \| p_x^{eq})
\end{aligned}$$

Decomposition of Mutual Information

下面考虑怎么建模前面的思想实验（Szilard Engine）, 我们将其建模为要被观测的物体（Obj）和观测设备 (Dev)，Obj 是一个随机变量 $x \in \{L, R\}$，Dev 同样是 $y \in \{L, R\}$。初始状态下系统的熵显然是 $s^{sys} = k_B \ln 4 = 2k_B \ln 2$，而在完成测量之后，系统的状态只能从 (L, L) 和 (R, R) 中任选其一，所以系统的熵减少了 $\ln 2$。为了不违反热力学第二定律有两种做法，要么增加另一个系统（例如存储介质）的熵，要么对系统做功 $k_{BT} \ln 2$ 并使得这些功完全被耗散掉。

下面考虑一般的情形。物体的状态以 x 标记，设备的状态以 y 标记，且 $\epsilon_y = 0$ 恒成立。测量的准确性由物体和设备之间的互信息表示：

$$I(obj; dev) = \sum_{xy} p_{xy} \frac{p_{xy}}{p_x p_y}$$

设测量过程中，物体的边缘概率 $p_x = \sum_y p_{xy}$ 不变，在 t_m 时刻测量完成，以及设测量开始时，设备和物体的状态是完全独立的，以至于互信息 $I(obj; dev) = 0$。下面考察由物体和设备组成的系统的熵变，计算：

$$\begin{aligned}
\Delta s^{sys} + k_B \Delta I = &-k_B \sum_{xy} p_{xy}(t_m) \ln p_{xy}(t_m) + k_B \sum_{xy} p_{xy}(t_0) \ln p_{xy}(t_0) \\
&k_B \sum_{xy} p_{xy} \ln p_{xy}(t_m) - k_B \sum_{xy} p_{xy}(t_m) \ln(p_x(t_m) p_y(t_m)) \\
&= k_B \sum_{xy} p_{xy}(t_0) \ln p_{xy}(t_0) - k_B \sum_{xy} p_{xy}(t_m) \ln(p_x(t_m) p_y(t_m)) \\
&= k_B \sum_{xy} p_{xy}(t_0) \ln(p_x(t_0) p_y(t_0)) - k_B \sum_{xy} p_{xy}(t_m) \ln(p_x(t_m) p_y(t_m)) \\
&= k_B \sum_{y} p_y(t_0) \ln p_y(t_0) - k_B \sum_{y} p_y(t_m) \ln p_y(t_m) \\
&= +\Delta s^{dev}
\end{aligned}$$

所以总结下来，我们得到了：

$$\Delta s^{sys} = -k_B I + \Delta s^{dev}$$

这意味着系统的熵变可以被分解为两部分：设备的熵变和二者间互信息的变化。有两个特殊情况：若 $\Delta s^{dev}=0$，则所有输入的功都被用于提高物体和设备间的互信息；若 $\Delta s^{sys}=0$ ，则 $k_B I=+\Delta s^{dev}$，这意味着设备自身的熵增被用于提高设备和系统间的互信息。

The Sagawa-Ueda Relation

下面我们考虑一个反馈控制的情形，根据测量到的状态 y 来调整控制策略，$\lambda=\lambda(t,y)$。定义一条轨迹上的随机互信息：

$$i_{xy}=\ln\frac{p_{xy}(t)}{p_x(t)p_y(t)}$$

考虑在 t_m 时刻测量一次，这个测量也是有随机性的，条件概率记为 $p_{y|x}(t_m)$。从熵产生的表达式开始（注意此时我们没有把测量设备建模到系统中）：

$$\frac{-s^{tot}}{k_B}=\ln\left(\frac{P_{\hat{x}}[\hat{\lambda}]}{P_x[\lambda]}\right)$$

两边同时减一个互信息：

$$-\frac{s^{tot}}{k_B}-i_{xy}(t_m)=\ln\left(\frac{P_{\hat{x}}[\hat{\lambda}]}{P_x[\lambda]}\right)-\ln\frac{p_{y|x}(t_m)}{p_y(t)}$$

