Simulated Annealing (SA)

Annealing

A known thermal process for obtaining low energy state of a solid in a heat bath.

- This process consists of two steps:

  Step 1:
  Increase the temperature of the heat bath to a maximum value at which the solid melts.

  Step 2:
  Decrease the temperature carefully until the particles arrange themselves in a ground state of the solid.

- Liquid phase: High energy state
  Ground state solid: Energy of the system is minimized.

- Literature
  - 1953 - Metropolis et al. developed a computer simulation model of the physical annealing.

    Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, Equation of State Calculations by Fast Computing Machines, J. Chem. Phys., Vol. 21, No. 6, pp. 1087 - 1092, 1953.

  - 1983 - Kirkpatrick et al. adopted these ideas for solving discrete optimization problems in the context of a local search procedure.

    Kirkpatrick, S., C. D. Gelatt Jr, and M. P. Vecchi, Optimization by Simulated Annealing, Science, Vol. 220, No. 4598, pp. 671 - 680, 1983.

Overview of Simulated Annealing

Simulated annealing is a generalization of a Monte Carlo method for examining the equations of state and frozen states of n-body systems [Metropolis et al. 1953]. The concept is based on the manner in which liquids freeze or metals recrystalize in the process of annealing. In an annealing process a melt, initially at high temperature and disordered, is slowly cooled so that the system at any time is approximately in thermodynamic equilibrium. As cooling proceeds, the system becomes more ordered and approaches a "frozen" ground state at $T = 0$. Hence the process can be thought of as an adiabatic approach to the lowest energy state. If the initial temperature of the system is too low or cooling is done insufficiently slowly the system may become quenched forming defects or freezing out in metastable states (ie. trapped in a local minimum energy state).

The original Metropolis scheme was that an initial state of a thermodynamic system was chosen at energy E and temperature T, holding T constant the initial configuration is perturbed and the change in energy dE is computed. If the change in energy is negative the new configuration is accepted. If the change in energy is positive it is accepted

with a probability given by the Boltzmann factor $exp - (dE/T)$. This processes is then repeated sufficient times to give good sampling statistics for the current temperature, and then the temperature is decremented and the entire process repeated until a frozen state is achieved at $T = 0$.

By analogy the generalization of this Monte Carlo approach to combinatorial problems is straight forward [Kirkpatrick et al. 1983, Cerny 1985]. The current state of the thermodynamic system is analogous to the current solution to the combinatorial problem, the energy equation for the thermodynamic system is analogous to at the objective function, and ground state is analogous to the global minimum. The major difficulty (art) in implementation of the algorithm is that there is no obvious analogy for the temperature T with respect to a free parameter in the combinatorial problem. Furthermore, avoidance of entrainment in local minima (quenching) is dependent on the "annealing schedule", the choice of initial temperature, how many iterations are performed at each temperature, and how much the temperature is decremented at each step as cooling proceeds.

Basic Idea

- A stochastic single point search technique.

- avoiding entrapment in poor local optima by allowing an occasional uphill move.
  - guides the search by specifying a cooling scheme that allows the acceptance of randomly generated neighbor solutions which are relatively unfavorable as the current solution.

Basic Idea

Given a current state i of the solid with energy level E_i, generate a subsequent state j randomly (by small perturbation).

- Let E_j be the energy level at state j
- If $E_j - E_i \leq 0$, then accept state j as the current state.
- If $E_j - E_i > 0$, then accept state j with the probability $e^{-\frac{E_j - E_i}{k_B T}}$ where k_B is the Boltzman constant.

