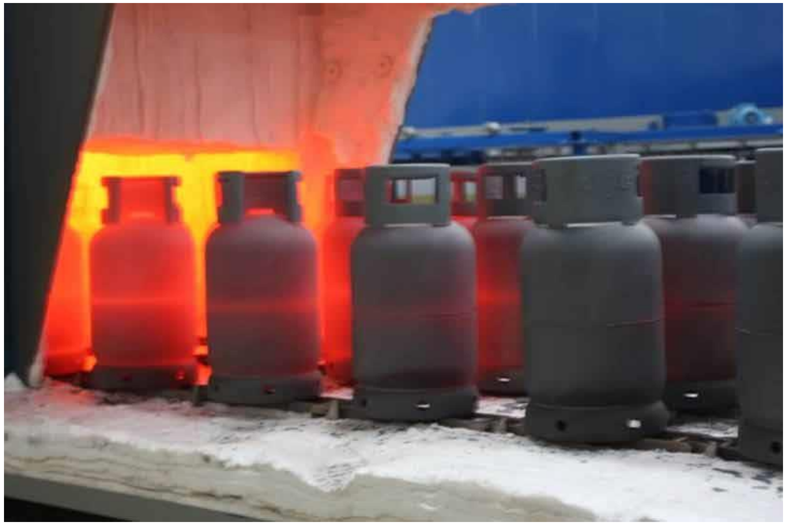
Advanced Search
Simulated annealing

Yingyu Liang
yliang@cs.wisc.edu
Computer Sciences Department
University of Wisconsin, Madison

[Based on slides from Jerry Zhu, Andrew Moore http://www.cs.cmu.edu/~awm/tutorials]



2. SIMULATED ANNEALING

Simulated Annealing

anneal

- To subject (glass or metal) to a process of heating and slow cooling in order to toughen and reduce brittleness.

Simulated Annealing

1. Pick initial state *s*
2. Randomly pick *t* in neighbors(*s*)
3. IF *f*(*t*) better THEN accept *s*🡐*t*.
4. ELSE /* *t* is worse than *s* */
5. accept *s*🡐*t* with a small probability
6. GOTO 2 until bored.

Simulated Annealing

1. Pick initial state *s*
2. Randomly pick *t* in neighbors(*s*)
3. IF *f*(*t*) better THEN accept *s*←*t*.
4. ELSE /* *t* is worse than *s* */
5. accept *s*←*t* with a small probability
6. GOTO 2 until bored.

What are the two key differences from Hill-Climbing?

Simulated Annealing

1. Pick initial state *s*
2. Randomly pick *t* in neighbors(*s*)
3. IF *f*(*t*) better THEN accept *s*🡐*t*.
4. ELSE /* *t* is worse than *s* */
5. accept *s*🡐*t* with a small probability
6. GOTO 2 until bored.

How to choose the small probability?

idea 1: $p = 0.1$

Simulated Annealing

1. Pick initial state *s*
2. Randomly pick *t* in neighbors(*s*)
3. IF *f*(*t*) better THEN accept *s*🡐*t*.
4. ELSE /* *t* is worse than *s* */
5. accept *s*🡐*t* with a small probability
6. GOTO 2 until bored.

How to choose the small probability?

idea 1: $p = 0.1$

What's the drawback?

Hint: consider the case when we are at the global optimum

Simulated Annealing

1. Pick initial state *s*
2. Randomly pick *t* in neighbors(*s*)
3. IF *f*(*t*) better THEN accept *s*←*t*.
4. ELSE /* *t* is worse than *s* */
5.   accept *s*←*t* with a small probability
6. GOTO 2 until bored.

How to choose the small probability?

idea 1: $p = 0.1$

idea 2: p decreases with time

Simulated Annealing

1. Pick initial state *s*
2. Randomly pick *t* in neighbors(*s*)
3. IF *f*(*t*) better THEN accept *s*←*t*.
4. ELSE /* *t* is worse than *s* */
5. accept *s*←*t* with a small probability
6. GOTO 2 until bored.

How to choose the small probability?

idea 1: $p = 0.1$

idea 2: *p* decreases with time

idea 3: *p* decreases with time, also as the 'badness' $|f(s)-f(t)|$ increases

Simulated Annealing

- If $f(t)$ better than $f(s)$, always accept t
- Otherwise, accept t with probability

$$\exp\left(-\frac{|f(s)-f(t)|}{Temp}\right)$$

Boltzmann distribution

Simulated Annealing

- If *f*(*t*) better than *f*(*s*), always accept *t*
- Otherwise, accept *t* with probability

$$\exp\left(-\frac{|f(s)-f(t)|}{Temp}\right)$$

Boltzmann distribution

- *Temp* is a temperature parameter that 'cools' (anneals) over time, e.g. $Temp \leftarrow Temp * 0.9$ which gives $Temp=(T_0)^{\#iteration}$

Simulated Annealing

- If $f(t)$ better than $f(s)$, always accept t
- Otherwise, accept t with probability

$$\exp\left(-\frac{| f(s) - f(t) |}{Temp} \right)$$

Boltzmann distribution

- *Temp* is a temperature parameter that 'cools' (anneals) over time, e.g. $Temp \leftarrow Temp*0.9$ which gives $Temp=(T_0)^{\#iteration}$
  - High temperature: almost always accept any *t*
  - Low temperature: first-choice hill climbing

Simulated Annealing

- If $f(t)$ better than $f(s)$, always accept t
- Otherwise, accept t with probability

$$\exp\left(-\frac{|f(s)-f(t)|}{Temp}\right)$$

Boltzmann distribution

- *Temp* is a temperature parameter that 'cools' (anneals) over time, e.g. $Temp \leftarrow Temp*0.9$ which gives $Temp=(T_0)^{\#iteration}$
  - High temperature: almost always accept any t
  - Low temperature: first-choice hill climbing
- If the 'badness' (formally known as energy difference) $|f(s)-f(t)|$ is large, the probability is small.

SA algorithm

// assuming we want to maximize f()

current = Initial-State(problem)

for t = 1 **to** ∞ **do**

T = Schedule(t) ; // T is the current temperature, which is monotonically decreasing with t

if T=0 **then return** current ; //halt when temperature = 0

next = Select-Random-Successor-State(current)

deltaE = f(next) - f(current) ; // If positive, next is better than current. Otherwise, next is worse than current.

if deltaE > 0 **then** current = next ; // always move to a better state

else current = next with probability p = exp(deltaE / T) ; // as T → 0, p → 0; as deltaE → -∞, p →0

end

Simulated Annealing issues

- Easy to implement.
- Intuitive: Proposed by **Metropolis** in 1953 based on the analogy that alloys manage to find a near global minimum energy state, when annealed slowly.

Simulated Annealing issues

- Cooling scheme important
- Neighborhood design is the real ingenuity, not the decision to use simulated annealing.
- Not much to say theoretically
  - With infinitely slow cooling rate, finds global optimum with probability 1.

- Try hill-climbing with random restarts first!