考虑以下平均：

$$\begin{aligned}\int\mathcal{D}x\mathrm{d}y\;\;p_{y|x}(t_m)P_x[\lambda]\frac{P_{\hat{x}}[\hat{\lambda}]}{P_x[\lambda]}\frac{p_y(t_m)}{p_{y|x}(t_m)}&=\int\mathcal{D}xP_{\hat{x}}[\hat{\lambda}]\int\mathrm{d}yp_y(t_m)\\&=\int\mathcal{D}\hat{x}P_{\hat{x}}[\hat{\lambda}]\int\mathrm{d}yp_y(t_m)=1\end{aligned}$$

从而得到 Sagawa-Ueda Relation：

$$\left\langle\exp\left(-\frac{s^{tot}(x)}{k_B}-i_{xy}(t_m)\right)\right\rangle=1$$

特别地，如果系统初态、末态都在平衡态上，那么 $s^{tot}=w-\Delta F$。所以通过测量，我们可以从系统中提取比自由能变更多的功！

The Mandal-Jarzynski Machine

下面考虑一种热机，这种热机只有 A,B,C 三态，一个读写头指向无限延长的纸带。机器在 A,C 态之间跃迁时会导致热机对外做功或者外界对热机做功，具体而言：从 A 态移动到 C 态时，与热机连接的重物下降一段距离（外界对系统做功）；反之，从 C 态移动到 A 态时，重物上升一段距离（系统对外界做功）。A,C 间的跃迁需要考虑与纸带上状态的耦合：我们允许 $(A,1)\to(C,0)$ 的跃迁，同时它的逆过程 $(C,0)\to(A,1)$ 也必须存在。这种耦合使得热机完成一轮 $A\to B\to C\to A$ 的过程时必须消耗一个 1，否则繁殖。 纸带上的每个比特在未与读写头接触时以 r 的概率为 1，纸带每隔时间 τ 前移一格。

写这个系统的动力学时，我们通常考虑 6 个状态。简便起见，定义 $\omega=\dfrac{W}{k_BT}$，系统的主方程为：

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}p_{A0}=p_{B0}-p_{A0}\\&\frac{\mathrm{d}}{\mathrm{d}t}p_{B0}=p_{A0}+p_{C0}-2p_{B0}\\&\frac{\mathrm{d}}{\mathrm{d}t}p_{C0}=p_{B0}-p_{C0}+\exp\left(\frac{\omega}{2}\right)p_{A1}-\exp\left(-\frac{\omega}{2}\right)p_{C0}\\&\frac{\mathrm{d}}{\mathrm{d}t}p_{A1}=p_{B1}-P_{A1}+\exp\left(-\frac{\omega}{2}\right)p_{C0}-\exp\left(\frac{\omega}{2}\right)P_{A1}\\&\frac{\mathrm{d}}{\mathrm{d}t}p_{B1}=p_{A1}+p_{c1}-2p_{B1}\\&\frac{\mathrm{d}}{\mathrm{d}t}p_{C1}=p_{B1}-p_{C1}\end{aligned}$$

计算的时候，需要先以 τ 为一个时间单位计算以上主方程的解，然后求 A, B, C 三态的边缘概率，再将系统与新的比特接触，重新赋予六个态上的概率。这个机器其实是麦克斯韦妖在马尔可夫动力学下的实现，在 ω 和 r 的不同取值下，它将从三个功能中选择其一：

- 信息擦除器：外界对热机做功，同时纸条熵减（类似于制冷机，从低温热源抽走熵）
- 信息引擎：热机对外界做功，同时纸条熵增（类似正常热机，向低温热源排放熵）
- 无用状态：外界对热机做功，同时纸条熵增

Copying Information

生物体内通常存在对于信息的精确复制行为，这种行为没有对信息进行不可逆的操作，因而不受到兰道尔原理的限制。考虑一个简单的描述 DNA 复制或者 RNA 转录过程的模型：有一个模板链和一个正在复制的链，你可以在复制链的一端加入正确或错误的单体，设加入 $(+)$ /移除 $(-)$ 正确 (r) /错误 (w) 单体的速率是 $k_{r/w}^{(+/-)}$。设定一种复制机制：将单体结合到链上的过程是有化学驱动的，但是摘下来的过程没有。这导致：

$$k_r^+ = w_r \exp\left(\frac{\delta}{k_BT}\right),\ k_r^- = w_r \exp\left(\frac{\Delta\epsilon_r}{k_BT}\right),\ k_w^+ = w_w \exp\left(\frac{\delta}{k_BT}\right),\ k_w^- = w_w \exp\left(\frac{\Delta\epsilon_w}{k_BT}\right),$$