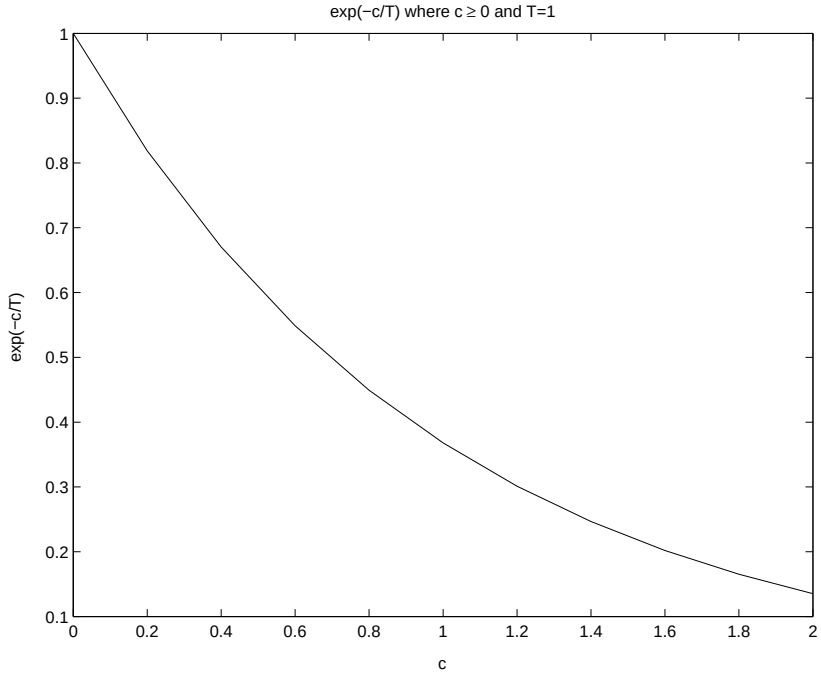


Figure 1: $e^{-\frac{c}{T}}$ vs. c when $T = 1$.

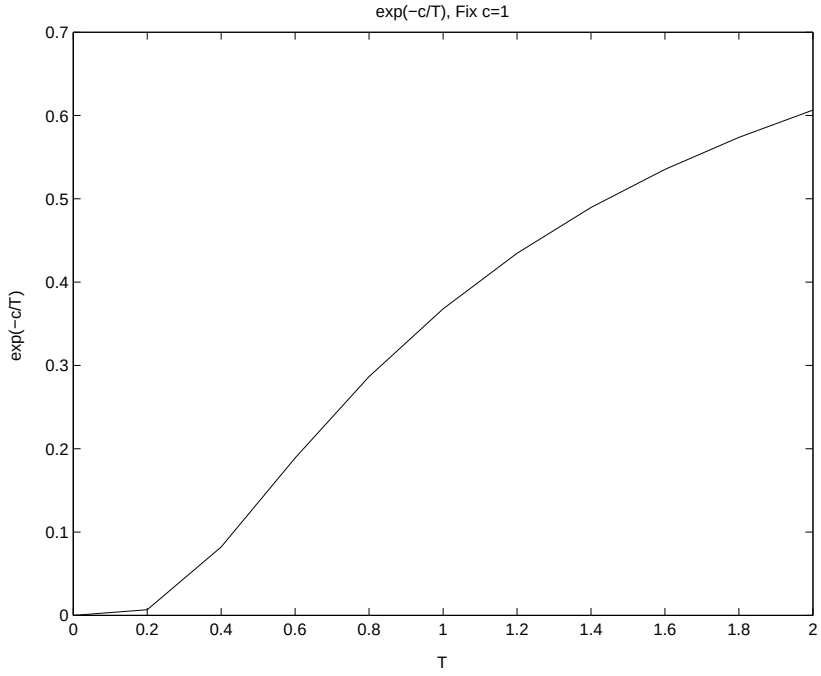


Figure 2: $e^{-\frac{c}{T}}$ vs. T when $c = 1$.

