Simulated Annealing

Dieter W. Heermann

Monte Carlo Methods

2015

- In the considered Hamiltonian there is frustration due to local free energy minima

$$H = -\sum_i (\epsilon_i \sum_j J_{i,i+1})) - \sum_{ij} K_{ij} \quad . \tag{1}$$

  Here ϵ_i is the energy of the ith residue, $J_{i,i+1}$ are nearest neighbour interactions and K_{ij} are short range interactions between residues.
- To reach the global minimum simulated annealing is one of the few known algorithms assuring convergence to a global minimum.
- It is often used in combination with efficient steepest descent methods, such as conjugate gradients, as a way for avoiding getting trapped in local minima.
- However much hinges on the choice of the annealing schedule.

$$\lim_{T\to 0} \pi_i(T) = \lim_{T\to 0} \frac{\exp\{-E_i/kT\}}{\sum_j \exp\{-E_j/kT\}} \tag{2}$$

$$= \lim_{T\to 0} \frac{\exp\left\{\frac{E^*-E_i}{kT}\right\}}{\sum_j \exp\left\{\frac{E^*-E_i}{kT}\right\}} \tag{3}$$

where,

$$E^* = \min_i E_i = \text{ Global min of energy} \tag{4}$$

- Thus, the exponents are always either zero or negative.

- In the limit when $T \to 0$ the terms with negative exponents disappear and we get

$$\lim_{T \to 0} \pi_i(T) = \begin{cases} \frac{1}{N^*} & \text{if } E_i = E^* \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where

$$N^* = |\{i : E_i = E^*\}| \quad . \tag{6}$$

- Thus, $\lim_{T \to 0} \pi_i(T)$ is uniformly distributed over the set of states of global minimum energy!

Simulated Annealing with Metropolis Monte Carlo

1: choose an initial configuration c
2: **for** $T = T_0, T_1, ..., T_m$ decreasing **do**
3: **for** mcs= 0; $i <$ mcsmax **do**
4: choose a new trial configuration c_t;
5: compute $W = \exp\{E(c)/T - E(c_t)/T\}$;
6: generate a random number R between 0 and 1
7: if $W > R$, accept the trial configuration as the new configuration of the system;
8: set $c = c_t$;
9: **end for**
10: **end for**

- Along the same line of though is constructed the *simulated tempering* [1] and the Metropolis-coupled Markov chain Monte Carlo [2].
- The basic idea is to use m different chains with distributions π_i $(i = 1, ..., m)$. From time to time we attempt to swap a state from chain i with one from chain j.

- E. Marinari and G. Parisi, Euro Phys. Lett **19**, 451-458 (1992)
- C.J. Geyer, in *Computing Science and Statistics: Proc. 23rd Symp Interface* (ed. E.M. Keramidas) pp. 156-163, Fairfax Station: Interface Foundation
- C.J. Geyer and E.A. Thomposon, J. Am. Statis. Ass **90**, 909-920 (1995)