为了保证复制的正确性，我们通常要求 $\Delta\epsilon_w \geq \Delta\epsilon_r$。不难看出系统的状态空间是无限大的，所以你会得到递归的主方程：

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}p_{\cdots r} &= k_r^+ p_{\cdots} + k_r^- p_{\cdots rr} + k_w^- p_{\cdots rw} - (k_r^+ + k_w^+ + k_r^-)p_{\cdots r}\\ \frac{\mathrm{d}}{\mathrm{d}t}p_{\cdots w} &= k_w^+ p_{\cdots} + k_r^- p_{\cdots wr} + k_w^- p_{\cdots ww} - (k_r^+ + k_w^+ + k_w^-)p_{\cdots w}\end{aligned}$$

我们一般无法处理这种递归的主方程，它也往往没有稳态分布（我们暂时并没有给模板链一个长度的上限）。因为每个单体的附着是不依赖于前面的单体的，所以不妨假设链很长时，错误比例是 η，从而一条链出现的概率是：

$$p_{\cdots} \propto (1-\eta)^{N_r} \cdot \eta^{N_w}$$

计算链延伸一个正确单体或者一个错误单体的净速率：

$$v_r = k_r^+ - k_r^-(1-\eta),\ \ v_w = k_w^+ - k_w^-\eta$$

错误单体所占的比例应该等于错误单体延伸的净速率：

$$\eta = \frac{v_w}{v_r + v_w}$$

这样我们就得到了关于 η 的一元二次方程。可以给出系统的临界状态，即：

$$v_r = 0,\ v_w = 0 \Rightarrow \frac{k_r^+}{k_r^-} = 1-\eta,\ \frac{k_w^+}{k_w^-} = \eta \Rightarrow \frac{k_r^+}{k_r^-} + \frac{k_w^+}{k_w^-} = 1$$

由此解出能维持链复制的最小化学驱动：

$$\delta_{stall} = -k_BT\ln\left(\exp\left(-\frac{\Delta\epsilon_r}{k_BT}\right) + \exp\left(-\frac{\Delta\epsilon_w}{k_BT}\right)\right)$$

以及此时的平稳分布：

$$\eta^{eq} = \frac{\exp\left(-\frac{\Delta\epsilon_w}{k_BT}\right)}{\left(\exp\left(-\frac{\Delta\epsilon_r}{k_BT}\right) + \exp\left(-\frac{\Delta\epsilon_w}{k_BT}\right)\right)}$$

此外，系统的链延长的速率显然是：

$$v = k_r^+ + k_w^+ - (1-\eta)k_r^- - \eta k_w^-$$

下面我们试着做出熵产生速率。如果我把这个系统的状态空间建模成二叉树，且每个节点 $\cdots$ 的左子节点记为 $\cdots r$，右子节点记为 $\cdots w$，那么所有从当前节点走向左子节点的速率之和是 v_r，所有从当前节点走向右子节点的速率之和是 v_w，那么（总）

熵产生速率是：

$$T\dot{s}^{tot}=k_BT\left(v_r\ln\frac{k_r^+}{k_r^-(1-\eta)}+v_w\ln\frac{k_w^+}{k_w^-\eta}\right)$$

Information Cost in Sensing

接下来考虑一个对设备的建模。设设备要跟踪一个信号 $e\in\{0,1\}$，设备的状态被 $a,m\in\{0,1\}$ 表示，但是 m 的弛豫时间远长于 a，或者说 m 在两个态间的跳转速度慢于 a。我们要求在 m 完全跟踪 e 时，无论 a 是什么样子，系统的能量都不变；在 m 与 e 不同时，如果 a 成功跟踪了 e，那么会获得一点奖励（让能量降低）。构造如下模型：

$$\epsilon_{a,m,e}=|e-m|(\Delta_m+|a-e|\Delta_a)\quad\Delta_{a/m}=-k_BT\ln\delta_{a/m}$$

简化起见，我们设 a,m 的跳转不会同时发生，系统有四个跳转速率：

$$k_{(-a,m)\to(a,m)}=w_a\exp\left(\frac{\epsilon_{-a,m,e}-\epsilon_{a,m,e}}{2k_BT}\right)$$
$$k_{(a,-m)\to(a,m)}=w_m\exp\left(\frac{\epsilon_{a,-m,e}-\epsilon_{a,m,e}}{2k_BT}\right)$$