Analogy

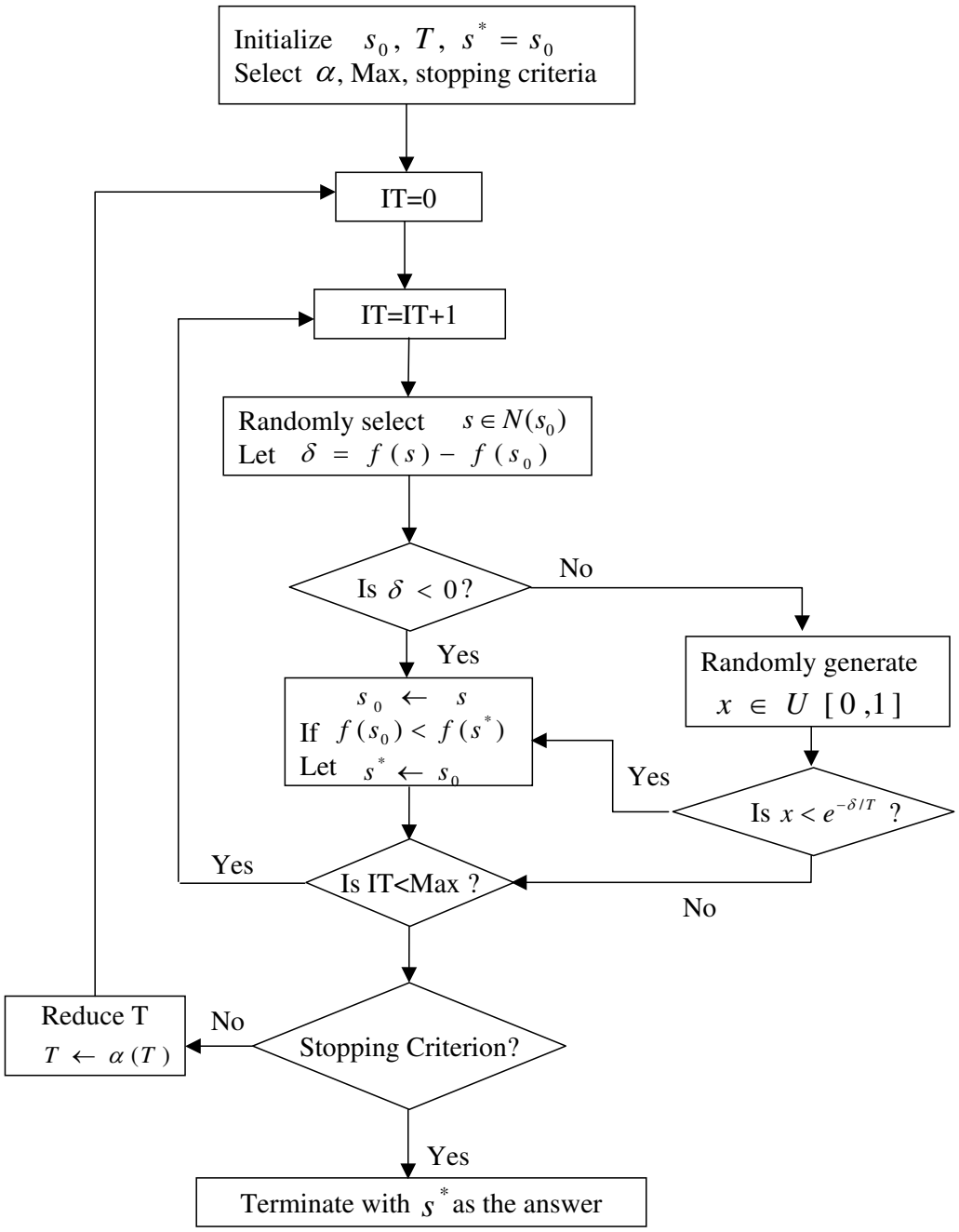
- Each solution corresponds to a state of the system.
- Cost corresponds to the energy level.
- Neighborhood corresponds to a set of subsequent states that the current state can reach.
- Control parameter corresponds to temperature.
- Optimal solution corresponds to ground state.

Simulated Annealing Algorithm

For a minimization problem with solution space S, objective function f and neighborhood structure N

Select an initial solution s_0 and $s^* = s_0$;
Select an initial temperature $T > 0$;
Select a temperature reduction function $\alpha(\cdot)$;
Select a maximum iteration count Max;
Repeat
  Set iteration count $IT = 0$
  Repeat
    $IT = IT + 1$
    Randomly select $s \in N(s_0)$
    set $\delta = f(s) - f(s_0)$;
    If $\delta < 0$
      then set $s_0 = s$ (downhill move)
    else
      generate random x uniformly in the range $[0, 1]$;
      If $x < \exp(-\delta/T)$
        then set $s_0 = s$ (uphill move)
    If $f(s_0) < f(s^*)$ then $s^* = s_0$.
  Until $IT =$ Max
  Set T=$\alpha(T)$;
Until stopping condition becomes true.
Output s^* as an approximation to the optimal solution.

