

Optimization Problems: Simulated Annealing

An important type of problem is to find the “optimum” configuration of a system which can exist in a very large number of possible states. For example:

Annealing: A perfect defect-free crystal is much tougher than a crystal with lots of defects. Such a crystal can be produced by a carefully designed schedule of heating and cooling. Quantitatively, the perfect crystal is a global minimum energy configuration, whereas the configurations with defects correspond to local minima of the energy. The optimization problem here is to find the global minimum of a complicated function of all of the atomic coordinates!

Traveling Salesman: A salesman needs to travel to 532 U.S. cities to sell his wares. He wants to minimize his travel distance and he does not want to visit any city more than once. There are $\sim 10^{1217}$ possible routes he can take!

Simulated Annealing

This technique for finding a global extremum (minimum or maximum) of a function was introduced by S. Kirkpartick, C.D. Gelatt and M.P. Vecchi *Science* **220**, 671 (1983). It has been used in many applications, e.g., for designing integrated circuits with with millions of elements placed so as to minimize interference between their connecting wires.

The method is inspired by experimental observations on crystalization from a melt. At high temperatures, the atoms in the melt are free to move around the sample. As the temperature is reduced, the atoms tend to crystallize into a solid. If the sample is *quenched*, i.e., cooled very rapidly, then the solid is usually polycrystalline or amorphous in form. If the sample is *annealed*, i.e., cooled slowly, then the sample stands a better chance of forming a perfect crystal, which is the global minimum energy configuration of the system. Defects cost energy, and samples with defects correspond to local minima of the energy. Since there are an enormous number of possible configurations with defects, the *energy landscape* of the cooled solid is very complicated with numerous hills and valleys. Quenching typically leads to the bottom of the nearest valley, which annealing allows the system to explore the landscape and settle down into one of the lower valleys. The key to successful annealing is to use a good *annealing schedule*, i.e., a protocol for gradually reducing the temperature of the sample.

Simulated annealing in statistical physics is based on the concept of thermal equilibrium at temperature T . The probability of a configuration of the system is determined by the Boltzmann factor

$$\text{Prob}(E) \sim \exp\left(-\frac{E}{k_{\text{B}}T}\right),$$

where E is the energy of the configuration. At $T = 0$ the system is stuck at a minimum. At finite T , it can move around from one configuration to another, i.e., it can explore the nearby hills and valleys on the energy landscape.

To implement simulated annealing, we need a *dynamics* for the system to move from one configuration to another at a particular temperature T . A natural algorithm to

use is the Metropolis et al., algorithm: a random trial step is taken; if the step is downhill on the energy landscape (i.e., if the probability increases) the step is always accepted; and if the step is uphill, it is accepted conditionally.

The program by Silverman and Adler applies this technique to a 2-D system of atoms interacting by the Lennard-Jones potential

$$V(r) = 4\epsilon \left[\left(\frac{r}{\sigma} \right)^{12} - \left(\frac{r}{\sigma} \right)^6 \right].$$

It can be shown theoretically that the minimum energy configuration for this system is a triangular lattice—the closest packed configuration of disks in 2-D. (A square lattice is actually unstable and corresponds to a saddle point of the energy.) Local minima correspond to triangular lattices with defects.

With N particles, the total potential energy

$$\sum_{\text{pairs } i,j} V(|\mathbf{r}_i - \mathbf{r}_j|)$$

is a function of $2N$ coordinate variables. At $T = 0$ the atoms have zero kinetic energy and the system is located at some point on the landscape. At finite temperature, the system hovers at some point above the landscape on account of the kinetic energy of the particles. One could move the system around using Molecular Dynamics, but the Metropolis algorithm is simpler and more efficient.

Initial Configuration: Either a square lattice or a random initial configuration of atoms. Units are chosen so that the equilibrium separation in the Lennard-Jones

potential $2\frac{1}{6}\sigma = 1$: the average interparticle spacing should of course be ~ 1 to ensure that the system is close to being a solid.

Boundary Conditions: Periodic boundary conditions can be used to model a bulk system. Free boundary conditions are more appropriate to model clusters of atoms.

Effective Temperature: Silverman and Adler found that using an effective temperature for each atom

$$T_{\text{eff}} = zT ,$$

where z is the effective *coordination number*, i.e., the number of atoms within a distance $a \simeq 1.5-2$ yielded good results. They argue that when the particle makes a trial move, the change in energy ΔE will be roughly proportional to the number of neighbors, so the acceptance probability $\exp(-\Delta E/zk_{\text{B}}T)$ will be less dependent on the number of neighbors.

Annealing Schedule: The temperature is measured in units of ϵ/k_{B} . Recall that *one Monte Carlo step per particle* is conventionally taken to be N Metropolis steps with the particles being chosen either at random or in sequence. An annealing schedule is a sequence of temperatures T_i and Monte Carlo steps per particle n_i , $i = 0, 1, 2, \dots$. There is evidently a lot of freedom in choosing a schedule. Finding a good schedule for a particular problem is something of an art. Once found, it can be applied to solve similar problems.