根据假设，需要 $w_a\gg w_m$。通过这样的跳转速率设置，我们显然实现了平衡分布：

$$p_{am|e}\propto\exp\left(-\frac{\epsilon_{a,m,e}}{k_BT}\right)$$

仍然使用设备和信号的互信息来说明设备是不是完全跟随了信号：

$$\begin{aligned}I(dev;env)&=\sum_{a,m,e}p_{a,m,e}\ln\frac{p_{a,m,e}}{p_{am}p_e}\\&=\sum_{a,m,e}p_{am|e}p_e\ln\frac{p_{am|e}}{p_{am}}\\&=\sum_{a,m,e}p_{m|e}p_e\ln\frac{p_{m|e}}{p_m}+\sum_{a,m,e}p_{a,m,e}\ln\frac{p_{a|me}}{p_{a|m}}\\&=I(m;e)+I(a;e|m)\end{aligned}$$

这里第一项是通过读取慢速记忆 m 获得的关于环境的信息，第二项是在已知记忆 m 的情况下，读取快速响应 a 获取的关于环境的信息。这种模型可以用于模拟生物的感知过程。
考虑如下情景：$t=0$ 时，e 从 $\{0,1\}$ 中随机取某值（由于系统的能量完全是根据 m,a 与 e 是否匹配决定的，不失一般性，其实可以直接设 e 初始取 0），系统已经处于平衡态，在时刻 t_f，系统状态以 $\frac{1}{2}$ 概率反转。令 $x=(m,a)$，定义如下两个量：

$$I(dev(t);env(t_f))=\sum_{x,e_i,e_f}=p_{x|e_i,e_f}(t)p_{e_i,e_f}\ln\frac{p_{x|e_f}(t)p_{e_f}}{p_x(t)p_{e_f}}$$

是 t 时刻设备和环境末态的互信息（注意这里 e_i 其实根本没用，可以直接求和掉）；

$$\begin{aligned}I(dev(t);env(t_0)|env(t_f))&=\sum_{x,e_i,e_f}p_{x|e_i,e_f}(t)p_{e_i,e_f}\ln\frac{p_{x,e_i|e_f}(t)}{p_{x|e_f}(t)p_{e_i|e_f}}\\&=\sum_{x,e_i,e_f}p_{x|e_i,e_f}(t)p_{e_i,e_f}\ln\frac{p_{x|e_ie_f}(t)}{p_{x|e_f}(t)}\end{aligned}$$

是给定末态后设备和初态的互信息。设备和环境末态的互信息的变化被用于衡量“设备测量了多少环境”：

$$\Delta I^{meas}(t)=I(dev(t);env(t_f))-I(dev(t_0);env(t_f))$$

而给定末态后设备与初态的互信息的变化则被用于衡量“系统从自己的内存中擦出了多少信息”：

$$\Delta I^{eras}(t)=I(dev(t_0);env(t_0)|env(t_f))-I(dev(t);env(t_0)|env(t_f))$$

在状态切换的一刻，外界把这个系统从 ϵ_{x,e_i} 拉到了另一个能级 ϵ_{x,e_f} 上，因此这一瞬间外界对系统所做之功：

$$\langle w\rangle = \sum_{e_i,e_f} = \left\langle \epsilon_{x,e_f} - \epsilon_{x,e_i}\right\rangle [p^{eq}_{x,e_i}] p_{e_i,e_f} \geq 0$$

现在看整个系统（设备和与这个设备接触的热库）的熵增。简单起见，记 $H(p) = -p\ln p$，考虑 e_i, e_f 给定，设备自身的熵增：

$$\Delta s^{dev}_{e_i,e_f}(t) = k_B(H(p_{x|e_i,e_f}(t)) - H(p_{x|e_i,e_f}(t_0)))$$

至于对外放热，完全是在环境的状态 e 发生跳变之后，在系统弛豫过程中才会进行：

$$\Delta s^{res}_{e_i,e_f}(t) = \frac{1}{T}(\left\langle \epsilon_{x,e_f}\right\rangle [p^{eq}_{x|e_i}] - \left\langle \epsilon_{x,e_f}\right\rangle [p^{eq}_{x|e_i,e_f}(t)])$$