Flow Chart of SA Algorithm



Convergence Theorem of SA

$$\text{(P)} \quad \begin{array}{l} \min f(x) \\ \\ \text{s.t.} \quad \text{x} \in \text{S}. \end{array}$$

Define $S_{opt} = \{x : x \in S \text{ and } f(x) \leq f(y), \forall y \in S\}$.

X : a stochastic variable denoting the current solution obtained by SA.

Corollary 2.1 [p.18 in Aarts and Korst]

Given an instance (S, f) of a combinatorial optimization problem, and a suitable neighborhood structure, after a sufficiently large number of transitions and a proper cooling scheme, we have

$$\lim_{t \to 0} P_t\{X = x\} = \begin{cases} \frac{1}{|S_{opt}|} & \text{if} \quad x \in S_{opt} \\ 0 & \text{otherwise} \end{cases}$$

Important Implementation Issues

- Speed of convergence
- Quality of solution
  - Problem-specific
    - Solution representation
    - Neighbors of a solution
    - Cost of a solution
    - Initial solution
  - Generic
    - Initial temperature (T)
    - Temperature reduction ($\alpha(\cdot)$)
    - Temperature length L (Max)
    - Stopping criteria

- Initial Temperature

  Must be high enough to allow almost free exchange of neighboring solutions.

  – Sometimes but rarely it can be determined theoretically.

  – Otherwise increase the temperature rapidly until a relatively large proportion of the neighbors are acceptable.

- Cooling Scheme

  – Max iteration at each temperature.

  Max: depends on the size of the NBHD

  - $k * |N|$ where k is size factor and N is the expected neighborhood size.

  - Temperature length will be proportional to the number of neighbors.

  – Temperature reduction function $\alpha(\cdot)$.

  Geometric: $\alpha(T) = \alpha * T$

- Stopping Criteria

  - Lundy and Meese

$$T \leq \frac{\epsilon}{\ln[(|s|-1)/\theta]}$$

    At equilibrium, the probability of being at a solution which is within ϵ of the optimal value is $1-\theta$.

  - Number of iterations without acceptance.

  - Total number of iterations.

Example

Maximum Cut Problem (MCP) [Aarts and Korst]

- Given an undirected graph $G = (V, E)$, with a weight d_{uv} associated with each $(u, v) \in E$.

- MCP: Find a partition of V into disjoint subsets A and B such that $A \cup B = V$ and the function $f(A, B)$ is maximized, where

$$f(A, B) = \sum_{u \in A, v \in B} d_{uv}$$

- Solution Space (representation)?

- Neighborhood?

  Given a solution $S = (A, B)$ a solution S' is in the neighborhood of S if....

  How to generate NBHD?

- Cost function?

References

1. Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, Equation of State Calculations by Fast Computing Machines, J. Chem. Phys., 21(6), 1087 - 1092, 1953.

2. Kirkpatrick, S., C. D. Gelatt Jr, and M. P. Vecchi, Optimization by Simulated Annealing, Science, 220(4598), 671 - 680, 1983.

3. Cerny, V., Thermodynamical Approach to the Traveling Salesman Problem: An Efficient Simulation Algorithm, J. Opt. Theory Appl., 45(1), 41-51, 1985.

4. Szu, H. and R. Hartley, Fast Simulated Annealing, Physics Letters A, 122(3), 157 - 162, 1987.

5. D. S. Johnson, C. R. Aragon, L. A. McGeoch, and C. Schevon. Optimization by Simulated Annealing: An Experimental Evaluation; Part I, Graph Partitioning. Operations Research, 37, 865-892, 1989.

6. Aarts, E. and J. Korst. Simulated Annealing and

Boltzmann Machines, Wiley, 1989.

7. Reeves, Colin R. Modern Heuristic Techniques for Combinatorial Optimization, Wiley, 1993.

8. McMullen, P. R. and G. V. Frozier, An SA Approach to Mixed Model Sequencing with Multiple Objectives, IIE Transactions, 32, 679-686, 2000.