退火过程

- 升温：随着温度的不断上升，粒子**逐渐脱离其平衡位置，变得越来越自由**，直到达到固体的溶解温度，粒子排列从原来的有序状态变为完全的无序状态。
- 退火：随着温度的下降，粒子的热运动逐渐减弱，粒子逐渐停留在不同的状态，其排列也从无序向有序方向发展，直至到温度很低时，重新以一定的结构排列。
- 粒子不同的排列结构，对应着不同的内能水平。如果退火过程是缓慢进行的，也就是说，温度的下降如果非常缓慢的话，使得在每个温度下，粒子的排列都达到一种平衡态，则当温度趋于绝对零度时，系统的内能将趋于最小值。

Metropolis准则

从状态i转换为状态j的转换准则：

如果$E(j) \leq E(i)$，则状态转换被接受；

如果$E(j) > E(i)$，则状态转移被接受的概率为：$\exp(\frac{E(i)-E(j)}{kT})$

Boltzmann分布律：$P_i(T) = \frac{e^{-\frac{E(i)}{kT}}}{Z_T}$

退火过程分析

1. 同一温度下两个内能不同的状态

假设$E(i) < E(j)$，那么$P_i(T) - P_j(T) > 0$。也即在任何温度下，系统**处于低内能状态的概率大于处于高内能状态的概率**。

2.高温下的情况

$$\lim_{T\to\infty} P_i(T) = \lim_{T\to\infty} \frac{\exp(-\frac{E(i)}{kT})}{\sum_{j\in S}\exp(-\frac{E(j)}{kT})} = \frac{1}{|S|} \tag{1}$$

（每一项都是1）

也就是说，高温情况下各个状态的概率是相等的。

3. 低温下的情况

$$
\begin{aligned}
\lim_{T\to 0} P_i(T) &= \lim_{T\to 0}\left[\frac{e^{-\frac{E(i)}{kT}}}{\sum_{j\in S} e^{-\frac{E(j)}{kT}}}\right] = \lim_{T\to 0}\left[\frac{e^{-\frac{E(i)-E_m}{kT}}}{\sum_{j\in S} e^{-\frac{E(j)-E_m}{kT}}}\right] \\
&= \lim_{T\to 0}\left[\frac{e^{-\frac{E(i)-E_m}{kT}}}{\sum_{j\in S_m} e^{-\frac{E(j)-E_m}{kT}} + \sum_{j\notin S_m} e^{-\frac{E(j)-E_m}{kT}}}\right] \\
&= \begin{cases} \frac{1}{|S_m|}, & i\in S_m \\ 0, & i\notin S_m \end{cases}
\end{aligned} \tag{2}
$$

S_m表示系统最小内能状态的集合，E_m表示系统的最小内能

即系统等概率趋近于几个内能最小的状态之一（系统达到内能最小状态的概率为1），而系统处于其他状态的概率为0。

4. 温度缓慢下降的情况

$$
\begin{aligned}
\frac{\partial P_i(T)}{\partial T} &= \frac{\partial}{\partial T}\left[\frac{e^{-\frac{E(i)}{kT}}}{Z_T}\right] \\
&= \frac{P_i(T)}{kT^2}[E(i) - \overline{E_T}] \\
&\begin{cases} >0, & E(i) > \overline{E_T}, \text{高能状态} \\ <0, & E(i) < \overline{E_T}, \text{低能状态} \end{cases} \\
\text{其中}\overline{E_T} &= \sum_{j\in S} E(j)P_j(T)\text{为状态内能的平均值}
\end{aligned} \tag{3}
$$

也就是说：系统处于低能状态的概率随着温度的下降单调上升，而系统处于高能状态的概率随着温度的下降单调下降。

从另一个角度说明了，当温度为绝对零度时，系统只能为内能最小的状态。

5. 退火过程的三个条件

初始温度必须足够高（各个状态都等概率）；

在每个温度下状态的交换必须足够充分；

温度T的下降必须足够缓慢。

组合优化问题与退火过程的类比

退火过程	**组合优化问题**
物理系统中的一个状态i	组合优化问题的解i
状态的内能$E(i)$	解的指标函数$f(i)$
内能最低状态	最优解
温度T	控制参数t
粒子的热运动	解在领域内的交换