特别地，对于设备的熵产生，我们可以推出一个分解关系：利用互信息满足的恒等式：

$$I(dev(t); e_i, e_f) = H(dev(t)) - H(dev(t)|e_i, e_f)$$

以及：

$$I(dev(t); e_i, e_f) = I(dev(t); e_f) + I(dev(t); e_i|e_f)$$

推出分解关系是：

$$\begin{aligned}\Delta s^{dev}_{e_i,e_f}(t) &= \Delta s^{dev}(t) - k_B\Delta I(dev(t); e_f) - k_B\Delta I(dev(t); e_i|e_f) \\ &= \Delta s^{dev}(t) - k_B\Delta I^{meas}(t) + k_B\Delta I^{eras}(t)\end{aligned}$$

也就是说设备自身的熵增（不考虑环境信息）可以看作设备的无条件熵增（考虑环境信息）减去观测获得的互信息加上擦除的互信息。换言之，观测减小了设备的混乱程度，而擦除反之。

我们考虑系统演化到时间无穷远处的情形，根据能级的对称性不难发现，设备从初态的平衡到末态的平衡过程中，自身的熵是不变的，所以热库要熵增，设备对热库必定放了正的热。根据热力学第一定律和自由能的定义，可以给出一个不等式：

$$Q = \langle w\rangle - \langle \Delta\epsilon\rangle \geq 0,\ \ \Delta F = \langle \Delta\epsilon\rangle - T\langle \Delta s\rangle \Rightarrow \langle w\rangle - \Delta F \geq T\langle \Delta s\rangle$$

利用系统能级的对称性还可以给出另外一个不等式：考虑刚刚完成状态跳转的那一刻的非平衡自由能：

$$\begin{aligned}F_{neq}(t_f^+) &= \langle w\rangle + U_{eq} - k_BT[H(dev(t_f)|e_f)] \\ &= \langle w\rangle + U_{eq} - k_BT[H(dev(t_f)) - I(dev(t_f)|e_f)]\end{aligned}$$

注意我们这里定义系统自由能时使用的是取定环境状态的条件熵，这样你才是只计算了系统本身的混乱程度，否则会将环境的影响也考虑进来。计算初态时的平衡自由能：

$$F_{eq}(t_0) = U_{eq} - k_BT\cdot H(dev(t_f)|e_i)$$

换掉上面的 U_{eq}：

$$\left\langle F_{neq}(t_f^+)\right\rangle = \langle w\rangle + F_{eq}(t_0) - k_BTI(dev(t_0); e_i) + k_BTI(dev(t_f); e_f)$$

在从 t_f^+ 向着无穷远弛豫的过程中，系统的自由能必然要下降，能级的对称性说明 $F_{eq}(+\infty) = F_{eq}(t_0)$，从而：

$$\langle w\rangle \geq k_BT(I(dev(t_0); e_i) - I(dev(t_f); e_f))$$

由于能级的对称性，末态的平衡互信息一定等于初态的平衡互信息，也就是说 $I(dev(t_0); e_i) = I(dev(+\infty); e_f)$；而刚刚到 t_f 时刻的时候，系统的状态还没来得及改变，所以 $I(dev(t_f); e_f) = I(dev(t_0); e_f)$。那么我们得到：

$$\langle w\rangle \geq k_BT\Delta I^{meas}(+\infty)$$

也就是说输入的功中的一部分用于提升设备与环境的互信息。

Information Reservoirs

下面我们对信息库做一个统一建模。简单起见只考虑两态系统，$\epsilon_d = 0, \epsilon_u = \epsilon$，初始时热浴在温度为 T 的恒温热库中，已经达到平衡态。现在使得系统还和一个信息库交互，交互速率为 γ：信息库中装了无限多的态（可以被视作一条无穷长的纸带），

其中 u 态的比例为 r，d 态为 $1-r$，当系统与信息库交互时，将从信息库中抽取一个态，跳到这个态上，而信息库中的态将被跃迁之前的态覆盖。从而由信息库引起的跳转速率是：

$$\bar{k}^+ = \gamma r,\ \ \bar{k}^- = \gamma(1-r)$$

系统的主方程是：

$$\frac{\mathrm{d}}{\mathrm{d}} p_u = (k^+ + \bar{k}^+) p_d - (k^- + \bar{k}^-) p_u$$
$$\frac{\mathrm{d}}{\mathrm{d}} p_d = (k^- + \bar{k}^-) p_u - (k^+ + \bar{k}^+) p_d$$

其中 $k^{+/-}$ 是热库引起的跃迁速率。记 $k = k^+ + k^-$，容易求出引入信息库后的平稳分布：

$$p_u^{st} = \frac{k^+ + \bar{k}^+}{k+\gamma},\ \ p_d^{st} = \frac{k^- + \bar{k}^-}{k+\gamma}$$

我们给出一些可以使用信息库研究的模型：考虑反馈控制系统，系统的状态是 $x \in \{0,1\}$，一个设备以 r 的速率测量系统状态，测量结果为 y，但是有 r 的概率测错。此时，上面信息库的状态 u, d 被用于表示 y 和 x 间的误差（设 d 是准确态，而 u 是误差态）。当然你可以直接使用信息库模型研究之前的 MJ 热机。

下面计算熵产生：两个概率流：

$$j^{th} = k^+ p_d - k^- p_u \quad j^{info} = \bar{k}^+ p_d - \bar{k}^- p_u$$

系统自身熵增：

$$\dot{s}^{sys} = (j^{th} + j^{info}) \ln \frac{p^d}{p^u}$$

热库熵增（注意：只有与热库交互引起的跃迁才会从热库吸放热，引起热库熵变）：

$$\dot{s}^{h-res} = j^{th} \ln \frac{k^-}{k^+} = j^{th} \ln \frac{\epsilon}{k_B T}$$

信息库熵增：交互后，纸带上的 u 态的比例将发生改变：

$$r' = r - \frac{j^{info}}{\gamma}$$

信息库的熵增率是输出纸带的熵减输入纸带的熵：

$$\dot{s}^{i-res} = \gamma(s^{out} - s^{in}) \approx j^{info} \ln \frac{r}{1-r}$$

这里需要 $\dfrac{j^{info}}{\gamma} \ll 1$。不难证明：

$$\dot{s}^{tot} = (k^+ p_d - k^- p_u) \ln \frac{k^+ p_d}{k^- p_u} + (\bar{k}^+ p_d - \bar{k}^- p_u) \ln \frac{\bar{k}^+ p_d}{\bar{k}^- p_u} \geq 0$$

Fluctuation Relations with Information Reservoirs

下面我们导出与信息库相关的一些涨落关系。为此需要把上面的信息库模型复制一份：考虑系统有 A, B 两部分，每次与信息库交互的时候，从一部分跳到另一部分。系统的主方程：

$$\frac{\mathrm{d}}{\mathrm{d}t} p_{uA} = k^+ p_{dA} - k^- p_{uA} + \bar{k}^+ p_{dB} + \bar{k}^u p_{uB} - (\bar{k}^- + \bar{k}^u) p_{uA}$$
$$\frac{\mathrm{d}}{\mathrm{d}t} p_{uB} = k^+ p_{dB} - k^- p_{uB} + \bar{k}^+ p_{dA} + \bar{k}^u p_{uA} - (\bar{k}^- + \bar{k}^u) p_{uB}$$
$$\frac{\mathrm{d}}{\mathrm{d}t} p_{dA} = k^- p_{uA} - k^+ p_{dA} + \bar{k}^- p_{uB} + \bar{k}^d p_{dB} - (\bar{k}^+ + \bar{k}^d) p_{dA}$$
$$\frac{\mathrm{d}}{\mathrm{d}t} p_{dB} = k^- p_{uB} - k^+ p_{dB} + \bar{k}^- p_{uA} + \bar{k}^d p_{dA} - (\bar{k}^+ + \bar{k}^d) p_{dB}$$

所有的涨落关系的推导，都来自将正向、逆向路径与熵产生关联起来，从而给出以熵产生为核心的宏观热力学量的涨落。简便起见，下文中我们总是设置：

$$(\bar{k}^- + \bar{k}^u) = (\bar{k}^+ + \bar{k}^d) = \gamma$$

如果你再设置：

$$\bar{k}^+ = \gamma r,\ \bar{k}^- = \gamma(1-r)$$

那么你就回到了上一节中的系统（只不过多了一个复制的部分）。路径概率当然要拆成正过程、逆过程跃迁速率的比，对于热库部分，这些比例直接对应热产生/热库熵产生，现在的问题是信息库熵产生怎么和跃迁速率之比联系起来。首先，信息库中 u 态的比例 r 应当满足：

$$r^{in} = \frac{\bar{k}^+ + \bar{k}^u}{\bar{k}^+ + \bar{k}^u + \bar{k}^- + \bar{k}^d} = \frac{\bar{k}^+ + \bar{k}^u}{2\gamma}$$