模拟退火算法

- 基本思想：在局部搜索算法中，从领域中随机选择一个解j，按照Metropolis准则接受该解。

1，随机选择一个解 i，k=0，$t_0=T_{max}$（初始温度），计算指标函数 f(i)。
2，$t=t_k$，如果满足结束条件，则转（15）。
3，Begin
4，　　如果在该温度内达到了平衡条件，则转（13）。
5，　　Begin
6，　　　　从 i 的邻域 N(i)中随机选择一个解 j。
7，　　　　计算指标函数 f(j)。
8，　　　　如果 f(j)≤f(i)，则接受 j，i=j，f(i)=f(j)，转（4）。
9，　　　　否则计算

$$P_{i\Rightarrow j}(t)=e^{-\frac{f(j)-f(i)}{t}}$$

10，　　　　　如果 Random(0, 1)<$P_{i\Rightarrow j}(t)$，则接受 j，i=j，f(i)=f(j)。
11，　　　　转（4）
12，　　End
13，　　对 t 降温，t_{k+1}=Drop(t_k)，k=k+1，转（2）。
14，End
15，输出结果。
16，结束。

算法需要解决的几个问题：

(1) 初始温度t_0

(2) 温度t的衰减函数

(3) 算法的终止准则

(4) 每个温度t下的马尔可夫链长度（什么叫做达到平衡）

初始温度

- 基本原则：足够高的初始温度，以等概率处于任何一个状态

$$\exp(-\frac{\Delta f(i,j)}{t_0}) = P_0 \approx 1 \tag{4}$$

得：

$$t_0 = \frac{\Delta f(i,j)}{\ln(P_0^{-1})} \tag{5}$$

其中$\Delta f(i,j)$为状态j和状态i的指标函数差，可由随机产生的序列S计算：

$$\Delta f(i,j) = \max_{i \in S} f(i) - \min_{i \in S} f(i) \tag{6}$$

$$\Delta f(i,j) = \frac{\sum_{i,j \in S} |f(i) - f(j)|}{|S|^2} \tag{7}$$

$$\Delta f(i,j) = \frac{\sum_{i=0}^{|S|-1} |f(S(i)) - f(S(i+1))|}{|S|} \tag{8}$$

另一种选取方式：假设在t_0下随机的生成一个状态序列，分别用m_1和m_2表示指标函数下降的状态数和指标函数上升的状态数，$\overline{\Delta f(i,j)}$表示指标函数增加的平均值。则m_2个状态中，被接受的个数为：

$$m_2 \exp(-\frac{\overline{\Delta f(i,j)}}{t_0}) \tag{9}$$

则有平均接受概率：

$$P_0 = \frac{m_1 + m_2 \exp(-\frac{\overline{\Delta f(i,j)}}{t_0})}{m_1 + m_2} \tag{10}$$

令$P_0 \to 1$，反解$t_0 = \frac{\overline{\Delta f(i,j)}}{\ln\left(\frac{m_2}{m_2 P_0 - m_1(1-P_0)}\right)}$。

温度的下降方法

- 基本原则：下降的足够缓慢

1. 等比例下降

$$t_{k+1} = \alpha t_k,\ 0 < \alpha < 1 \tag{11}$$

2. 等值下降

$$t_{k+1} = t_k - \Delta t \tag{12}$$

每一温度下的停止准则

- 在每个温度下要有足够的交换次数

固定长度方法：在每一个温度下，都使用相同的L_k（交换次数）。一般可以将L_k选择为**关于问题规模n的一个多项式函数**。

算法的终止原则

- 基本原则：温度足够低

1. 零度法：$t < \epsilon$时结束

2. 循环总控制法：循环下降次数达到K时结束

3. 无变化控制法：在相邻的n个温度中得到的指标函数值无变化*（一般来说两个温度相等就可以了）*

TSP

解空间

- 10城市：$10! = 3628800$
- 20城市：$20! \approx 2.43 \times 10^{18}$，不可能穷举

指标函数

$$f(\pi_1, \ldots, \pi_n) = \sum_{i=1}^{n} d_{\pi_i \pi_{i+1}} \tag{13}$$
$$\pi_{n+1} = \pi_1$$

新解的产生

- 逆序交换法

当前解$(\pi_1, ..., \pi_u, \pi_{u+1}, ..., \pi_{v-1}, \pi_v, ..., \pi_n)$

交换后$((\pi_1, ..., \pi_u, \pi_{v-1}, ..., \pi_{u+1}, \pi_v, ..., \pi_n))$（相当于中间`reverse`了一次）

$$
\begin{aligned}
P_{i \to j}(t) = \exp(-\frac{f(j) - f(i)}{t}) = \exp(-\frac{\Delta f}{t}) \\
\Delta f = f(j) - f(i) = (d_{\pi_u \pi_{v-1}} + d_{\pi_{u+1} \pi_v}) - (d_{\pi_u \pi_{u+1}} + d_{\pi_{v-1} \pi_v})
\end{aligned}
\tag{14}
$$