不妨设 R_{ud} 是由于和信息库交互产生的，从 d 态跳跃到 u 态的净概率流，那么 $\mathrm{d}t$ 时间内有 $R_{ud}\mathrm{d}t$ 个 Bit 被从 u 态反转到 d 态。而 $\mathrm{d}t$ 时间内，系统从 A 跳向 B 以及从 B 跳向 A 的次数之和（也就是 $\mathrm{d}t$ 时间内和信息库的交互次数，也就是信息库吐出的新比特数）是：

$$v^{info}\mathrm{d}t = ((\bar{k}^- + \bar{k}^u)(p_{uA} + p_{uB}) + (\bar{k}^+ + \bar{k}^d)(p_{dA} + p_{dB}))\mathrm{d}t$$

也就是说，输出的比特流的 u 态比例相比于输入比特流将发生变化，变化幅度是：

$$r^{out} = r^{in} - \frac{R_{ud}}{v^{info}}$$

设这个变化幅度很小，那么单位时间内的熵产生率：

$$s^{out} - s^{in} = (-\ln r^{in} - 1 + \ln(1-r^{in}) + 1)\frac{R_{ud}}{v^{info}} \cdot v^{info} = \ln\frac{r^{in}}{1-r^{in}}R_{ud} = \ln\frac{\bar{k}^+ + \bar{k}^u}{\bar{k}^- + \bar{k}^d}R_{ud}$$

后面我们会以 $R_{xx'}$ 标记跳转速率，并且对所有 $x \neq x'$ 求和，所以会有一个 $\frac{1}{2}$ 系数。所以信息库的熵产生并不是和正、逆向跳率 $\ln\frac{\bar{k}^+}{\bar{k}^-}$ 绑定在一起的。

当然，仍可形式化地写下涨落定理：

$$\frac{P_x}{P_{\hat{x}}} = \frac{p_{x_0}(t_0)}{p_{x_f}(t_f)}\exp\left(\frac{1}{k_B}s^{res}[x]\right)$$

$$s^{res} = \frac{1}{2}k_B\sum_{x\neq x'}\left(J_{xx'}\ln\frac{k_{xx'}}{k_{x'x}} + R_{xx'}\ln\frac{\bar{k}_{xx'}}{\hat{k}_{x'x}}\right)$$

其中的两个跃迁速率：

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

$$\begin{aligned}R_{xx'} = &\sum_{l=0}^{n}(\delta_{xA,x_{l+1}\alpha_{l+1}}\delta_{x'B,x_l\alpha_l} + \delta_{x'B,x_{l+1}\alpha_{l+1}}\delta_{xA,x_l\alpha_l})\\ &- (\delta_{x'A,x_{l+1}\alpha_{l+1}}\delta_{xB,x_l\alpha_l} + \delta_{xB,x_{l+1}\alpha_{l+1}}\delta_{x'A,x_l\alpha_l})\end{aligned}$$

我这样写的时候，实际上我已经引入了一个参考过程：该过程的与热库交互的跳率仍是 $k_{xx'}$，而与信息库交互的跳率换成了 $\hat{k}_{xx'}$。这里的 $P_{\hat{x}}$ 是使用参考过程算出来的，左侧的 s^{res} 也只是一个轨迹不可逆性的广义量度，不一定是有物理意义的熵产生。通过设置不同的参考过程，可以获得涨落定理的不同具体形式。例如，如果强制要求：

$$\hat{k}_{du} = \bar{k}^+\frac{1-r^{in}}{r^{in}},\ \ \hat{k}_{ud} = \hat{k}^-\frac{r^{in}}{1-r^{in}}$$

可以做出关于真正熵产生的涨落定理 $\left\langle \exp\left(-\frac{s^{tot}}{k_B}\right)\right\rangle = 1$。通过将逆向轨迹服从的真实过程替换为参考过程，我们可以在等式左侧制造各种各样的量（而不只是物理意义下的熵产生！），从而拓展涨落定理的研究范围。